TECHNICAL RESEARCH REPORT

Optimal Output Feedback Control Using Two Remote Sensors over Erasure Channels

by Vijay Gupta, Nuno C Martins and John S Baras

TR 2007-7

Optimal Output Feedback Control Using Two Remote Sensors over Erasure Channels

Vijay Gupta, Nuno C. Martins and John S. Baras

Abstract

Consider a discrete-time, linear time-invariant plant, two sensors and one controller. The plant's state is observed in the presence of noise by the sensors, which are connected to the controller via links that feature erasure. If a link transmits successfully then a finite-dimensional vector of real numbers is conveyed from the sensor to the controller. If an erasure event occurs, then any information conveyed over the link is lost. This paper addresses the problem of designing the maps that specify the processing at the controller and at the sensors to minimize a quadratic cost function. When the information is lost over the links either in an independent and identically distributed (i.i.d.) or (time-homogeneous) Markovian fashion, we derive necessary and sufficient conditions for the existence of maps such that the plant is stabilized in the bounded second moment sense. We also solve the optimal design problem in the presence of delayed noiseless acknowledgment signals at the sensors from the controller for any arbitrary packet drop pattern. We provide explicit recursive schemes to implement our solution. We also indicate how our approach can be extended to situations when more than two sensors are available, when the sensors can cooperate and when the acknowledgment link features erasure. The analysis also carries over to the case when each point-to-point erasure link connecting the sensors and the controller is replaced by a network of erasure links.

I. Introduction

Recently, significant attention has been directed towards networked control systems in which components communicate over wireless links or communication networks that may also be used for transmitting other unrelated data (see, e.g., [2], [6], [21] and the references therein). The estimation and control performance in such systems is severely affected by the properties of the communication channels. Communication links introduce many potentially detrimental phenomena, such as quantization error, random delays, data loss and data corruption to name a few, that may lead to performance degradation or even stability loss.

In this work, we are specifically interested in the problem of estimation and control across communication links that exhibit data loss. We consider a dynamical process evolving in time that is being observed by two sensors. The sensors need to transmit the data over communication links to a remote node, which can either be an estimator or a controller. However information transmitted over the links is erased stochastically. Preliminary work in this area has largely concentrated on the case when only one sensor is present. Within the one-sensor framework, both

The authors are affiliated with the Department of Electrical and Computer Engineering and the Institute for Systems Research at the University of Maryland, College Park. Contact author: Vijay Gupta. Email address: vijay@cds.caltech.edu.

stability [38], [46] and performance [26], [38] have been analyzed. Approaches to compensate for the data loss to counteract the degradation in performance have also been proposed. As some representative examples, Hadjicostis and Touri [17] proposed applying zero control if sensor data is lost, Nilsson [32] proposed using the previous control input or time-updating the previous estimate in case of data loss, Ling and Lemmon [26] posed the problem as a filter-design through a non-linear optimization for scalar observations and Smith and Seiler [40] proposed a sub-optimal but computationally efficient estimator for packet drops occurring according to a Markov chain. Also relevant are the works of Azimi-Sadjadi [3], Schenato et al. [37] and Imer et al. [22] who looked at controller structures to minimize quadratic costs for systems in which both sensor-controller and controller-actuator channels exhibit erasure. The related problem of optimal estimation across an erasure link was considered by Sinopoli et al. in [39] for the case of one sensor and erasures occurring in an i.i.d. fashion, while Gupta et al. [14] considered multiple sensors and more general erasure models.

Most of the above designs aimed at designing a packet-loss compensator. The compensator accepts those packets that the link successfully transmits and propagates an estimate of the plant's state when data sent over the link is lost. If the estimator is used inside a control loop, the estimate is then used by the controller. We take a more general approach to the control of networked control systems. It has often been recognized that typical network / communication data packets have much more space for carrying information than required inside a traditional control loop. For instance, the minimum size of an ethernet data packet is 72 bytes, while a typical data point will only consume 2 bytes [11]. Many other examples are given in Lian et al. [25]. Moreover, many of the devices used in networked control systems have processing and memory capabilities on account of being equipped to communicate across wireless channels or networks. Thus the question arises if we can use this possibility of pre-processing information prior to transmission and transmission of extra data to combat the effects of packet delays, loss and so on and improve the performance of a networked control system. In Gupta et al. [16] it was shown that pre-processing (or encoding) information before transmission over the communication link can indeed yield significant improvements in terms of stability and performance. Moreover, for a given performance level, it can also lead to a reduced amount of communication. The benefits incurred become even more apparent when the communication link is replaced by a *network* of communication links [15]. This effect can also be seen in the recent works on receding horizon networked control, in which a few future control inputs are transmitted at every time step by the controller and buffered at the actuator to be used in case subsequent control updates are dropped by the network and do not arrive at the actuator(s), see, e.g., [12], [13], [23], [29], [30].

In this work, we extend this idea to the case when multiple sensors are present. Suppose a process is observed using two sensors that transmit the data over erasure links to a controller. If the sensors can share their measurements, there is effectively only one sensor. We look at the case when cooperation between the sensors is either not permitted, or occurs over erasure links. We obtain necessary and sufficient conditions of stabilizability in terms of the state-space representation of the plant and the probability of erasure at the links. We also give optimal performance achieving algorithms, under the assumption that the sensors have access to noiseless acknowledgment regarding the erasure process at the links connecting them to the controller. If acknowledgment signals are not present or if they are conveyed via erasure links then we give algorithms that, though not optimal, still perform better than the case when sensors transmit measurements without any processing.

The problem involving the presence of multiple sensors transmitting data in an aperiodic fashion is much more complicated than the problem involving only a single sensor. The problem of finding optimal encoding algorithms for the multi-sensor case and analyzing their performance is similar to the problems of fusion of data from multiple sensors and of track-to-track fusion that have long been open. A usual starting point for the works that address these problems is an attempt to decentralize the Kalman filter as, e.g., in [43]. However this approach requires that data about the global estimate be sent from the fusion node to the local sensors. This difficulty was overcome in [9], [41] and further in [18] where both the measurement and time update steps of the Kalman filter were decentralized. Alternative approaches for data fusion from many nodes include using the Federated filter [7], Bayesian methods [10], a scattering framework [24], algorithms based on decomposition of the information form of the Kalman filter [33] and so on.

However all these approaches assume a fixed communication topology among the nodes where a link, if present, is perfect (no erasure). In our case, information is lost randomly by the erasure links. This random loss of information reintroduces the problem of correlation between the estimation errors of various nodes [4] and renders the approaches proposed in the literature as sub-optimal. An approach to solve this problem was proposed in [5] in the context of track-to-track fusion through exchange of state estimates based on each sensor's own local measurements but the specific scheme that was used was not proved to be optimal. It was subsequently proved in [8] that the technique was based on an assumption that was not met in general. There are special cases for which the solution is known, e.g., when the process noise is absent [44] or when one of the sensors transmits data over a channel that does not erase information [16]. However, as stated earlier, in general, the problem is still open. Owing to a separation principle that we present, our results also carry over to the multi-sensor fusion problem.

The paper is organized as follows. We begin in the next section by describing the problem set-up and our notation. In Section III, we present a summary of the stabilizability results for the case when two sensors transmit data over erasure channels. In Section IV, we present a separation principle that allows us to consider an alternative estimation problem. For the case when the sensors have access to acknowledgments from the controller, we provide a recursive algorithm which is optimal with respect to every possible realization of the erasure process in Section V-A. Stability analysis of this algorithm allows us to prove the necessity of the stabilizability conditions in Section V-B. In Section V-C, we then prove that the conditions are sufficient as well, by presenting a sub-optimal algorithm that stabilizes the system even when acknowledgments from the controller are not available. Section VI generalizes the results in various directions. The case of more than two sensors being present is treated in Section VI-A. Section VI-B considers the case when the sensors transmit information to the controller over networks of erasure links. This also allows us to treat the case when the sensors can co-operate over erasure channels. In Section VI-C, we present an algorithm that can be used if the acknowledgments from the controller are also transmitted over an erasure channel. Although the algorithm is not optimal, it yields performance that is continuous with respect to the erasure probabilities. Finally, in Section VI-D we analyze the case when the channel erasures are correlated in time and can be described by a Markov chain. We finish with some possible directions for future work.

II. Framework Description and Problem Formulation

A. Modeling Assumptions

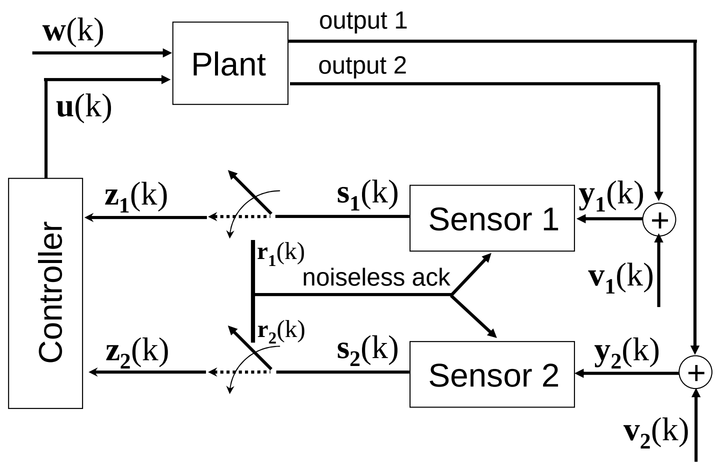

Fig. 1. Basic framework for output feedback using two remote sensors, in the presence of erasure channels. The process and measurement noises at the plant are represented by $\mathbf{w}(k)$ and $\mathbf{v}_{1 \text{ or } 2}(k)$, respectively. The erasure process, in the links connecting the sensors to the controller, is governed by $\mathbf{r}_{1 \text{ or } 2}(k)$.

Consider the set-up of Fig 1, and the following associated assumptions regarding the plant, the external sources of randomness and the erasure links that connect the sensors to the controller.

The plant is described by a discrete-time state-space representation of the following type:

$$\mathbf{x}(k+1) = A\mathbf{x}(k) + B\mathbf{u}(k) + \mathbf{w}(k), \;\; k \geq 1 \tag{1}$$

where $\mathbf{x}(k) \in \mathbb{R}^n$ is the process state, $\mathbf{u}(k) \in \mathbb{R}^l$ is the control input and $\mathbf{w}(k)$ is the process noise assumed to be white, Gaussian, zero mean with covariance $\mathbf{R}_w > 0$. The initial state $\mathbf{x}(0)$ is a zero mean and Gaussian random variable with covariance matrix $\mathbf{P}(0)$. The process state is observed using two sensors that generate measurements of the form

$$\mathbf{y}_1(k) = C_1\mathbf{x}(k) + \mathbf{v}_1(k), \;\; k \geq 0 \tag{2}$$

$$\mathbf{y}_2(k) = C_2\mathbf{x}(k) + \mathbf{v}_2(k), \;\; k \geq 0 \tag{3}$$

where $\mathbf{y}_1(k) \in \mathbb{R}^{m_1}$ and $\mathbf{y}_2(k) \in \mathbb{R}^{m_2}$. The measurement noises $\mathbf{v}_1(k)$ and $\mathbf{v}_2(k)$ are also assumed to be white, Gaussian, zero mean with positive definite covariance matrices $\Sigma_{v,1}$ and $\Sigma_{v,2}$ respectively. For ease of notation, we adopt the concatenation $\mathbf{v}(k)^T \overset{def}{=} [\mathbf{v}_1(k)^T \;\, \mathbf{v}_2(k)^T]^T$ and denote the covariance matrix of $\mathbf{v}(k)$ by $\mathbf{R}_v$. Similarly, we define

$$C = \begin{bmatrix} C_1 \\ C_2 \end{bmatrix} \qquad\qquad \mathbf{y}(k) = \begin{bmatrix} \mathbf{y}_1(k) \\ \mathbf{y}_2(k) \end{bmatrix}.$$

In addition, we adopt the following assumption for simplicity.

Assumption 1. *The pairs (A, C_1) and (A, C_2) are not observable. In addition, we assume that the overall system is observable, i. e., the pair (A, C) is observable.*

Assumption 1 corresponds to the more difficult scenario where the controller might have to combine the information gathered from $\mathbf{y}_1$ and $\mathbf{y}_2$. Later we show that the stability analysis, for the case where (A, C_1) and (or) (A, C_2) are observable, constitutes a particular case of our analysis, indicating that our Assumption 1 comes at no loss of generality.

Definition II.1. *(***Erasure Link Model** *) Consider that $\{\mathbf{r}_1(k)\}_{k=0}^{\infty}$ and $\{\mathbf{r}_2(k)\}_{k=0}^{\infty}$ represent Bernoulli stochastic processes taking values in the set $\{\mathbf{1}, \emptyset\}$ and characterized by a probability mass function of the following type:*

$$p_{i,j} \stackrel{def}{=} Pr\left(\mathbf{r}(k) = (i,j)\right), \ (i,j) \in \{\mathbf{1}, \emptyset\}^2$$

where $\mathbf{r}(k) \stackrel{def}{=} (\mathbf{r}_1(k), \mathbf{r}_2(k))$. The process $\mathbf{r}(k)$ governs the state of the links that connect the sensors to the controller. More specifically, the relationship between sensor i's output $\mathbf{s}_i(k)$ and the controller's input $\mathbf{z}_i(k)$ is described by:

$$\mathbf{z}_i(k) = \begin{cases} \emptyset & \text{if } \mathbf{r}_i(k) = \emptyset \\ \mathbf{s}_i(k) & \text{if } \mathbf{r}_i(k) = \mathbf{1} \end{cases}, \ i \in \{1,2\} \tag{4}$$

where we adopt the symbol $\emptyset$ to represent erasure, i.e., it indicates that the information sent from sensor i to the controller was lost.

Note that, in general, we do *not* assume that the erasure events in the channels are independent. However, we presuppose that the sources of randomness $\mathbf{x}(0)$, $\{\mathbf{r}(k)\}_{k=0}^{\infty}$, $\{\mathbf{v}(k)\}_{k=0}^{\infty}$ and $\{\mathbf{w}(k)\}_{k=0}^{\infty}$ are mutually independent.

We consider sensors with the following functional structure:

Definition II.2. (Sensor map classes $\mathbb{S}_q$ and $\mathbb{S}_q^{NAK}$) *For any given positive integer q, we define $\mathbb{S}_q$ as the set containing all sensor maps with the following structure:*

$$\mathbf{s}_i(k) = \begin{cases} \mathcal{S}\left(k, \mathbf{y}_i(0), \dots, \mathbf{y}_i(k), \mathbf{r}(0), \dots, \mathbf{r}(k-1)\right) & k \geq 1 \\ \mathcal{S}\left(0, \mathbf{y}_i(0)\right) & k = 0 \end{cases} \tag{5}$$

where i is in the set $\{1,2\}$ and $\mathbf{s}_i(k)$ takes values in $\mathbb{R}^q$. Notice that we assume that $\{\mathbf{r}(i)\}_{i=0}^{i=k-1}$ is made available to the sensor via noiseless acknowledgments. In addition, we consider another set $\mathbb{S}_q^{NAK}$ of sensor maps with the following structure:

$$\mathbf{s}_i(k) = \begin{cases} \mathcal{S}^{NAK}\left(k, \mathbf{y}_i(0), \dots, \mathbf{y}_i(k)\right) & k \geq 1 \\ \mathcal{S}^{NAK}\left(0, \mathbf{y}_i(0)\right) & k = 0 \end{cases} \tag{6}$$

where i is in the set $\{1,2\}$ and $\mathbf{s}_i(k)$ takes values in $\mathbb{R}^q$. Notice that $\mathbb{S}_q^{NAK}$ is the subset of $\mathbb{S}_q$ consisting of the sensor structures that do not rely on the knowledge of past values of the erasure process $\{\mathbf{r}(i)\}_{i=0}^{i=k-1}$. Equivalently,

$\mathbb{S}_q^{NAK}$ *can be specified as the set of sensor maps for which the sensor does not have access to acknowledgment signals.*

In the sequel, we will also refer to the choice of the sensor maps $\mathcal{S}_1$ and $\mathcal{S}_2$ as an encoding algorithm or an information processing algorithm and to the sensors as encoders.

Definition II.3. (Controller class) *Consider stochastic processes* $\mathbf{z}_1(k)$ *and* $\mathbf{z}_2(k)$ *taking values in* $\mathbb{R}^q \bigcup \{\mathbf{1}, \emptyset\}$. *We define the controller class* $\mathbb{K}$ *as the set of all controllers with the following structure:*

$$\mathbf{u}(k) = \mathcal{K}(k, \mathbf{z}_1(0), \mathbf{z}_2(0), \ldots, \mathbf{z}_1(k), \mathbf{z}_2(k)) \tag{7}$$

where $\mathbf{u}(k)$ *takes values in* $\mathbb{R}^l$ *and* l *is the dimension of the control input to the plant specified in (1).*

B. Problem formulation

Given the description of the plant and the erasure link statistics, specified by the probability mass function $p_{\,i,j}$, we want to minimize a quadratic cost function using controllers and sensor maps, in the classes $\mathbb{S}_{\,q}$ or $\mathbb{S}_q^{NAK}$ and investigate conditions for the existence of such controllers and sensor maps that stabilize the plant in the following sense.

Definition II.4. *(* **Cost Function and Stability Criterion)** *Consider the set-up of Figure 1 and assume that the matrices* A*,* B*,* C_1*,* C_2 *and the erasure link statistics* $p_{i,j}$ *are given. We wish to find the integer* q*, controller* $\mathcal{K}$ *(in the specified class* $\mathbb{S}_q$ *or* $\mathbb{S}_q^{NAK}$*) and sensor maps* $\mathcal{S}_1$ *and* $\mathcal{S}_2$ *that minimize the familiar quadratic cost function*

$$J_K = \sum_{k=0}^{K} E_{\beta(k),\mathbf{x}(0)}[\mathbf{x}(k)^T Q \mathbf{x}(k) + \mathbf{u}(k)^T R \mathbf{u}(k)], \tag{8}$$

where Q *and* R *are positive definite matrices,* $\mathbf{x}(k)$ *is the state of the plant and the set* $\beta(k) \overset{def}{=} \{x(0), \mathbf{v}(i), \mathbf{w}(i)\}_{i=0}^{i=k}$ *is used to indicate that the expectation is taken with respect to the initial condition, the process noise and the measurement noises. Further, a selection of controller* $\mathcal{K}$*, integer* q *and sensor maps* $\mathcal{S}_1$ *and* $\mathcal{S}_2$ *is stabilizing if and only if*

$$E_{\mathbf{r}_i}[J_\infty] \overset{def}{=} \lim_{K\to\infty} E_{\mathbf{r}_i}\left[\frac{J_K}{K}\right] < \infty, \tag{9}$$

where $E_{\mathbf{r}_i}[.]$ *indicates that the expectation is further taken with respect to the erasure events.*

We wish to point out that the expectation in equation (8) is not taken with respect to the erasure processes $\mathbf{r}(i)$; the design we propose will often be optimal for any realization of the packet dropping process. For future reference, we will denote the minimal cost achieved in equation (8) by using the optimal controller for a particular encoding algorithm $\mathcal{A}$ as $J_K^{\star,\mathcal{A}}$. Finally, for a comparison between our framework and existing work on stabilizability of decentralized control under data-rate constraints see Section III-A.

III. Necessary and sufficient conditions for stability

In this section, we state the necessary and sufficient conditions for stabilization, in terms of the state space representation of the plant and the probabilistic description of the erasure process at the links that connect the sensors to the controller. The proofs of the results will be constructed in several stages, going from the proof of a separation principle in Section IV to the description of the optimal control algorithm in Section V. We will rely on the following Proposition, regarding the state space representation of linear systems of the form (1)-(3) that is proved in the Appendix.

Proposition III.1. *Consider an n dimensional linear and time-invariant system satisfying Assumption 1 and let $\mathbf{y}_1(k)$ and $\mathbf{y}_2(k)$, taking values in $\mathbb{R}^{m_1}$ and $\mathbb{R}^{m_2}$, constitute a bi-partition of the system's output. We can always construct a state-space representation with the structure (1)-(3), where the matrices $A \in \mathbb{R}^{n\times n}$, $B \in \mathbb{R}^{n\times l}$, $C_1 \in \mathbb{R}^{m_1\times n}$ and $C_2 \in \mathbb{R}^{m_2\times n}$ are written in one and only one of the following forms, which we refer to as* **type I** *and* **type II**. *The first possibility, denoted as* **type I**, *is given by:*

$$A = \begin{bmatrix} A_{1,1} & A_{1,2} \\ \mathbf{0}^{n_2\times n_1} & A_{2,2} \end{bmatrix} \tag{10}$$

$$C_1 = \begin{bmatrix} \mathbf{0}^{m_1\times n_1} & C_{1,2} \end{bmatrix} \tag{11}$$

$$C_2 = \begin{bmatrix} C_{2,1} & \mathbf{0}^{m_2\times n_2} \end{bmatrix} \tag{12}$$

where $A_{i,j} \in \mathbb{R}^{n_i\times n_j}$, $C_{i,j} \in \mathbb{R}^{m_i\times n_j}$ and $n_1 + n_2 = n$.

*The following is the second possibility (***type II***) :*

$$A = \begin{bmatrix} A_{1,1} & A_{1,2} & A_{1,3} \\ \mathbf{0}^{n_2\times n_1} & A_{2,2} & A_{2,3} \\ \mathbf{0}^{n_3\times n_1} & \mathbf{0}^{n_3\times n_2} & A_{3,3} \end{bmatrix} \tag{13}$$

$$C_1 = \begin{bmatrix} \mathbf{0}^{m_1\times n_1} & C_{1,2} & C_{1,3} \end{bmatrix} \tag{14}$$

$$C_2 = \begin{bmatrix} C_{2,1} & \mathbf{0}^{m_2\times n_2} & C_{2,3} \end{bmatrix} \tag{15}$$

where $A_{i,j} \in \mathbb{R}^{n_i\times n_j}$, $C_{i,j} \in \mathbb{R}^{m_i\times n_j}$ and $n_1 + n_2 + n_3 = n$.

Remark III.1. *In the above representations (of types I or II), $A_{1,1}$ describes the dynamics of the state subspace that is not observable from $\mathbf{y}_1(k)$, while the modes that are not observable by $\mathbf{y}_2(k)$ follow the dynamics of $A_{2,2}$. If the representation is of type II, then $A_{3,3}$ specifies the dynamics of the modes that are observable by both $\mathbf{y}_1(k)$ and $\mathbf{y}_2(k)$. Such representations are particularly convenient for the purposes of this paper. Alternative modal decomposition for decentralized systems are also possible, see, for instance, [1].*

Using the representation above, we can state the necessary conditions for stabilizability of the system as follows. The proof is provided in Section V-B.

Theorem III.2. *(**Necessary Conditions for Stabilizability**) Consider the scheme of Fig 1 and let $A \in \mathbb{R}^{n\times n}$, $B \in \mathbb{R}^{n\times l}$, $C_1 \in \mathbb{R}^{m_1\times n}$ and $C_2 \in \mathbb{R}^{m_2\times n}$ be given matrices specifying the state-space representation for the plant. In addition, assume that the plant satisfies Assumption 1 and that the statistics of the erasure links is specified by a given probability mass function $Pr(\mathbf{r}(k) = (i,j))$, with $(i,j) \in \{\mathbf{1}, \emptyset\}^2$ that is independent of the time index k. Suppose that the state-space representation can be written as in (10)-(12) (type I). There exists a controller in the class $\mathbb{K}$, a positive integer q and sensors in the class $\mathbb{S}_q$ such that the closed loop system is stable only if the following inequalities hold:*

$$\varrho(A_{1,1})^2 Pr(\mathbf{r}_2(k) = \emptyset) < 1 \tag{16}$$

$$\varrho(A_{2,2})^2 Pr(\mathbf{r}_1(k) = \emptyset) < 1 \tag{17}$$

where $\varrho(A_{i,i})$ represents the spectral radius of the matrix $A_{i,i}$. If, instead, the state-space representation is of type II, i. e. of the form (13)-(15), then necessary conditions for stabilization also include the following additional inequality:

$$\varrho(A_{3,3})^2 Pr\left(\mathbf{r}(k) = (\emptyset, \emptyset)\right) < 1 \tag{18}$$

Remark III.2. *The case when Assumption 1 does not hold and the system is observable using only one sensor has already been considered in the literature [16]. Our results can be applied to this case if we adopt the convention that the spectral radius of an empty matrix is 0. Thus, e.g., if the entire state is observable from $\mathbf{y}_1(k)$, then the spectral radius of $A_{1,1}$ is assumed to be 0. A similar statement can be said about the sufficiency conditions given below as well. Thus we will assume that Assumption 1 holds in our analysis from now on.*

It turns out that the above conditions are also sufficient for stabilizability for sensors in the class $\mathbb{S}_q^{NAK}$ (and hence in the class $\mathbb{S}_q$). We have the following result that will be proved in Section V-C.

Theorem III.3. *(**Sufficient conditions for stabilizability**) Consider the set-up of Figure 1 and let $A \in \mathbb{R}^{n\times n}$, $B \in \mathbb{R}^{n\times l}$, $C_1 \in \mathbb{R}^{m_1\times n}$ and $C_2 \in \mathbb{R}^{m_2\times n}$ be given matrices specifying the state-space representation for the plant. In addition, assume that the plant is controllable and that it satisfies Assumption 1. In addition, let the statistics of the erasure link, given by the probability mass function $Pr(\mathbf{r}(k) = (i,j))$, $(i,j) \in \{\mathbf{1}, \emptyset\}^2$, be given. Consider that the state space representation can be written as in (10)-(12) (type I). There exists a controller of class $\mathbb{K}$, a positive integer q and sensors of class $\mathbb{S}_q^{NAK}$ such that the feedback system is stable, if the following two inequalities hold:*

$$\varrho(A_{1,1})^2 Pr(\mathbf{r}_2(k) = \emptyset) < 1 \tag{19}$$

$$\varrho(A_{2,2})^2 Pr(\mathbf{r}_1(k) = \emptyset) < 1 \tag{20}$$

where $\varrho(A_{i,i})$ represents the spectral radius of the matrix $A_{i,i}$. If the state-space representation is of type II, i.e. it is of the form (13)-(15), then stability is assured by requiring that the following additional inequality also holds:

$$\varrho(A_{3,3})^2 Pr\left(\mathbf{r}(k) = (\emptyset, \emptyset)\right) < 1 \tag{21}$$

Remark III.3. *The inequalities in Theorems III.2 and III.3 are identical. However, notice that Theorem III.3 states that if such inequalities hold then stabilization is achievable by using sensors of class $\mathbb{S}_q^{NAK}$, while Theorem III.2 characterizes the necessary condition for stabilization by allowing sensors of class $\mathbb{S}_q$. This subtle difference, and the fact that $\mathbb{S}_q^{NAK} \subset \mathbb{S}_q$, lead to the interesting conclusion that the use of acknowledge signals $\{\mathbf{r}(i)\}_{i=0}^{i=k-1}$ at the sensors does not impact stabilizability. The use of $\{\mathbf{r}(i)\}_{i=0}^{i=k-1}$ is crucial, however, in the optimal control strategy that will be identified in Section V-A.*

Remark III.4. *The stabilizability conditions make intuitive sense. The quantity $\varrho(A_{1,1})^2$ measures the rate of increase of the second moment of the modes that are observable using only sensor 2. To keep the estimate error covariance of these modes bounded, we need the information from sensor 2 to arrive at a large enough rate. Equation (16) formalizes this relation. Similarly, the inequality in (17) places a constraint on the drop rate of information from sensor 1 in terms of the rate of increase of the modes that are observable solely through sensor 1. Finally, the relation in (18) places a constraint on the arrival rate of information from at least one of the sensors in terms of the modes that are observable from either sensor.*

If the erasure processes $\mathbf{r}_1(k)$ and $\mathbf{r}_2(k)$ are independent, then the inequalities in Theorems III.2 and III.3 lead to a log-convex stabilizability region, in terms of the erasure probabilities (see Fig 2).

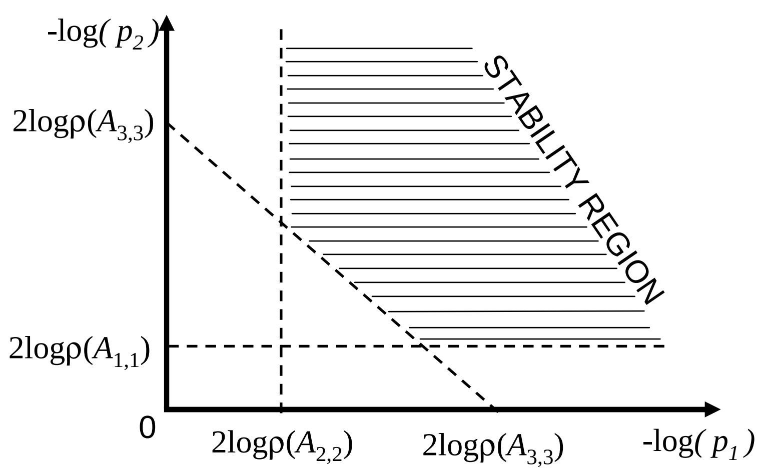

Fig. 2. Log-convex stabilizability region in terms of $p_1 = Pr(\mathbf{r}_1(k) = \emptyset)$ and $p_2 = Pr(\mathbf{r}_2(k) = \emptyset)$, under the assumption that $\mathbf{r}_1(k)$ and $\mathbf{r}_2(k)$ are independent.

A. Comparison with existing results on decentralized stabilizability under data-rate constraints

The authors of [28], [31], [42], [45] have derived necessary and sufficient conditions under a similar framework, but considering finite data-rate communication links. It is interesting to note that [42, Fig 3.] also defines a convex stabilizability region, in terms of data-rates, which is similar to our Fig 2. However, we must stress that our result cannot be derived from the finite data-rate framework of any of these works because of the following main reasons.

1) The work of [28], [31], [42], [45] considers communication links featuring deterministic and finite data-rate constrains. Their results are derived based on information theoretic ideas and counting arguments. In contrast, our results cannot be derived using information theoretic or counting arguments because our erasure links have infinite capacity (in the information theoretic sense), provided that the probability of erasure at the links is strictly less than one[1]. Existing work for finite data-rate (stochastic) erasure channels addresses only a single link and it does not provide guarantees of optimality [27], [35].
2) Our work considers measurement noise and disturbances while the authors [28], [31], [42], [45] focus on autonomous asymptotic stability.
3) We also derive optimal control strategies, while the work of [42], [45] addresses solely stabilizability.

IV. A Separation Principle in the presence of noise free acknowledgments

We begin by presenting a separation principle that allows us to consider an equivalent estimation problem instead of the control problem formulated above. At any time k, define the time-stamp corresponding to sensor i as

$$\mathbf{t}_i(k) = \max\{j \in \{0,1,2,\ldots\} \mid j \leq k, \quad \mathbf{r}_i(j) = \mathbf{1}\}.$$

Thus the time-stamp denotes the latest time at which transmission was possible from sensor i. Using the time-stamp, we can define the maximal information set $\mathcal{I}_i^{\max}(k)$ for each sensor as

$$\mathcal{I}_i^{\max}(k) = \{\mathbf{y}_i(0), \mathbf{y}_i(1), \cdots, \mathbf{y}_i(\mathbf{t}_i(k)), \mathbf{r}_i(0), \mathbf{r}_i(1), \cdots, \mathbf{r}_i(k)\}.$$

The maximal information set is the largest set of measurements from sensor i that the controller can possibly have access to at time k. For any encoding algorithm $\mathcal{A}$ followed by the sensors, we will also define the information set corresponding to sensor i at time k as

$$\mathcal{I}_i^{\mathcal{A}}(k) = \{\mathbf{s}_i(0), \mathbf{s}_i(t_i(1)), \mathbf{s}_i(t_i(2)), \ldots, \mathbf{s}_i(t_i(k))\},$$

where $\mathbf{s}_i(k)$ is the output of the sensor i at time k, when the algorithm $\mathcal{A}$ is followed at sensor i. From the definition of the time stamp $t_i(k)$ it follows that $\mathcal{I}_i^{\mathcal{A}}(k)$ comprises the time-samples of $\mathbf{s}_i(k)$ which can be recovered at the controller via the output of the erasure link from sensor i. For any encoding strategy followed by the sensors, the following inclusion holds:

$$\mathbb{I}_i^{\mathcal{A}}(k) \subseteq \mathbb{I}_i^{\max}(k).$$

where $\mathbb{I}_i^{\mathcal{A}}(k)$ and $\mathbb{I}_i^{\max}(k)$ are the smallest sigma algebras (filtrations) generated by $\mathcal{I}_i^{\mathcal{A}}(k)$ and $\mathcal{I}_i^{\max}(k)$, respectively.

Consider two encoding algorithms $\mathcal{A}_1$ and $\mathcal{A}_2$ that guarantee at every time step

$$\mathbb{I}_1^{\mathcal{A}_1}(k) \subseteq \mathbb{I}_1^{\mathcal{A}_2}(k), \qquad \mathbb{I}_2^{\mathcal{A}_1}(k) \subseteq \mathbb{I}_2^{\mathcal{A}_2}(k).$$

[1]As explained in the Introduction, we assume that quantization error is not an issue since typically a communication packet will assign a large number of bits for the data transmitted by the sensors.

With the optimal controller design for the two algorithms, the values of cost J_K achieved using the two algorithms will satisfy

$$J_K^{\star,\mathcal{A}_1} \leq J_K^{\star,\mathcal{A}_2}.$$

Now consider an algorithm $\bar{\mathcal{A}}$ under which, at every time step k the encoder for sensor i transmits the set

$$S_i(k) = \{\mathbf{y}_i(0), \mathbf{y}_i(1), \cdots, \mathbf{y}_i(k), \mathbf{r}_i(0), \mathbf{r}_i(1), \cdots, \mathbf{r}_i(k)\}.$$

Note that the algorithm $\bar{\mathcal{A}}$ does not specify valid sensor maps $\mathbb{S}_q$ since the dimension of the transmitted vectors cannot be bounded by any constant q. However, if algorithm $\bar{\mathcal{A}}$ is followed, at any time step k, the decoder (and the controller) would have access to the maximal information sets $\mathcal{I}_1^{\max}(k)$ and $\mathcal{I}_2^{\max}(k)$. This implies that for any other encoding algorithm $\mathcal{A}$, the cost function

$$J_K^{\star,\bar{\mathcal{A}}} \leq J_K^{\star,\mathcal{A}}.$$

Thus, in particular, one way to achieve the optimal value of J_K is through the combination of an encoding algorithm that makes the information sets $\mathcal{I}_i^{\max}(k)$'s available to the controller and a controller that optimally utilizes the information set. Further, one such information processing algorithm is the algorithm $\bar{\mathcal{A}}$ described above. However, this algorithm relies on the transmission of real vectors whose dimension increases linearly over time. In the sequel, we show that this difficulty can be avoided in the presence of noiseless acknowledgments. In particular, we prove that optimal performance can be achieved by using sensors of the class $\mathbb{S}_q$, where q is a finite constant quantifying the dimension of the transmitted vectors.

To this end, we begin by a statement of the familiar separation principle when algorithm $\bar{\mathcal{A}}$ is used. For any random variable $\alpha(k)$, denote by $\hat{\alpha}\left(k|\beta(k)\right)$ the minimum mean squared error (mmse) estimate of $\alpha(k)$ given the information $\beta(k)$.

Proposition IV.1. *[Separation Principle] Consider the problem defined in section II-B. Suppose that the encoding algorithm $\bar{\mathcal{A}}$ as described above is followed, so that the controller has access to the maximal information sets $\mathcal{I}_i^{\max}(k)$'s at every time step k. Then, for an optimizing choice of the control, the control and estimation costs decouple. Specifically, the optimal control input at time k is calculated by using the relation*

$$\mathbf{u}(k) = \hat{\mathbf{u}}_{LQ}\left(k|\mathcal{I}_1^{\max}(k), \mathcal{I}_2^{\max}(k), \{\mathbf{u}(t)\}_{t=0}^{k-1}\right) \stackrel{def}{=} E\left[\mathbf{u}(k)|\mathcal{I}_1^{\max}(k), \mathcal{I}_2^{\max}(k), \{\mathbf{u}(t)\}_{t=0}^{k-1}\right],$$

where $\mathbf{u}_{LQ}(k)$ denotes the optimal LQ control law corresponding to the cost function (8).

Proof: The proof is along the lines of the standard separation principle (see, e.g., [20, Chapter 9]; see also [16]) and is omitted for space constraints. ☐

There are two reasons this separation principle is useful to us:

1) We recognize that the optimal controller does not need to have access to the information sets $\mathcal{I}_i^{\max}(k)$'s at every time step k. The encoders and the decoder only need to ensure that the controller receives the quantity $\hat{\mathbf{u}}_{LQ}\left(k|\mathcal{I}_1^{\max}(k), \mathcal{I}_2^{\max}(k), \{\mathbf{u}(t)\}_{t=0}^{k-1}\right)$, or equivalently, $\hat{\mathbf{x}}\left(k|\mathcal{I}_1^{\max}(k), \mathcal{I}_2^{\max}(k), \{\mathbf{u}(t)\}_{t=0}^{k-1}\right)$.

2) If we can ensure that the controller has access to this quantity, the controller design part of the problem is solved. The optimal controller is the solution to the LQ control problem.

Our next result allows us to make another simplification in the problem by separating the dependence of the estimate on measurements from the effect of the control inputs. In the context of our problem, this is useful since the encoders do not have access to the control inputs[2]. Thus, the effect of the previous control inputs has to be included by the controller that has access to all previous control inputs. To this end, we state the following Proposition.

Proposition IV.2. *[Separation of Control and Measurement Effects] Consider the problem formulation as defined in section II-B. The quantity $\hat{\mathbf{u}}_{LQ}\left(k|\mathcal{I}_1^{\max}(k), \mathcal{I}_2^{\max}(k), \{\mathbf{u}(t)\}_{t=0}^{k-1}\right)$, can be calculated as the sum of two quantities*

$$\hat{\mathbf{u}}_{LQ}\left(k|\mathcal{I}_1^{\max}(k), \mathcal{I}_2^{\max}(k), \{\mathbf{u}(t)\}_{t=0}^{k-1}\right) = \bar{\mathbf{x}}(k) + \psi(k),$$

where $\bar{\mathbf{x}}(k)$ depends only on the information sets $\mathcal{I}_i^{\max}(k)$'s (and not on the control inputs $\{\mathbf{u}(t)\}_{t=0}^{k-1}$) and $\psi(k)$ depends only on the control inputs (but not on the information sets).

Proof: Assume, without loss of generality that $\mathbf{t}_1(k) \leq \mathbf{t}_2(k)$. We will denote the error covariance in calculating the estimate of $\mathbf{x}(k)$ using measurements from sensor 1 till time j, from sensor 2 till time m and previous control inputs till time $k-1$ by $\mathbf{P}(k|j,m)$.

To begin with, note that

$$\hat{\mathbf{x}}\left(k|\mathcal{I}_1^{\max}(k), \mathcal{I}_2^{\max}(k), \{\mathbf{u}(t)\}_{t=0}^{k-1}\right) = \hat{\mathbf{x}}\left(k|\{\mathbf{y}_1(t)\}_{t=0}^{\mathbf{t}_1(k)}, \{\mathbf{y}_2(t)\}_{t=0}^{\mathbf{t}_2(k)}, \{\mathbf{u}(t)\}_{t=0}^{k-1}\right).$$

The quantity on the right hand side can be calculated using the Kalman filter in three steps.

1) First calculate $\hat{\mathbf{x}}\left(\mathbf{t}_1(k)|\{\mathbf{y}_1(t)\}_{t=0}^{\mathbf{t}_1(k)}, \{\mathbf{y}_2(t)\}_{t=0}^{\mathbf{t}_1(k)}, \{\mathbf{u}(t)\}_{t=0}^{\mathbf{t}_1(k)-1}\right)$ by processing the measurements from both the sensors and the control inputs from time $m=0$ to $\mathbf{t}_1(k)-1$ as follows:

   Measurement Update for the Kalman filter:

$$(\mathbf{P}(m|m,m))^{-1} = (\mathbf{P}(m|m-1,m-1))^{-1} + C^T\mathbf{R}_v^{-1}C \tag{22}$$

$$\begin{aligned}&(\mathbf{P}(m|m,m))^{-1}\hat{\mathbf{x}}\left(m|\{\mathbf{y}_1(t)\}_{t=0}^{m}, \{\mathbf{y}_2(t)\}_{t=0}^{m}, \{\mathbf{u}(t)\}_{t=0}^{m-1}\right)\\ &\qquad = (\mathbf{P}(m|m-1,m-1))^{-1}\hat{\mathbf{x}}\left(m|\{\mathbf{y}_1(t)\}_{t=0}^{m-1}, \{\mathbf{y}_2(t)\}_{t=0}^{m-1}, \{\mathbf{u}(t)\}_{t=0}^{m-1}\right) + C^T\mathbf{R}_v^{-1}\mathbf{y}(m).\end{aligned}$$

   Time Update for the Kalman filter:

$$\mathbf{P}(m|m-1,m-1) = A\mathbf{P}(m-1|m-1,m-1)A^T + \mathbf{R}_w \tag{23}$$

$$\begin{aligned}&\hat{\mathbf{x}}\left(m|\{\mathbf{y}_1(t)\}_{t=0}^{m-1}, \{\mathbf{y}_2(t)\}_{t=0}^{m-1}, \{\mathbf{u}(t)\}_{t=0}^{m-1}\right)\\ &\qquad = A\hat{\mathbf{x}}\left(m|\{\mathbf{y}_1(t)\}_{t=0}^{m-1}, \{\mathbf{y}_2(t)\}_{t=0}^{m-1}, \{\mathbf{u}(t)\}_{t=0}^{m-2}\right) + B\mathbf{u}(m-1).\end{aligned}$$

[2]Note that we do not assume knowledge of even the cost function at the encoders.

The initial conditions for the Kalman filter are given by $\hat{\mathbf{x}}\left(0|\mathbf{y}_1(-1),\mathbf{y}_2(-1),\mathbf{u}(-1)\right)=0$ and $\mathbf{P}(0|-1,-1)=\mathbf{P}(0)$.

2) Calculate $\hat{\mathbf{x}}\left(\mathbf{t}_2(k)|\{\mathbf{y}_1(t)\}_{t=0}^{\mathbf{t}_1(k)},\{\mathbf{y}_2(t)\}_{t=0}^{\mathbf{t}_2(k)},\{\mathbf{u}(t)\}_{t=0}^{\mathbf{t}_2(k)-1}\right)$ by processing the measurements from sensor 2 and the control inputs from time $m=\mathbf{t}_1(k)$ to $m=\mathbf{t}_2(k)-1$ again using the Kalman filter equations, but only considering new measurements to arrive from sensor 2.

3) Construct the estimate $\hat{\mathbf{x}}\left(k|\{\mathbf{y}_1(t)\}_{t=0}^{\mathbf{t}_1(k)},\{\mathbf{y}_2(t)\}_{t=0}^{\mathbf{t}_2(k)},\{\mathbf{u}(t)\}_{t=0}^{k-1}\right)$ as

$$\begin{aligned}&\hat{\mathbf{x}}\left(k|\{\mathbf{y}_1(t)\}_{t=0}^{\mathbf{t}_1(k)},\{\mathbf{y}_2(t)\}_{t=0}^{\mathbf{t}_2(k)},\{\mathbf{u}(t)\}_{t=0}^{k-1}\right)\\&\quad=A^{k-\mathbf{t}_2(k)}\hat{\mathbf{x}}\left(\mathbf{t}_2(k)|\{\mathbf{y}_1(t)\}_{t=0}^{\mathbf{t}_1(k)},\{\mathbf{y}_2(t)\}_{t=0}^{\mathbf{t}_2(k)},\{\mathbf{u}(t)\}_{t=0}^{j}\right)+\sum_{i=0}^{k-\mathbf{t}_2(k)-1}A^iB\mathbf{u}(k-i-1).\end{aligned}\tag{24}$$

The effect of the control inputs appears linearly and can be considered separately. To see this, calculate the quantity $\acute{\mathbf{x}}(j|\{\mathbf{y}_1(t)\}_{t=0}^{\mathbf{t}_1(k)},\{\mathbf{y}_2(t)\}_{t=0}^{\mathbf{t}_2(k)})$ using the measurements from sensor 1 from time 0 to j_1 and sensor 2 from time 0 to j_2 according to a *modified Kalman filter*.

1) First calculate $\acute{\mathbf{x}}\left(\mathbf{t}_1(k)|\{\mathbf{y}_1(t)\}_{t=0}^{\mathbf{t}_1(k)},\{\mathbf{y}_2(t)\}_{t=0}^{\mathbf{t}_1(k)}\right)$ by processing the measurements from both the sensors.

Measurement Update for the modified Kalman filter:

$$\left(\mathbf{P}(m|m,m)\right)^{-1}=\left(\mathbf{P}(m|m-1,m-1)\right)^{-1}+C^T\mathbf{R}_{\mathbf{v}}^{-1}C\tag{25}$$

$$\begin{aligned}&\left(\mathbf{P}(m|m,m)\right)^{-1}\acute{\mathbf{x}}(m|\{\mathbf{y}_1(t)\}_{t=0}^{m},\{\mathbf{y}_2(t)\}_{t=0}^{m})\\&\quad=\left(\mathbf{P}(m|m-1,m-1)\right)^{-1}\acute{\mathbf{x}}(m|\{\mathbf{y}_1(t)\}_{t=0}^{m-1},\{\mathbf{y}_2(t)\}_{t=0}^{m-1})+C^T\mathbf{R}_{\mathbf{v}}^{-1}\mathbf{y}(m),\end{aligned}$$

Time Update for the modified Kalman filter:

$$\begin{aligned}\mathbf{P}(m|m-1,m-1)&=A\mathbf{P}(m-1|m-1,m-1)A^T+\mathbf{R}_w\\\acute{\mathbf{x}}(m|\{\mathbf{y}_1(t)\}_{t=0}^{m-1},\{\mathbf{y}_2(t)\}_{t=0}^{m-1})&=A\acute{\mathbf{x}}(m-1|\{\mathbf{y}_1(t)\}_{t=0}^{m-1},\{\mathbf{y}_2(t)\}_{t=0}^{m-1}).\end{aligned}\tag{26}$$

The initial conditions are given by $\acute{\mathbf{x}}\left(0|\mathbf{y}_1(-1),\mathbf{y}_2(-1)\right)=0$ and $\mathbf{P}(0|-1,-1)=\mathbf{P}(0)$. Note that calculation of the terms $\mathbf{P}(m|m,m)$ and $\mathbf{P}(m|m-1,m-1)$ does not require knowledge of either the measurements $\mathbf{y}_i(m)$'s or the control inputs $\mathbf{u}(m)$'s and these terms can even be calculated offline. The equations for the modified Kalman filter are identical to the ones for the Kalman filter given in (22) and (23) except that the control input $\mathbf{u}(m)$ is assumed to be 0.

2) Calculate $\acute{\mathbf{x}}\left(\mathbf{t}_2(k)|\{\mathbf{y}_1(t)\}_{t=0}^{\mathbf{t}_1(k)},\{\mathbf{y}_2(t)\}_{t=0}^{\mathbf{t}_2(k)}\right)$ by processing the measurements from sensor 2 again using the modified Kalman filter equations, but only considering new measurements to arrive from sensor 2.

3) Include the effect of the control inputs through the term $\tilde{\psi}(j)$ that evolves as

$$\tilde{\psi}(m)=B\mathbf{u}(m-1)+\Gamma(m-1)\tilde{\psi}(m-1),$$

where

$$\Gamma(m)=\begin{cases}A\left(\mathbf{P}(m-1|m-1,m-1)\right)^{-1}\mathbf{P}(m-1|m-2,m-2) & m\leq\mathbf{t}_1(k)+1\\A\left(\mathbf{P}(m-1|\mathbf{t}_1(k),m-1)\right)^{-1}\mathbf{P}(m-1|\mathbf{t}_1(k),m-2) & \text{otherwise,}\end{cases}$$

and the initial condition is $\tilde{\psi}(0) = 0$.

It can readily be verified that

$$\hat{\mathbf{x}}\left(\mathbf{t}_2(k)|\{\mathbf{y}_1(t)\}_{t=0}^{\mathbf{t}_1(k)}, \{\mathbf{y}_2(t)\}_{t=0}^{\mathbf{t}_2(k)}, \{\mathbf{u}(t)\}_{t=0}^{k-1}\right) = \acute{\mathbf{x}}\left(\mathbf{t}_2(k)|\{\mathbf{y}_1(t)\}_{t=0}^{\mathbf{t}_1(k)}, \{\mathbf{y}_2(t)\}_{t=0}^{\mathbf{t}_2(k)}\right) + \tilde{\psi}(\mathbf{t}_2(k)+1). \quad (27)$$

The estimate at time step k can then once again be calculated using (24). Comparing the two methods, we see that

$$\begin{aligned}\hat{\mathbf{x}}\left(k|\{\mathbf{y}_1(t)\}_{t=0}^{\mathbf{t}_1(k)}, \{\mathbf{y}_2(t)\}_{t=0}^{\mathbf{t}_2(k)}, \{\mathbf{u}(t)\}_{t=0}^{k-1}\right) &= A^{k-\mathbf{t}_2(k)}\acute{\mathbf{x}}\left(\mathbf{t}_2(k)|\{\mathbf{y}_1(t)\}_{t=0}^{\mathbf{t}_1(k)}, \{\mathbf{y}_2(t)\}_{t=0}^{\mathbf{t}_2(k)}\right) \\ &\quad + A^{k-\mathbf{t}_2(k)}\tilde{\psi}(\mathbf{t}_2(k)) + \sum_{i=0}^{k-\mathbf{t}_2(k)-1} A^i B\mathbf{u}(k-i-1),\end{aligned}$$

where the term $\acute{\mathbf{x}}(.|.,.)$ depends only on the measurements and the terms $\tilde{\psi}(.)$ and $\mathbf{u}(.)$ depend only on the control inputs. To complete the proof, we simply identify

$$\begin{aligned}\bar{\mathbf{x}}(k) &= A^{k-\mathbf{t}_2(k)}\acute{\mathbf{x}}\left(\mathbf{t}_2(k)|\{\mathbf{y}_1(t)\}_{t=0}^{\mathbf{t}_1(k)}, \{\mathbf{y}_2(t)\}_{t=0}^{\mathbf{t}_2(k)}\right) \\ \psi(k) &= A^{k-\mathbf{t}_2(k)}\tilde{\psi}(\mathbf{t}_2(k)) + \sum_{i=0}^{k-\mathbf{t}_2(k)-1} A^i B\mathbf{u}(k-i-1).\end{aligned}$$

□

As mentioned above, the advantage of separating the effects of measurements and the control inputs is that the sensors do not need access to the control inputs. The controller (which has access to all the control inputs) can calculate $\psi(k)$ and, in turn, the estimate of the state. Finally, note that the term $\bar{\mathbf{x}}(k)$ that the network needs to deliver is, in fact, the mmse estimate of the state $\mathbf{x}(k)$ of a process evolving as

$$\mathbf{x}(k+1) = A\mathbf{x}(k) + \mathbf{w}(k), \quad (28)$$

given the measurements $\mathbf{y}_1(0)$, $\mathbf{y}_1(1)$, $\cdots$, $\mathbf{y}_1(\mathbf{t}_1(k))$ and $\mathbf{y}_2(0)$, $\mathbf{y}_2(1)$, $\cdots$, $\mathbf{y}_2(\mathbf{t}_2(k))$ that are assumed to originate from sensors of the form (3).

Thus consider **an alternative estimation problem**. A process of the form (28) is observed by sensors of the form (3). The sensors transmit data across channels with the erasure link models defined in Definition II.1. The sensor maps are also as defined above. There is an estimator across the links that needs to estimate the state $\mathbf{x}(k)$ of the process in the mmse sense at every time step k. We can once again define the information set $\mathcal{I}_i(k)$ that contains the information from the sensor i that the controller has access to at time k and the corresponding maximal information set $\mathcal{I}_i^{\max}(k)$. What are the optimal sensor maps that allow the estimator to calculate the estimate of $\mathbf{x}(k)$ based on the information sets $\mathcal{I}_i^{\max}(k)$'s? By the arguments above, the optimal sensor maps in the original control problem and this estimation problem are identical. The stability criterion presented in Definition II.4 for the control problem corresponds to the condition that for the estimate $\hat{\mathbf{x}}(k)$ of the estimator at step k,

$$\sup_{k\geq 0} E_{\sigma(k),\mathbf{x}(0)}\left[(\mathbf{x}(k) - \hat{\mathbf{x}}(k))^T (\mathbf{x}(k) - \hat{\mathbf{x}}(k))\right] < \infty. \quad (29)$$

For the presentation of the algorithm and analysis of its properties, we consider this equivalent estimation problem while keeping in mind that, to solve the original control problem the controller can then calculate $\psi(k)$ to include the effect of the previous control inputs and finally compute the new control input $\mathbf{u}(k)$ by utilizing the separation principle.

V. Optimal Algorithm and Proof of the Necessary and Sufficient Conditions for Stabilizability

In this section, we will first present a recursive encoding algorithm for sensors in the class $\mathbb{S}_q$ that allows the estimator to calculate the estimate $\hat{\mathbf{x}}\left(k|\mathcal{I}_1^{\max}(k),\mathcal{I}_2^{\max}(k)\right)$ and is thus optimal. We will then analyze the stability of the algorithm which shall prove Theorem III.2. Finally, we shall prove that the conditions required for stability using the optimal algorithm are sufficient for a particular algorithm in the sensor class $\mathbb{S}_q^{NAK}$, which shall prove Theorem III.3.

A. A Recursive Algorithm for Optimal Performance

In this section, we consider the sensor class $\mathbb{S}_q$ and present an optimal encoding algorithm for the sensors. To begin with, we note the following result.

Proposition V.1. *Consider a process of the form (28) being observed by two sensors of the form (3). Let $\hat{\mathbf{x}}_i(k|l)$ denote the mmse estimate of $\mathbf{x}(k)$ based on all the measurements of sensor i up to time l. Denote the corresponding error covariance by $\mathbf{P}_i(k|l)$. The estimate $\hat{\mathbf{x}}(k|l,m)$ of the state based on measurements from sensor 1 till time l and sensor 2 till time m can be calculated using a relation of the form*

$$\hat{\mathbf{x}}(k|l,m)=f\left(I_{1,l,m}(k),I_{2,l,m}(k)\right),$$

where $I_{1,l,m}(k)$ does not depend on measurements from sensor 2 and $I_{2,l,m}(k)$ does not depend on measurements from sensor 1.

Proof: Proof is based on the algorithm proposed in [16]. Assume, without loss of generality, that $l\leq m$. The quantity $I_{1,l,m}(k)$ is calculated using the following algorithm. At each time step $j\leq k$,

1) Obtain the estimate $\hat{\mathbf{x}}_1(j|j)$ and $\mathbf{P}_1(j|j)$ through a Kalman filter. For $j\leq l$, use the measurement $\mathbf{y}_1(j)$. For $j>l$, assume that the sensor 1 did not take any measurement at time step j.
2) Calculate
$$\lambda_1(j)=\left(\mathbf{P}_1(j|j)\right)^{-1}\hat{\mathbf{x}}_1(j|j)-\left(\mathbf{P}_1(j|j-1)\right)^{-1}\hat{\mathbf{x}}_1(j|j-1).$$
3) Calculate global error covariance matrices $\mathbf{P}(j|j,j)$ and $\mathbf{P}(j|j-1,j-1)$ using the relation
$$\begin{aligned}&\left(\mathbf{P}(j|j,j)\right)^{-1}\\ &\qquad=\begin{cases}\left(\mathbf{P}(j|j-1,j-1)\right)^{-1}+\left(C_1\right)^T\left(\Sigma_{v,1}\right)^{-1}\left(C_1\right)+\left(C_2\right)^T\left(\Sigma_{v,2}\right)^{-1}\left(C_2\right) & j\leq l\\ \left(\mathbf{P}(j|j-1,j-1)\right)^{-1}+\left(C_2\right)^T\left(\Sigma_{v,2}\right)^{-1}\left(C_2\right) & l<j\leq m\\ \left(\mathbf{P}(j|j-1,j-1)\right)^{-1} & \text{otherwise,}\end{cases}\end{aligned}$$

$$\mathbf{P}(j|j-1,j-1) = A\mathbf{P}(j-1|j-1,j-1)A^T + \mathbf{R}_w.$$

4) Obtain

$$\gamma(j) = (\mathbf{P}(j|j-1,j-1))^{-1} A\mathbf{P}(j-1|j-1,j-1).$$

5) Finally calculate

$$I_{1,l,m}(j) = \lambda_1(j) + \gamma(j)I_{1,l,m}(j-1), \tag{30}$$

with $I_{1,l,m}(-1) = 0$.

The quantity $I_{2,l,m}(k)$ is calculated by a similar algorithm except using the local estimates $\hat{\mathbf{x}}_2(j|j)$ and covariance $\mathbf{P}_2(j|j)$. Finally, the estimate $\hat{\mathbf{x}}(k|l,m)$ is calculated using the relation

$$(\mathbf{P}(k|k,k))^{-1}\hat{\mathbf{x}}(k|l,m) = I_{1,l,m}(k) + I_{2,l,m}(k), \tag{31}$$

where $\mathbf{P}(k|k,k)$ is calculated as above. That $\hat{\mathbf{x}}(k|l,m)$ is indeed the mmse estimate given all the measurements from sensor 1 till time l and from sensor 2 till time m can be proved by utilizing the block diagonal structure of the matrix $\mathbf{R}_v$ as in the proof of Theorem 2 in [16]. □

The above result identifies the quantities that need to be transmitted by the two sensors to calculate the mmse estimate of $\mathbf{x}(k)$. The quantities depend only on local measurements at the sensors; however, an implicit assumption is that each sensor is informed about the times l and m.

Using this result, we can now provide the optimal encoding algorithm for the sensor class $\mathbb{S}_q$. Let $\mathbf{P}(k|l,m)$ denote the error covariance of the mmse estimate of the state $\mathbf{x}(k)$ calculated using measurements from sensor 1 till time l and from sensor 2 till time m. Now consider the following algorithm denoted for future reference as the algorithm $\mathcal{A}_{ack}$. At each time step k

- *Encoder for Sensor 1:* Because of the noiseless acknowledgments, sensor 1 can calculate the time stamp $\mathbf{t}_2(k-1)$ as

  $$\mathbf{t}_2(k-1) = \max\ j \text{ such that } \mathbf{r}_2(j) = \mathbf{1}.$$

  Encoder 1 calculates and transmits two quantities: $I_{1,k,k}(k)$ and $I_{1,k,\mathbf{t}_2(k-1)}(k)$.
- *Encoder for Sensor 2:* Sensor 2 calculates and transmits $I_{2,k,k}(k)$ and $I_{2,\mathbf{t}_1(k-1),k}(k)$.
- *Decoder at the Estimator:* The estimator maintains three quantities.
  - the estimate $\hat{\mathbf{x}}^{dec}(k)$ with the initial value $\hat{\mathbf{x}}^{dec}(-1) = 0$,
  - a vector $I_1^{dec}(k)$ for the contribution from sensor 1 with the initial value $I_1^{dec}(-1) = 0$.
  - a vector $I_2^{dec}(k)$ for the contribution from sensor 2 with the initial value $I_2^{dec}(-1) = 0$.

  At every time step k, it faces one of four situations.

  1) $\mathbf{r}_1(k) = \mathbf{r}_2(k) = \emptyset$: The decoder calculates

$$\begin{aligned} I_1^{dec}(k) &= (\mathbf{P}(k|\mathbf{t}_1(k-1),k))^{-1} A\left(\mathbf{P}(k-1|\mathbf{t}_1(k-1),k-1)\right) I_1^{dec}(k-1) \\ I_2^{dec}(k) &= (\mathbf{P}(k|k,\mathbf{t}_2(k-1)))^{-1} A\left(\mathbf{P}(k-1|k-1,\mathbf{t}_2(k-1))\right) I_2^{dec}(k-1) \\ \hat{\mathbf{x}}^{dec}(k) &= A\hat{\mathbf{x}}^{dec}(k-1). \end{aligned}$$

2) $\mathbf{r}_1(k) = \emptyset, \mathbf{r}_2(k) = \mathbf{1}$: The decoder calculates

$$\begin{aligned} I_1^{dec}(k) &= \left(\mathbf{P}(k|\mathbf{t}_1(k-1),k)\right)^{-1} A \left(\mathbf{P}(k-1|\mathbf{t}_1(k-1),k-1)\right) I_1^{dec}(k-1) \\ I_2^{dec}(k) &= I_{2,k,k} \\ \hat{\mathbf{x}}^{dec}(k) &= \mathbf{P}(k|\mathbf{t}_1(k),k) \left(I_1^{dec}(k) + I_{2,\mathbf{t}_1(k-1),k}(k)\right). \end{aligned}$$

3) $\mathbf{r}_1(k) = \mathbf{1}, \mathbf{r}_2(k) = \emptyset$: The decoder calculates

$$\begin{aligned} I_1^{dec}(k) &= I_{1,k,k} \\ I_2^{dec}(k) &= \left(\mathbf{P}(k|k,\mathbf{t}_2(k-1))\right)^{-1} A \left(\mathbf{P}(k-1|k-1,\mathbf{t}_2(k-1))\right) I_2^{dec}(k-1) \\ \hat{\mathbf{x}}^{dec}(k) &= \mathbf{P}(k|k,\mathbf{t}_2(k)) \left(I_{1,k,\mathbf{t}_2(k-1)}(k) + I_2^{dec}(k)\right). \end{aligned}$$

4) $\mathbf{r}_1(k) = \mathbf{r}_2(k) = \mathbf{1}$: The decoder calculates

$$\begin{aligned} I_1^{dec}(k) &= I_{1,k,k} \\ I_2^{dec}(k) &= I_{2,k,k} \\ \hat{\mathbf{x}}^{dec}(k) &= \mathbf{P}(k|k,k) \left(I_{1,k,k}(k) + I_{2,k,k}(k)\right). \end{aligned}$$

We can state the following result.

Theorem V.2. *In the algorithm $\mathcal{A}_{ack}$, $\hat{\mathbf{x}}^{dec}(k) = \hat{x}\left(k|\mathcal{I}_1^{\max}(k), \mathcal{I}_2^{\max}(k)\right)$, where $\mathcal{I}_i^{\max}(k)$'s are the maximal information sets defined earlier.*

Proof: The proof is straight-forward given Proposition V.1. At any time step k, the term $I_1^{dec}(k)$ equals $I_{1,\mathbf{t}_1(k),k}(k)$ and $I_2^{dec}(k)$ equals $I_{2,k,\mathbf{t}_2(k)}(k)$. For any of the four possibilities of channel outputs, it can be verified that the estimate is calculated according to equation (31). □

Note that the algorithm is optimal, yet involves a constant amount of transmission and processing. Each sensor can calculate the terms it transmits using a recursive algorithm of the form outlined in equation (30).

Remark V.1 (Optimality for any Drop Sequence and the 'Washing Away' Effect)**.** *So far, we have made no assumptions on the realization of the erasure process nor on the knowledge of the statistics of the erasure events at any of the nodes. The algorithm provides the optimal estimate for an arbitrary realization of the erasure process, irrespective of whether the erasure process can be modeled as i.i.d. or as a more sophisticated model like a Markov chain. The algorithm results in the optimal estimate at every time step for any realization of the erasure process, not merely in the optimal average performance. Also note that if data is received from sensor i at any time step k, the effect of all previous erasures from that sensor is 'washed away'. The estimate at the receiving node becomes identical to the case when all measurements $\mathbf{y}_i(0)$, $\mathbf{y}_i(1)$, $\cdots$, $\mathbf{y}_i(k)$ were available, irrespective of which previous data had been lost (erased).*

B. Necessary Conditions for Stabilizability

By analyzing the stability of the optimal algorithm $\mathcal{A}_{ack}$, we can obtain necessary conditions for stability for any encoding algorithm in the class $\mathbb{S}_q$ (and, in turn, $\mathbb{S}_q^{NAK}$). We shall need the following result that can be proved along the lines of Theorem 4 in [14].

Proposition V.3. *Consider the system in equation (28) being observed by a sensor of the form*

$$\bar{\mathbf{y}}(k) = \bar{C}\mathbf{x}(k) + \bar{\mathbf{v}}(k),$$

where $\bar{\mathbf{v}}(k)$ is white Gaussian noise with zero mean and covariance $\mathbf{R}$. Let $f(X)$ denote the Ricatti recursion corresponding to this sensor as applied on the matrix X, thus,

$$f(X) = AXA^T + \mathbf{R}_w - AX\bar{C}^T\left(\bar{C}X\bar{C}^T + R\right)^{-1}\bar{C}XA^T. \tag{32}$$

Further, let $f^m(X)$ denote the above Ricatti recursion applied m times on the matrix X, i.e.,

$$f^m(X) = \underbrace{f\left(f\left(\cdots f\left(X\right)\cdots\right)\right)}_{f \text{ applied } m \text{ times}}. \tag{33}$$

Finally, let p be a scalar. Then, the sum

$$S = X + pf(X) + p^2f^2(X) + p^3f^3(X) + \cdots + p^mf^m(X), \tag{34}$$

is bounded as $m \to \infty$ if and only if

$$p \mid \varrho(\bar{A}) \mid^2 < 1,$$

where $\varrho(\bar{A})$ is the spectral radius of the state subspace that is unobservable from $\bar{\mathbf{y}}(k)$. In particular, if the matrix $\bar{C} = 0$, so that the Ricatti recursion (32) corresponds to the Lyapunov recursion

$$f(X) = AXA^T + \mathbf{R}_w,$$

then the sum (34) converges if and only if

$$p \mid \varrho(A) \mid^2 < 1,$$

where $\varrho(A)$ is the spectral radius of matrix A.

Proof: Omitted for space constraints. □

Proof of Theorem III.2: To begin with, note that because of the separation principle stated in Propositions IV.1 and IV.2 and the optimality of the algorithm $\mathcal{A}_q$ as proved in Theorem V.2, we can consider the stability of the error covariance of the equivalent estimation problem, assuming that the process evolves as in (28) and algorithm $\mathcal{A}_q$ is used. For ease of notation, we define the Ricatti operators $f_1(.)$, $f_2(.)$ and $f_\emptyset(.)$ in a fashion similar to equation (32) when sensor 1, sensor 2 and no sensor is used, respectively. We also define $f_1^m(.)$, $f_2^m(.)$ and $f_\emptyset^m(.)$ analogously. Finally we define $M(k)$ to be the error covariance of the mmse estimate of $\mathbf{x}(k+1)$ when all the measurements from sensors 1 and 2 till time step k are available. Because of the assumption on observability of (A, C), $M(k)$ converges exponentially to a steady-state value denoted by $M^\star$.

For ease of notation, we will consider the stability of the expected error covariance $E[\mathbf{P}(k)]$ of the estimate of the state $\mathbf{x}(k+1)$ as calculated at time k. The conditions required for stabilizability of the error covariance $E[\mathbf{P}(k)]$ are identical to those required for stabilizability of the expected error covariance of the estimate of the state $\mathbf{x}(k)$ calculated at time k. We will condition the expected error covariance on events $\mathcal{E}_{mn}$ where the subscript m denotes the time at which the last transmission was successfully received from sensor 1 and n denotes the time at which the last transmission was successfully received from sensor 2. Obviously $0 \leq m, n \leq k$. We also allow the indices to attain the value -1 to denote the event when transmission from the corresponding sensor was never possible till time k. Denote the error covariance conditioned on the event $\mathcal{E}_{mn}$ happening by $\mathbf{P}_{mn}$. Due to Theorem V.2, $\mathbf{P}_{mn}$ is the error covariance in estimating $\mathbf{x}(k+1)$ based on measurements $\mathbf{y}_1(0)$, $\mathbf{y}_1(1)$, $\cdots$, $\mathbf{y}_1(m)$ from sensor 1 and $\mathbf{y}_2(0)$, $\mathbf{y}_2(1)$, $\cdots$, $\mathbf{y}_2(n)$ from sensor 2. Let p_{mn} be the probability of the event $\mathcal{E}_{mn}$ occurring.

We can thus write

$$E[\mathbf{P}(k)] = \sum_{m=-1}^{k} \sum_{n=-1}^{k} p_{mn}\mathbf{P}_{mn}.$$

Since each term in the above summation is positive semi-definite, a necessary condition for the sum to be bounded is that any sub-sequence in the sum is bounded. We will consider three particular sub-sequences and show that the conditions in (16-18) are necessary for stabilizability. First consider the sequence

$$\begin{aligned} S_1(k) &= \sum_{m=0}^{k} p_{mk}\mathbf{P}_{mk} \\ &= Pr(\mathbf{r}_1(k)=\mathbf{1})Pr(\mathbf{r}_2(k)=\mathbf{1})\Big(M(k) + Pr(\mathbf{r}_1(k)=\emptyset)f_2\left(M(k-1)\right) \\ &\quad + (Pr(\mathbf{r}_1(k)=\emptyset))^2 f_2^2\left(M(k-2)\right) + \cdots + (Pr(\mathbf{r}_1(k)=\emptyset))^k f_2^k\left(M(0)\right)\Big). \end{aligned}$$

Since $M(k)$ converges exponentially to $M^\star$ as $k \to \infty$, we can substitute $M^\star$ for the conditional error covariances to study the convergence. Thus, we obtain

$$\lim_{k\to\infty} S_1(k) = Pr(\mathbf{r}_1(k)=\mathbf{1})Pr(\mathbf{r}_2(k)=\mathbf{1}) \sum_{m=0}^{\infty} (Pr(\mathbf{r}_1(k)=\emptyset))^m f_2^m\left(M^\star\right).$$

Thus, using Proposition V.3, we can prove that this sum converges only if (16) holds.

In a similar fashion, we can prove that the condition in (17) is necessary by considering the sub-sequence

$$S_2(k) = \sum_{n=0}^{k} p_{kn}\mathbf{P}_{kn}.$$

Finally, the sub-sequence

$$S_2(k) = \sum_{n=0}^{k} p_{nn}\mathbf{P}_{nn}$$

yields the necessary condition

$$\varrho(A)^2 Pr\left(\mathbf{r}(k) = (\emptyset,\emptyset)\right) < 1. \quad (35)$$

Since $\varrho(A) = \max\{\varrho(A_{i,i})\}$, there are two cases.

1) If $\varrho(A) = \varrho(A_{3,3})$, equation (35) reduces to equation (18) and the proof is complete.

2) If either $\varrho(A) = \varrho(A_{1,1})$ or $\varrho(A) = \varrho(A_{2,2})$, equation (35) is subsumed by either equation (17) or equation (16). Moreover, equation (35) implies equation (18). Thus, the proof is complete in this case as well.

$\square$

C. *Sufficient Conditions for Stabilizability*

We now present the proof of Theorem III.3 by considering a particular algorithm in the class $\mathbb{S}_q^{NAK}$. Note that we have proved the separation principle only for an encoding/decoding algorithm that guarantees that the controller can estimate the state given the maximal information sets $\mathcal{I}_i^{\max}(k)$'s. However, even for other encoding/decoding algorithms, we note that

1) If an encoding and decoding algorithm guarantees that the state $\mathbf{x}(k)$ of the process evolving as in (1) can be estimated at the controller with bounded second moment error, then any controller of the type

$$\mathbf{u}(k) = F\hat{\mathbf{x}}(k)$$

with

$$\varrho\left(A + F\right) < 1,$$

where $\hat{\mathbf{x}}(k)$ is the estimate of state at time k, will guarantee stability of the closed loop system.

2) Moreover, since all the previous control inputs are known to the controller, the encoding algorithm only needs to ensure that the state $\mathbf{x}(k)$ of a process evolving as in (28) when observed by sensors of the form (3) can be estimated at the controller end using sensors of the form $\mathbb{S}_q^{NAK}$.

We shall now propose such an algorithm, denoted by $\mathcal{A}_{nack}$ for future reference. Due to Proposition III.1, we can consider the system to be either of **type I** or of **type II**. We can also partition the state space $\mathbf{x}(k)$ of the process in one of two ways.

1) If the system is of **type I**, denote

$$\mathbf{x}(k) = \begin{bmatrix} \mathbf{x}_1(k)^{n_1\times 1} \\ \mathbf{x}_2(k)^{n_2\times 1} \end{bmatrix}. \tag{36}$$

2) If the system is of **type II**, denote

$$\mathbf{x}(k) = \begin{bmatrix} \mathbf{x}_1(k)^{n_1\times 1} \\ \mathbf{x}_2(k)^{n_2\times 1} \\ \mathbf{x}_3(k)^{n_3\times 1} \end{bmatrix}. \tag{37}$$

Now consider the following algorithm. At each time step k

- *Encoder for Sensor 1:*
  - If the system is of **type I**, sensor 1 calculates and transmits the estimate $\hat{\mathbf{x}}_2^{loc,1}(k)$ of the modes $\mathbf{x}_2(k)$ of the process using its local measurements $\mathbf{y}_1(0)$, $\mathbf{y}_1(1)$, $\cdots$, $\mathbf{y}_1(k)$.
  - If the system is of **type II**, sensor 1 calculates and transmits the estimate $\hat{\mathbf{x}}_2^{loc,1}(k)$ and $\hat{\mathbf{x}}_3^{loc,1}(k)$ of the modes $\mathbf{x}_2(k)$ and $\mathbf{x}_3(k)$ of the process using its local measurements $\mathbf{y}_1(0)$, $\mathbf{y}_1(1)$, $\cdots$, $\mathbf{y}_1(k)$.

- *Encoder for Sensor 2:*
  - If the system is of **type I**, sensor 2 calculates and transmits the estimate $\hat{\mathbf{x}}_1^{loc,2}(k)$ of the modes $\mathbf{x}_1(k)$ of the process using its local measurements $\mathbf{y}_2(0)$, $\mathbf{y}_2(1)$, $\cdots$, $\mathbf{y}_2(k)$.
  - If the system is of **type II**, sensor 2 calculates and transmits the estimate $\hat{\mathbf{x}}_1^{loc,2}(k)$ and $\hat{\mathbf{x}}_3^{loc,2}(k)$ of the modes $\mathbf{x}_1(k)$ and $\mathbf{x}_3(k)$ of the process using its local measurements $\mathbf{y}_2(0)$, $\mathbf{y}_2(1)$, $\cdots$, $\mathbf{y}_2(k)$.
- *Decoder:*
  - If the system is of **type I**, the decoder maintains an estimate $\hat{\mathbf{x}}_1(k)$ of the modes $\mathbf{x}_1(k)$ and $\hat{\mathbf{x}}_2(k)$ of the modes $\mathbf{x}_2(k)$. At every time step k, the decoder takes the following actions.
    1) If $\mathbf{r}_1(k) = \emptyset$, $\hat{\mathbf{x}}_1(k) = A\hat{\mathbf{x}}_1(k-1)$, else $\hat{\mathbf{x}}_1(k) = \hat{\mathbf{x}}_1^{loc,2}(k)$.
    2) If $\mathbf{r}_2(k) = \emptyset$, $\hat{\mathbf{x}}_2(k) = A\hat{\mathbf{x}}_2(k-1)$, else $\hat{\mathbf{x}}_2(k) = \hat{\mathbf{x}}_2^{loc,1}(k)$.

    It then constructs the estimate $\hat{\mathbf{x}}(k)$ by stacking the estimates $\hat{\mathbf{x}}_1(k)$ and $\hat{\mathbf{x}}_2(k)$.
  - If the system is of **type II**, the decoder maintains estimates $\hat{\mathbf{x}}_1(k)$, $\hat{\mathbf{x}}_2(k)$ and $\hat{\mathbf{x}}_3(k)$ of the modes $\mathbf{x}_1(k)$, $\mathbf{x}_2(k)$ and $\mathbf{x}_3(k)$ respectively. At every time step k, the decoder takes one of the following actions:
    1) If $(\mathbf{r}_1(k), \mathbf{r}_2(k)) = (\mathbf{1}, \mathbf{1})$,
    $$\begin{aligned} \hat{\mathbf{x}}_1(k) &= \hat{\mathbf{x}}_1^{loc,2}(k) \\ \hat{\mathbf{x}}_2(k) &= \hat{\mathbf{x}}_2^{loc,1}(k) \\ \hat{\mathbf{x}}_3(k) &= \hat{\mathbf{x}}_3^{loc,1}(k). \end{aligned}$$
    2) If $(\mathbf{r}_1(k), \mathbf{r}_2(k)) = (\mathbf{1}, \emptyset)$,
    $$\begin{aligned} \hat{\mathbf{x}}_1(k) &= \hat{\mathbf{x}}_1^{loc,2}(k) \\ \hat{\mathbf{x}}_2(k) &= A\hat{\mathbf{x}}_2(k-1) \\ \hat{\mathbf{x}}_3(k) &= \hat{\mathbf{x}}_3^{loc,1}(k). \end{aligned}$$
    3) If $(\mathbf{r}_1(k), \mathbf{r}_2(k)) = (\emptyset, \mathbf{1})$,
    $$\begin{aligned} \hat{\mathbf{x}}_1(k) &= A\hat{\mathbf{x}}_1(k-1) \\ \hat{\mathbf{x}}_2(k) &= \hat{\mathbf{x}}_2^{loc,1}(k) \\ \hat{\mathbf{x}}_3(k) &= \hat{\mathbf{x}}_3^{loc,2}(k). \end{aligned}$$
    4) If $(\mathbf{r}_1(k), \mathbf{r}_2(k)) = (\emptyset, \emptyset)$,
    $$\begin{aligned} \hat{\mathbf{x}}_1(k) &= A\hat{\mathbf{x}}_1(k-1) \\ \hat{\mathbf{x}}_2(k) &= A\hat{\mathbf{x}}_2(k-1) \\ \hat{\mathbf{x}}_3(k) &= A\hat{\mathbf{x}}_3(k-1). \end{aligned}$$

  It then constructs the estimate $\hat{\mathbf{x}}(k)$ by stacking the estimates $\hat{\mathbf{x}}_1(k)$, $\hat{\mathbf{x}}_2(k)$ and $\hat{\mathbf{x}}_3(k)$.

We shall now prove that under the conditions (19-21), the estimate $\hat{\mathbf{x}}(k)$ of the state $\mathbf{x}(k)$ is stable in the sense of (29).

Proof of Theorem III.3 We give the proof if the system is of **type II**. The proof for **type I** is similar. By construction, the estimates $\hat{\mathbf{x}}_2^{loc,1}(k)$, $\hat{\mathbf{x}}_1^{loc,2}(k)$, $\hat{\mathbf{x}}_3^{loc,1}(k)$ and $\hat{\mathbf{x}}_3^{loc,2}(k)$ are stable. Denote the corresponding error covariance matrices by $K_1(k)$, $K_2(k)$, $K_3(k)$ and $K_4(k)$ respectively.

1) For the modes $\mathbf{x}_3(k)$, the error covariance evolves as follows:

$$\mathbf{P}_3(k) = \begin{cases} K_3(k) & \text{with probability } Pr(\mathbf{r}_1(k) = \mathbf{1}) \\ K_4(k) & \text{with probability } Pr(\mathbf{r}_1(k) = \emptyset)Pr(\mathbf{r}_2(k) = \mathbf{1}) \\ A_{3,3}\mathbf{P}_3(k-1)A_{3,3}^T + \mathbf{R}_{w,3} & \text{with probability } Pr(\mathbf{r}(k) = (\emptyset, \emptyset)), \end{cases}$$

where $\mathbf{R}_{w,3}$ is covariance matrix of the process noise entering the evolution of the modes $\mathbf{x}_3(k)$. Thus if (21) is satisfied, then the error for the modes $\mathbf{x}_3(k)$ will remain stable.

2) For the modes $\mathbf{x}_2(k)$, the error covariance in estimating the modes $\mathbf{x}_3(k)$ can thus be considered to be additional noise with bounded covariance. The error covariance for these modes evolves as

$$\mathbf{P}_2(k) = \begin{cases} K_2(k) & \text{with probability } Pr(\mathbf{r}_1(k) = \mathbf{1}) \\ A_{2,2}\mathbf{P}_2(k)A_{2,2}^T + \mathbf{R}_{w,2} & \text{with probability } Pr(\mathbf{r}_1(k) = \emptyset), \end{cases}$$

where $\mathbf{R}_{w,2}$ denotes the covariance of noise and error incurred through the estimation of modes $\mathbf{x}_3(k)$. Thus if (20) is satisfied, the error for the modes $\mathbf{x}_2(k)$ will be stable.

3) A similar argument shows that if (19) is satisfied, the error for the modes $\mathbf{x}_1(k)$ will be stable.

$\square$

In the next section, we consider some generalizations of the results that we have presented above.

VI. Extensions and Generalizations

A. Case of Multiple Sensors

It is fairly obvious that the proof techniques of Theorems III.2 and III.3 can be generalized to the case when N sensors are present. We present the following stability result while omitting the proof.

Proposition VI.1. *Consider the process in (1) being observed by N sensors, such that the i-th sensor generates measurements according to the model*

$$\mathbf{y}_i(k) = C_i\mathbf{x}(k) + \mathbf{v}_i(k), \qquad 1 \le i \le N.$$

The sensors transmit data over erasure channels, with the event of erasure in the i-th channel being denoted by $\mathbf{r}_i = \emptyset$. Consider the 2^N possible ways of choosing m out of the N sensors, for all values of m between 0 and N. For the j-th such way, let the sensors chosen be denoted by n_1, n_2, $\cdots$, n_j and sensors not chosen by m_1, m_2, $\cdots$, m_{N-j}. Denote by $\mathcal{C}^j$ the matrix formed by stacking the matrices C_{m_1}, C_{m_2}, $\cdots$, $C_{m_{N-j}}$. Finally, denote by ϱ^j the spectral radius of the unobservable part of matrix A when the pair $(A, \mathcal{C}^j)$ is put in the observer canonical *form. A necessary and sufficient condition for the existence of a positive integer q, an encoding algorithm of either the type $\mathbb{S}_q$ or $\mathbb{S}_q^{NAK}$ and a controller that stabilize the process is that the following 2^N inequalities be satisfied:*

$$Pr\left(\mathbf{r}_{n_1} = \emptyset, \mathbf{r}_{n_2} = \emptyset, \cdots, \mathbf{r}_{n_j} = \emptyset\right) \mid \varrho^j \mid^2 < 1, \qquad 1 \leq j \leq 2^N.$$

The optimal encoding strategy for the class $\mathbb{S}_q^{NAK}$ can also be identified as in the case of 2 sensors. However, it turns out to involve transmitting exponentially (in the number of sensors) increasing amount of data. This observation has also been made in [36].

B. Communication over Networks of Erasure Channels

We can also consider the case when sensors transmit information not over erasure channels, but over networks, in which each link is modeled using the erasure model described above. It is fairly obvious that the algorithms used for proving the necessity of the stabilizability conditions in Theorem III.2 and for proving the sufficiency in Theorem III.3 can be generalized to this case, provided there is a provision for time-stamping the transmitted vectors. As an example, consider the algorithm used to prove sufficiency.

- If the networks connecting the two sensors to the controller are disjoint, each link in the two networks carries two quantities as above.
    1) Sensor 1 calculates and transmits the estimates $\hat{\mathbf{x}}_1^{loc,1}(k)$ and $\hat{\mathbf{x}}_3^{loc,1}(k)$ at every time step. Similarly, sensor 2 calculates and transmits the estimates $\hat{\mathbf{x}}_2^{loc,2}(k)$ and $\hat{\mathbf{x}}_3^{loc,2}(k)$ at every time step. The time-stamps correspond to the latest measurements used in calculating these estimates.
    2) Every node in the network checks the time-stamp of data received over the incoming edges and the estimate in its memory. It chooses the data with the latest time-stamp, transmits it along outgoing links and stores it in the memory for the next time step.
    3) The controller constructs the estimate in the same way as in the two-channel case.
- If the networks share links, however, each link carries four quantities. While the sensors calculate and transmit local estimates, each node in the network transmits data corresponding to the latest received values of all the four estimates: $\hat{\mathbf{x}}_1^{loc,1}(.)$, $\hat{\mathbf{x}}_3^{loc,1}(.)$, $\hat{\mathbf{x}}_2^{loc,2}(.)$ and $\hat{\mathbf{x}}_3^{loc,2}(.)$. Using this data, the controller can calculate the estimate.

Note that the intermediate nodes in the network do not need acknowledgments from the controller.

Moreover, we can use the techniques used in [15] for the case when only one sensor is present and extend the stability conditions to this case. We state the following result without proof.

Proposition VI.2. *Consider the set-up of Figure 1 with the erasure links being replaced by networks in which each link is modeled as a erasure link with given probability of erasure. Let $A \in \mathbb{R}^{n\times n}$, $B \in \mathbb{R}^{n\times l}$, $C_1 \in \mathbb{R}^{m_1\times n}$ and $C_2 \in \mathbb{R}^{m_2\times n}$ be given matrices specifying the state-space representation for the plant. In addition, assume that the plant is observable and controllable and that its state-space representation is of type I or type II. If the state space representation is of type I, then there exists a controller of class $\mathbb{K}$, a positive integer q and sensors of class $\mathbb{S}_q$ or*

$\mathbb{S}_q^{NAK}$ *such that the feedback system is stable if and only if the following inequalities hold*

$$\varrho(A_{1,1})^2 p_{maxcut,2} < 1 \tag{38}$$

$$\varrho(A_{2,2})^2 p_{maxcut,1} < 1, \tag{39}$$

where $\varrho(A_{i,i})$ *represents the spectral radius of the matrix* $A_{i,i}$*. If the state-space representation is of type II then the necessary and sufficient conditions for stabilizability include the following additional inequality:*

$$\varrho(A_{3,3})^2 p_{maxcut,12} < 1. \tag{40}$$

In the above inequalities, the terms $p_{maxcut,i}$ *denote the max-cut probabilities of the network. For the case when the erasure over distinct links are independent events, they can be calculated as follows:*

1) *To calculate* $p_{maxcut,1}$*, form a cut by partitioning the node set of the network connecting sensor 1 and the network into two sets: the source set containing the sensor 1 and the sink set containing the controller. For this cut, consider the edges going from the source set to the sink set and calculate the cut-probability by multiplying the erasure probabilities for these edges. The maximum such cut-probability yields* $p_{maxcut,1}$.
2) *To calculate* $p_{maxcut,2}$*, proceed as above. However, the source set now contains sensor 2 instead of sensor 1.*
3) *To calculate* $p_{maxcut,12}$*, proceed as above. However, the source set now contains both sensor 1 and sensor 2.*

A special case of the network arises when each sensor transmits data over a single link to the controllers. However, in addition, the sensors can cooperate by communicating with each other over a link. If the link is perfect (i.e., does not exhibit erasure), then the two sensors, in effect, form one sensor and the results of [16] apply. However, if this link also exhibits erasure, then we obtain the following stability conditions:

Corollary VI.3 (Sensors cooperating over a erasure link)**.** *Consider the set-up of Figure 1 with an additional bidirectional link connecting the two sensors. Let* $A \in \mathbb{R}^{n\times n}$, $B \in \mathbb{R}^{n\times l}$, $C_1 \in \mathbb{R}^{m_1\times n}$ *and* $C_2 \in \mathbb{R}^{m_2\times n}$ *be given matrices specifying the state-space representation for the plant. In addition, assume that the plant is observable and controllable and that its state-space representation is of type I or type II. Let the event of erasure over the link connecting sensor 1 to the controller be denoted as before by* $\mathbf{r}_1(k) = \emptyset$*, over the link connecting sensor 2 to the controller by* $\mathbf{r}_2(k) = \emptyset$*, and over the link connecting the two sensors by* $\mathbf{r}_3(k) = \emptyset$*. If the state space representation is of type I, then there exists a controller of class* $\mathbb{K}$*, a positive integer* q *and sensors of class* $\mathbb{S}_q$ *or* $\mathbb{S}_q^{NAK}$ *such that the feedback system is stable if and only if the following inequalities hold*

$$\varrho(A_{2,2})^2 \max\left(Pr(\mathbf{r}_1(k) = \emptyset), Pr(\mathbf{r}_2(k) = \emptyset, \mathbf{r}_3(k) = \emptyset)\right) < 1 \tag{41}$$

$$\varrho(A_{1,1})^2 \max\left(Pr(\mathbf{r}_2(k) = \emptyset), Pr(\mathbf{r}_1(k) = \emptyset, \mathbf{r}_3(k) = \emptyset)\right) < 1, \tag{42}$$

where $\varrho(A_{i,i})$ represents the spectral radius of the matrix $A_{i,i}$. If the state-space representation is of type II then the necessary and sufficient conditions for stabilizability include the following additional inequality:

$$\varrho(A_{3,3})^2 Pr(\mathbf{r}_1(k) = \emptyset, \mathbf{r}_2(k) = \emptyset) < 1. \tag{43}$$

C. Performance Analysis in the Presence of Noisy Acknowledgments

While we were able to propose an algorithm that achieves the optimal performance in the presence of acknowledgments, there are no such guarantees for the algorithm $\mathcal{A}_{nack}$. The problem of identifying the algorithm that achieves the optimal performance when acknowledgments are not available is largely open and the solution is known only for special cases. As an example, the case when there is no process noise entering the system in (1) was considered in [44]. Similarly, the case when one of the links does not exhibit erasure was considered in [16]. The algorithm $\mathcal{A}_{ack}$ can be extended to another such special case. Consider the arrangement in which the sensors are of class $\mathbb{S}_{NAK}$, i.e., acknowledgments are not available, however, the two sensors transmit information over the same channel. Thus, at each time step, either both sensors transmit successfully or they both suffer erasure. In this case, we can state the following.

Proposition VI.4. *Consider a process evolving as in (28) being observed by two sensors of the form (3) that transmit information to an encoder over erasure links of the type described in Definition II.1. Further, let the erasure in the links be described by a single variable, so that $\mathbf{r}_1(k) = \mathbf{r}_2(k)$ at all time steps. Let $A \in \mathbb{R}^{n \times n}$, $B \in \mathbb{R}^{n \times l}$, $C_1 \in \mathbb{R}^{m_1 \times n}$ and $C_2 \in \mathbb{R}^{m_2 \times n}$ be given matrices specifying the state-space representation for the plant. Suppose that the following encoding algorithm is carried out at every time step.*

1) *Each sensor i calculates the information vector $I_{i,k,k}(k)$.*
2) *The decoder maintains a variable $\hat{\mathbf{x}}^{dec}(k)$ with the initial value $\hat{\mathbf{x}}^{dec}(-1) = 0$. If $\mathbf{r}_1(k) = \mathbf{r}_2(k) = \emptyset$, the variable is updated as*
$$\hat{\mathbf{x}}^{dec}(k) = A\hat{\mathbf{x}}^{dec}(k-1),$$
otherwise, it is updated as
$$\hat{\mathbf{x}}^{dec}(k) = \mathbf{P}(k|k,k)\left(I_{1,k,k}(k) + I_{2,k,k}(k)\right),$$
where $\mathbf{P}(k|k,k)$ is the mmse error covariance in estimating $\mathbf{x}(k)$ evolving as in (1) given the variables $\mathbf{y}_1(0)$, $\cdots$, $\mathbf{y}_1(k)$, $\mathbf{y}_2(0)$, $\cdots$, $\mathbf{y}_2(k)$.

Then, $\hat{\mathbf{x}}^{dec}(k) = \hat{\mathbf{x}}\left(k|\mathcal{I}_1^{\max}(k), \mathcal{I}_2^{\max}(k)\right)$, where $\mathcal{I}_i^{\max}(k)$'s are the maximal information sets.

Proof: Omitted for space constraints. □

In general, however, when the links' erasure events are independent and noiseless acknowledgments are not available at every time step, the optimal algorithm is not known. We now consider the case when acknowledgments are transmitted over a erasure link. We propose an algorithm whose performance improves as the probability of the acknowledgment erasure decreases (for given probabilities of erasure in the channels from the sensors to the decoder) or as the probability of erasures decreases (for a given probability of acknowledgment erasure). The algorithm still requires a constant amount of transmission and memory at the encoders, even though the memory requirement at the decoder can be high if one of the channels is much worse than the other.

Denote the stochastic event of acknowledgment being received at the sensors at time step k by $\mathbf{r}_{ack}(k) = \mathbf{1}$ and the event of acknowledgment being not received by $\mathbf{r}_{ack}(k) = \emptyset$. We shall once again consider the state space representation of the system to be partitioned as in either (36) or (37) according to the process being of **type I** or **type II**. The algorithm proceeds as follows. At each time step k

- *Encoder for Sensor 1:* The encoder for sensor 1 maintains an estimate $\hat{\mathbf{t}}_2(k)$ of the last time step at which sensor 2 was able to communicate to the decoder. The estimate is updated as

$$\hat{\mathbf{t}}_2(k) = \begin{cases} k-1 & \text{if } \mathbf{r}_2(k-1) = \mathbf{1} \text{ and } \mathbf{r}_{ack}(k-1) = \mathbf{1} \\ \hat{\mathbf{t}}_2(k-1) & \text{otherwise,} \end{cases}$$

  with $\hat{\mathbf{t}}_2(0) = -1$. Similar to the algorithm $\mathcal{A}_{ack}$, the encoder then calculates two quantities: $I_{1,k,k}(k)$ and $I_{1,k,\hat{\mathbf{t}}_2(k)}(k)$. Note that both these quantities can be calculated recursively. Finally, the sensor also calculates the estimate $\hat{\mathbf{x}}_2^{loc,1}(k)$ of the modes $\mathbf{x}_2(k)$ of the process using its local measurements $\mathbf{y}_1(0)$, $\mathbf{y}_1(1)$, $\cdots$, $\mathbf{y}_1(k)$. The sensor then transmits four quantities: $\hat{\mathbf{t}}_2(k)$, $I_{1,k,k}(k)$, $I_{1,k,\hat{\mathbf{t}}_2(k)}(k)$ and $\hat{\mathbf{x}}_2^{loc,1}(k)$.
- *Encoder for Sensor 2:* In a similar fashion, the encoder for the sensor 2 calculates and transmits the quantities $\hat{\mathbf{t}}_1(k)$, $I_{2,k,k}(k)$, $I_{2,\hat{\mathbf{t}}_1(k),k}(k)$ and $\hat{\mathbf{x}}_1^{loc,2}(k)$.
- *Decoder:*
  - The decoder maintains two sets $\mathcal{S}_1$ and $\mathcal{S}_2$ that are updated as follows. If $\mathbf{r}_1(k) = \mathbf{1}$, then take the following actions:

$$\begin{aligned} \mathcal{S}_1(k) &= \{I_{2,k,\hat{\mathbf{t}}_2}(k)\} \\ \mathcal{S}_2(k) &= \mathcal{F}_{2,k}\left(\mathcal{S}_2(k-1)\right) \bigcup I_{1,k,k}(k). \end{aligned}$$

    Similarly, if $\mathbf{r}_2(k) = \mathbf{1}$, then take the following actions:

$$\begin{aligned} \mathcal{S}_2(k) &= \{I_{1,\hat{\mathbf{t}}_1(k),k}(k)\} \\ \mathcal{S}_1(k) &= \mathcal{F}_{1,k}\left(\mathcal{S}_1(k-1)\right) \bigcup I_{2,k,k}(k). \end{aligned}$$

    In the above equations, the initial conditions are $\mathcal{S}_1(-1) = \emptyset$ and $\mathcal{S}_2(-1) = \emptyset$. The operation $\mathcal{F}_{2,k}(S)$ for a set S denotes that every element of S that is of the type $I_{1,j,k-1}(k-1)$ is converted to $I_{1,j,k}(k)$. The operation $\mathcal{F}_{1,k}(.)$ is similarly defined. It is easily verified that the sets $\mathcal{S}_1(k)$ and $\mathcal{S}_2(k)$ only have elements of the type $I_{2,j,k}(k)$ and $I_{1,k,j}(k)$ respectively.
  - The decoder also maintains two variables that it updates as follows.

$$\alpha(k) = \begin{cases} \hat{\mathbf{t}}_2(k) & \text{if } \mathbf{r}_1(k) = \mathbf{1} \\ \alpha(k-1) & \text{otherwise.} \end{cases}$$

$$\beta(k) = \begin{cases} \hat{t}_1(k) & \text{if } \mathbf{r}_2(k) = \mathbf{1} \\ \beta(k-1) & \text{otherwise.} \end{cases}$$

The initial values are $\alpha(0) = \beta(0) = 0$. Once again, it may be verified that the sets $\mathcal{S}_1(k)$ and $\mathcal{S}_2(k)$ at any time step k contain the elements $I_{2,k,\alpha(k)}(k)$ and $I_{1,\beta(k),k}(k)$ respectively.

– The decoder faces one of four possibilities.

1) If $\mathbf{r}_1(k) = \mathbf{1}$ and $\mathbf{r}_2(k) = \mathbf{1}$, the decoder calculates the estimate as

$$\hat{\mathbf{x}}^{dec}(k) = \mathbf{P}(k|k,k)\left(I_{1,k,k}(k) + I_{2,k,k}(k)\right),$$

where $\mathbf{P}(k|k,k)$ is the mmse error covariance in estimating $\mathbf{x}(k)$ evolving as in (28) given the variables $\mathbf{y}_1(0), \cdots, \mathbf{y}_1(k), \mathbf{y}_2(0), \cdots, \mathbf{y}_2(k)$.

2) If $\mathbf{r}_1(k) = \emptyset$ and $\mathbf{r}_2(k) = \emptyset$, the decoder updates its previous estimate as

$$\hat{\mathbf{x}}^{dec}(k) = A\hat{\mathbf{x}}^{dec}(k-1).$$

3) If $\mathbf{r}_1(k) = \mathbf{1}$ and $\mathbf{r}_2(k) = \emptyset$, the decoder calculates

$$\hat{\mathbf{x}}^{dec,t}(k) = \mathbf{P}(k|k,\alpha(k))\left(I_{1,k,\alpha(k)}(k) + I_{2,k,\alpha(k)}(k)\right),$$

where the term $I_{1,k,\alpha(k)}(k)$ has been transmitted by the sensor 1 and the term $I_{1,k,\alpha(k)}(k)$ can be obtained from the set $\mathcal{S}_2(k)$. The estimates for the modes $\mathbf{x}_2(k)$ and $\mathbf{x}_3(k)$ can now be isolated from $\hat{\mathbf{x}}^{dec,t}(k)$ by accessing the last $n_2 + n_3$ elements. Also, the decoder uses the estimate $\hat{\mathbf{x}}^{dec}(k-1)$ to evaluate the estimate of the modes $\mathbf{x}_1(k-1)$ by isolating the first n_1 components. It can then calculate

$$\hat{\mathbf{x}}_1(k) = A_{1,1}\hat{\mathbf{x}}_1(k-1),$$

and concatenate this estimate of the modes $\mathbf{x}_1(k)$ with the estimates for the modes $\mathbf{x}_2(k)$ and $\mathbf{x}_3(k)$ to obtain $\hat{\mathbf{x}}^{dec}(k)$.

4) The case when $\mathbf{r}_2(k) = \mathbf{1}$ and $\mathbf{r}_1(k) = \emptyset$ can be treated similarly.

Note that when $\mathbf{r}_{ack}(k) \equiv \emptyset$ at all time steps, the above algorithm is a version of the algorithm $\mathcal{A}_{nack}$ presented earlier. As $k \to \infty$, it is fairly obvious that the algorithm yields optimal performance for the case when acknowledgments are always available. As noted earlier, The stability conditions are unaffected by the availability of acknowledgments.

D. Markov Drops

While the algorithm $\mathcal{A}_{ack}$ was optimal for arbitrary realizations of the erasure process, the stability analysis so far assumed that erasure events were i.i.d.. This condition can be relaxed. A popular model for the bursty nature of packet drops in a wireless channel is according to a Markov chain. The simplest such model is the classical Gilbert-Elliot channel model. In this model, the channel is assumed to exist in one of two possible modes: state 0 corresponding to a packet drop and state 1 corresponding to no packet drop. The channel transitions between the two states according to a Markov chain. Suppose that erasures in the each of the two links in our model be described by such a Markov chain. Let the variable $\mathbf{r}_1(k)$ be governed by a Markov chain with transition probability matrix Q_1 and the variable $\mathbf{r}_2(k)$ by a Markov chain with transition probability matrix Q_2. Further, for simplicity, let the erasure events at the two links be independent. We have the following result.

Proposition VI.5. *(***Necessary and sufficient conditions for stabilizability for Markovian packet drops***) Consider the set-up of Figure 1 and let $A \in \mathbb{R}^{n\times n}$, $B \in \mathbb{R}^{n\times l}$, $C_1 \in \mathbb{R}^{m_1\times n}$ and $C_2 \in \mathbb{R}^{m_2\times n}$ be given matrices specifying the state-space representation for the plant. In addition, assume that the plant is observable and controllable and that its state-space representation is of type I or type II. In addition, let the statistics of the erasure links 1 and 2 be described by Markov chains with transition probability matrices Q_1 and Q_2 respectively, with the element $q_{00,i}$ denoting the probability of two consecutive erasures in the i-th link. Finally, let the erasures over the two channels be independent. If the state space representation is of type I, then there exists a controller of class $\mathbb{K}$, a positive integer q and sensors of class $\mathbb{S}_q$ or $\mathbb{S}_q^{NAK}$ such that the feedback system is stable if and only if the following inequalities hold*

$$\varrho(A_{1,1})^2 q_{00,2} < 1 \tag{44}$$

$$\varrho(A_{2,2})^2 q_{00,1} < 1 \tag{45}$$

where $\varrho(A_{i,i})$ represents the spectral radius of the matrix $A_{i,i}$. If the state-space representation is of type II then stability is assured if and only if the following additional inequality also holds:

$$\varrho(A_{3,3})^2 q_{00,1} q_{00,2} < 1 \tag{46}$$

VII. Conclusions and Future Work

In this paper, we considered the problem of controlling a plant using measurements from multiple sensors. The information from the sensors to the controller is transmitted over links where erasure (data loss) is governed by a stochastic process. We identified necessary and sufficient conditions for the stabilizability of a linear and time-invariant plant, in a mean square sense. The allowable stabilization policies at the sensors are constrained to place vectors of constant dimension for possible transmission over the erasure links. Under the assumption that the controller is able to transmit acknowledgments back to the sensors, we identified an encoding algorithm that minimizes a quadratic cost. We also considered various extensions, such as when sensors are able to co-operate over an erasure link, and/or when acknowledgments are transmitted from the controller to the sensors via an erasure channel. In this paper, we have also proved a multi-sensor version of the separation principle, which makes our results relevant also to sensor fusion problems.

There are various directions in which the present work may be extended. Our algorithm can be extended for the case when more than 2 sensors are present; however, this requires an exponentially increasing size of the memory required at the encoders and of the transmission required. It would be interesting to explore algorithms that require less data to be transmitted, possibly with a more relaxed notion of optimal performance. It would also be important to compute the performance of the sub-optimal encoding algorithm that was proposed in Section VI-C. We are currently also looking at finding the optimal encoding algorithms for other channels such as AWGN or discrete memoryless channels.

ACKNOWLEDGMENTS

The first author would like to thank Amir F. Dana for helpful discussions. The second author would like to thank Richard La (UMD) for a discussion that motivated the extension of our work so as to address noisy acknowledgments. The material is based upon the work supported by the National Aeronautics and Space Administration cash university support for the first author. The second has been supportted by the NSF (CAREER) award ECCS 0644764.

Any opinions, findings and conclusions or recommendations expressed in this material are those of the author(s) and do not necessarily reflect the views of the National Aeronautics and Space Administration.

APPENDIX

Proof of Proposition III.1 Consider that we are given a linear and time-invariant system with the properties specified in the statement of the Proposition and whose state-space representation is specified by matrices $A^{s0} \in \mathbb{R}^{n\times n}$, $B^{s0} \in \mathbb{R}^{n\times l}$, $C_1^{s0} \in \mathbb{R}^{m_1\times n}$ and $C_2^{s0} \in \mathbb{R}^{m_2\times n}$. Here C_1^{s0} and C_2^{s0} represent a bipartition of the output and $s0$ stands for *stage zero*. Below we outline a procedure, comprising three stages, that will lead to an equivalent state-space of type I or type II, as defined in the statement of the Proposition.

First stage Since, from Assumption 1, the system is not observable from $y_1(t)$ alone, or equivalently the pair $(A^{(s0)}, C_1^{(s0)})$ is not observable, we can use the *canonical structure theorem* [34, page 340, eq. (22)] to conclude that there exists a transformation $P^{0\to 1} \in \mathbb{R}^{n\times n}$ such that the matrices $A^{(s1)} \overset{def}{=} P^{0\to 1}A^{(s0)}\left(P^{0\to 1}\right)^{-1}$ and $C_1^{(s1)} \overset{def}{=} C_1^{(s0)}\left(P^{0\to 1}\right)^{-1}$ have the following structure:

$$A^{(s1)} = \begin{bmatrix} A_{1,1}^{(s1)} & A_{1,2}^{(s1)} \\ \mathbf{0}^{n_1'\times n_1} & A_{2,2}^{(s1)} \end{bmatrix} \tag{47}$$

$$C_1^{(s1)} = \begin{bmatrix} \mathbf{0}^{m_1\times n_1} & C_{1,2}^{(s1)} \end{bmatrix} \tag{48}$$

where $n_1 + n_1' = n$, $A_{2,2}^{(s1)} \in \mathbb{R}^{n_1'\times n_1'}$ and $C_{1,2}^{(s1)} \in \mathbb{R}^{m_1\times n_1'}$. Notice that n_1 is a strictly positive integer because the pair $(A^{(s0)}, C_1^{(s0)})$ is not observable. The remaining matrices defining the new state-space representation are given by $C_2^{(s1)} \overset{def}{=} C_2^{(s0)}\left(P^{0\to 1}\right)^{-1}$ and $B^{(s1)} \overset{def}{=} P^{0\to 1}B^{(s0)}$.

Second stage We start by partitioning $C_2^{(s1)}$ in the following way:

$$C_2^{(s1)} = \begin{bmatrix} C_{1,1}^{(s1)} & C_{1,2}^{(s1)} \end{bmatrix} \tag{49}$$

where $C_{1,2}^{(s1)} \in \mathbb{R}^{m_2\times n_1'}$. We can now apply, once again, the *canonical structure theorem* [34, page 340, eq. (22)] to show the existence of a transformation $P^{1\to 2} \in \mathbb{R}^{n_1'\times n_1'}$ such that the matrix $C_{2,2}^{(s2)} \overset{def}{=} C_{2,2}^{(s1)}\left(P^{1\to 2}\right)^{-1}$ features one of the following structures below:

$$C_{2,2}^{(s2)} = \begin{cases} \begin{bmatrix} \mathbf{0}^{m_2\times n_2} & C_{2,3}^{(s3)} \end{bmatrix},\ C_{2,3}^{(s3)} \in \mathbb{R}^{m_2\times n_3},\ n_2+n_3=n_1' & \text{if } C_{2,2}^{(s2)} \text{ is not zero (type II)} \\ \mathbf{0}^{m_2\times n_1'} & \text{otherwise (type I)} \end{cases} \tag{50}$$

Similarly, for the matrix $A_{2,2}^{(s2)} \stackrel{def}{=} P^{0\to 1}A^{(s1)}\left(P^{0\to 1}\right)^{-1}$ the following holds:

$$(\text{system is type II}) \implies A_{2,2}^{(s2)} = \begin{bmatrix} A_{2,2}^{(s3)} & A_{2,3}^{(s3)} \\ \mathbf{0}^{n_3\times n_2} & A_{3,3}^{(s3)} \end{bmatrix} \tag{51}$$

where $A_{2,2}^{(s3)} \in \mathbb{R}^{n_2\times n_2}$ and $A_{3,3}^{(s3)} \in \mathbb{R}^{n_3\times n_3}$. Recall that according to Assumption 1 the system is not observable from $\mathbf{y}_2$ alone, implying that n_2 is a nonzero positive integer.

Third stage Consider the following state transformation:

$$P \stackrel{def}{=} \begin{bmatrix} I & \mathbf{0}^{n_1\times n_1'} \\ \mathbf{0}^{n_1'\times n_1} & P^{1\to 2} \end{bmatrix} P^{0\to 1} \tag{52}$$

We can now use the previous analysis to show, by inspection, that the state-space representation given by the matrices $A^{(s3)} \stackrel{def}{=} PA^{(s0)}P^{-1}$, $B^{(s3)} \stackrel{def}{=} PB^{(s0)}$, $C_1^{(s3)} \stackrel{def}{=} C_1^{(s0)}P^{-1}$ and $C_2^{(s3)} \stackrel{def}{=} C_2^{(s0)}P^{-1}$ is in the form specified by (10)-(12) if the system is type I and that otherwise the matrices will have the structure (13)-(15). $\square$

REFERENCES

[1] B.D.O. Anderson and J.B. Moore, "Time-Varying Feedback Laws for Decentralized Control," IEEE Transactions on Automatic Control, Vol AC-26, No. 5, 1981, pp. 1133-1139

[2] P. Antsaklis and J. Baillieul, "Special issue on networked control systems," *IEEE Transactions on Automatic Control*, vol. 49, no. 9, 2004.

[3] B. Azimi-Sadjadi, "Stability of networked control systems in the presence of packet losses," in *Proceedings of 2003 IEEE Conference on Decision and Control*, Dec 2003.

[4] Y. Bar-Shalom, *"On the track-to-track correlation problem"*, IEEE Trans. Automat. Contr., AC-26(2):571-572, 1981.

[5] Y. Bar-Shalom and L. Campo, *The effect of the common process noise on the two-sensor fused-track covariance*, IEEE Trans. Aerospace and Electronic Systems, AES-22(6):803-805, 1986.

[6] L. Bushnell, "Special issue on networks and control," *IEEE Control Systems Magazine*, vol. 21, no. 1, February 2001.

[7] N. A. Carlson, *Federated square root filter for decentralized parallel processes*, IEEE Trans. Aerospace and Electronic Systems, AES-26(3):517-525, 1990.

[8] K. C. Chang, R. K. Saha and Y. Bar-Shalom, *On optimal track-to-track fusion*, IEEE Trans. Aerospace and Electronic Systems, AES-33(4):1271-1276, 1997.

[9] C. Y. Chong, *"Hierarchical Estimation"*, Proc. 2nd MIT/ONR C^3 Workshop, 1979.

[10] C. Y. Chong, S. Mori and K. C. Chang, *"Information Fusion in Distributed Sensor Networks"*, Proc. American Control Conf., pp 830-835, 1985.

[11] D. Georgiev and D. M. Tilbury, *"Packet-based Control: The H_2-optimal Solution"*, Automatica, 42(1), January 2006, pp. 137-144.

[12] S. Graham, G. Baliga and P. R. Kumar, *"Issues in the convergence of control with communication and computing: Proliferation, architecture, design, services, and middleware"*, Proceedings of the 43rd IEEE Conference on Decision and Control, Bahamas, 2004, pp. 1466-1471.

[13] V. Gupta, S. Adalakha and B. Sinopoli, *"Towards Receding Horizon Networked Control"*, Automatica, Submitted. See also V. Gupta, S. Adalakha and B. Sinopoli, *"Towards Receding Horizon Networked Control"*, Allerton conference on communication, control and computing, 2006.

[14] V. Gupta, T. Chung, B. Hassibi and R. M. Murray, "On a Stochastic Sensor Selection Algorithm with Applications in Sensor Scheduling and Dynamic Sensor Coverage," *Automatica*, Vol. 42, No. 2, February 2006, Pages: 251-260.

[15] V. Gupta, A. F. Dana, J. P. Hespanha, R. M. Murray and B. Hassibi, "Data Transmission over Networks for Estimation and Control," *IEEE Transactions on Automatic Control*, 2006, Submitted.

[16] V. Gupta, B. Hassibi, and R. M. Murray, "Optimal LQG control across a packet-dropping link," *Systems and Controls Letters*, 2007, To appear. See also, V. Gupta, D. Spanos, B. Hassibi, and R. M. Murray, "Optimal LQG control across a packet-dropping link," *Proc. of the American Control Conference (ACC)*, 2006.

[17] C. N. Hadjicostis and R. Touri, "Feedback control utilizing packet dropping network links," in *Proc. of the IEEE Conference on Decision and Control*, 2002.

[18] H. R. Hashemipour, S. Roy and A. J. Laub, *"Decentralized structures for parallel Kalman filtering"*, IEEE Trans. Automat. Contr., AC-33(1): 88-94, 1988.

[19] A. Hassibi, S. P. Boyd, and J. P. How, "Control of asynchronous dynamical systems with rate constraints on events," in *Proc. IEEE Conf. Decision and Control*, Dec 1999, pp. 1345–1351.

[20] B. Hassibi, A. H. Sayed, and T. Kailath, *Indefinite-Quadratic Estimation and Control*. Studies in Applied and Numerical Mathematics, 1999.

[21] J. P. Hespanha, P. Naghshtabrizi and Y. Xu, "Networked Control Systems: Analysis and Design," To appear in the Proc. of IEEE, Special Issue on Networked Control Systems, 2007.

[22] O.C. Imer, S. Yuksel and T. Basar, *"Optimal Control of LTI Systems over Communication Networks"*, Automatica, 42(9):1429-1440, September 2006.

[23] P. A. Kawka and A. G. Alleyne, *"Stability and performance of packet-based feedback control over a Markov channel"*, Proceedings of the American Control Conference, 2006, pp. 2807-2812.

[24] B. C. Levy, D. A. Castanon, G. C. Verghese and A. S. Willsky, *"A scattering framework for decentralized estimation problem"*, Automatica, 19(4): 373-384, 1983.

[25] F.L. Lian, J.R. Moyne and D.M. Tilbury, *"Performance evaluation of control networks"*, IEEE Control Systems Magazine 21 (2001) (1), pp. 6683.

[26] Q. Ling and M. D. Lemmon, "Power spectral analysis of netwroked control systems with data droputs," *IEEE Transactions on Automatic control*, vol. 49, no. 6, pp. 955–960, June 2004.

[27] N. C. Martins, M. .A. Dahleh, N Elia,"Feedback stabilization of uncertain systems in the presence of a direct link," IEEE Transactions on Automatic Control, Volume 51, Issue 3, March 2006 Page(s): 438-447

[28] A.S. Matveev, A.V. Savkin, "Decentralized Stabilization of Linear Systems via Limited Capacity Communication Networks," Proceedings of the IEEE Conference on Decision and Control, 2005 Page(s):1155 - 1161

[29] L. A. Montestruque and P. Antsaklis, *"On the Model-based Control of Networked Systems"*, Automatica, 39(10), October 2003, pp. 1837-1843.

[30] P. Naghshtabrizi and J. Hespanha, *"Anticipative and Non-anticipative Controller Design for Network Control Systems"*, In Panos J. Antsaklis, Paulo Tabuada, Networked Embedded Sensing and Control, volume 331 of Lect. Notes in Contr. and Inform. Sci., pages 203218, 2006.

[31] G.N. Nair, R.J. Evans, P.E. Caines," Stabilising Decentralised Linear Systems under Data Rate Constraints," Proceedings of the IEEE Conference on Decision and Control, 2004, pp. 3992-3997

[32] J. Nilsson, "Real-time control systems with delays," Ph.D. dissertation, Department of Automatic Control, Lund Institute of Technology, 1998.

[33] B. S. Rao and H. F. Durrant-Whyte, *"Fully decentralized algorithm for multi-sensor Kalman filtering"*, IEE proceedings, 138(5), 1991.

[34] W.J. Rugh, "Linear System Theory", Prentice Hall Information and Systems Sciences Series, 2nd eddition, 1996

[35] A. Sahai; S Mitter;"The Necessity and Sufficiency of Anytime Capacity for Stabilization of a Linear System Over a Noisy Communication LinkPart I: Scalar Systems," IEEE Transactions on Information Theory, Volume 52, Issue 8, Aug. 2006 Page(s):3369 - 3395

[36] V. Saligrama and D. A. Castanon, "Reliable Distributed Estimation with Intermittent Communications," *Proc. of the IEEE Conference on Decision and Control (CDC)*, 2006, Pages: 6763-6768.

[37] L. Schenato, B. Sinopoli, M. Franceschetti, K. Poolla and S. S. Sastry, *"Foundations of Control and Estimation over Lossy Networks"*, Proc. of the IEEE, To Appear 2006.

[38] P. J. Seiler, "Coordinated control of unmanned aerial vehicles," Ph.D. dissertation, University of California, Berkeley, 2002.

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

[40] S.C. Smith and P. Seiler, *Estimation with lossy measuraments: jump estimators for jump systems*. IEEE Transaction on Automatic Control, 48(12):1453-1464, 2003.

[41] J. L. Speyer, *"Computation and transmission requirements for a decentralized linear-quadratic Gaussian control problem"*, IEEE Trans. Automat. Contr., AC-24, 1979, pp. 266-269.

[42] S Tatikonda, "Some scaling properties of larger distributed control systems,"Proceedings of the IEEE Conference of Decision and Control, 2003, pp. 3142-3147

[43] D. Willner, C. B. Chang and K. P. Dunn, *"Kalman Filter Algorithms for a Multisensor System"*, Proc. IEEE Conference on Decision and Control, 1976, pp. 570-574.

[44] A. Willsky, M. Bello, D. Castanon, B. Levy and G. Verghese "Combining and updating of local estimates and regional maps along sets of one-dimensional tracks," *IEEE Transactions on Automatic Control* Volume 27, Issue 4, Aug 1982 Page(s):799 - 813.

[45] S Yuksel and T Basar, "Optimal Signaling Policies for Decentralized Multi-Controller Stabilizability over Communication Channels," accepted for publication in the IEEE Transactions on Automatic Control

[46] W. Zhang, M. S. Branicky, and S. M. Philips, "Stability of networked control systems," *IEEE Control System Magazine*, vol. 21, no. 1, pp. 84–89, Feb 2001.