Optimal scheduling of multiple sensors which transmit measurements over a dynamic lossy network

Johnson Carroll, Hassan Hmedi, and Ari Arapostathis

***Abstract*— Motivated by various distributed control applications, we consider a linear system with Gaussian noise observed by multiple sensors which transmit measurements over a dynamic lossy network. We characterize the stationary optimal sensor scheduling policy for the finite horizon, discounted, and long-term average cost problems and show that the value iteration algorithm converges to a solution of the average cost problem. We further show that the suboptimal policies provided by the rolling horizon truncation of the value iteration also guarantee geometric ergodicity and provide near-optimal average cost. Lastly, we provide qualitative characterizations of the multidimensional set of measurement loss rates for which the system is stabilizable for a static network, significantly extending earlier results on intermittent observations.**

I. Introduction

Distributed systems with multiple sensors require control of both the system as well as the scheduling of observations. This work addresses a system with both intermittent observations and multiple sensors.

A fundamental problem with distributed sensing is accounting for the possibility of lost or intermittent measurements. In the seminal work of [1], it was shown that for a discrete time linear system with appropriate Gaussian noise, the error covariance is bounded provided the measurement loss rate is below a particular critical value. A number of additional studies have sought to further characterize the behavior of the error covariance for particular systems [2], or with additional assumptions [3]–[7].

Sensor schedules aim to maintain system stability while optimizing system performance. Some approaches scheduled sensor transmissions randomly according to a predetermined (possibly random) schedule [8]–[10]. Dynamic sensor scheduling, based on the information available to the scheduler, can lead to significantly better performance but is of course more complex [11]–[14].

The intersection of these two areas, namely optimal sensor scheduling with intermittent network links, has been largely neglected. Among the few papers in the literature we cite [15]–[19].

This work was supported in part by the Army Research Office through grant W911NF-17-1-001, in part by the National Science Foundation through grant DMS-1715210, and in part by Office of Naval Research through grant N00014-16-1-2956 and was approved for public release under DCN #43-4982-19.

J. Carroll is with the Faculty of Engineering and the Built Environment, University of Johannesburg, Johannesburg, 2006, South Africa (e-mail: jcarroll@uj.ac.za)

‡ H Hmedi and A. Arapostathis are with the Electrical and Computer Engineering Department, 2501 Speedway, EER 7.824, The University of Texas at Austin, Austin, TX 78712 (e-mail: {hmedi,ari}@utexas.edu).

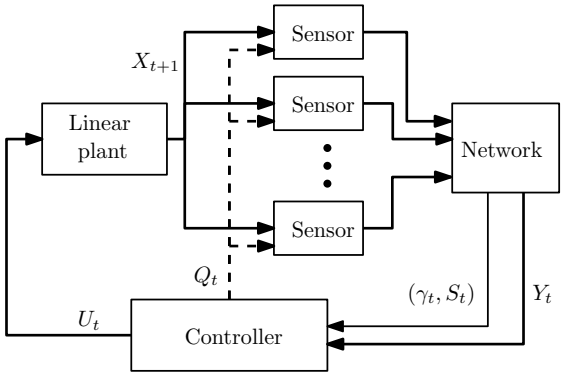

Fig. 1. Overview of the system detailed in Section II. An observation Y_t is lost ($\gamma_t = 0$) with probability λ_t, which depends on which sensor is queried (Q_t) and network state (S_t).

In this work, we consider a discrete-time linear quadratic Gaussian (LQG) system observed by a finite number of sensors. When queried, a sensor attempts to transmit the measurement to the controller over a noisy network which intermittently loses the measurement. Further, the network has its own query-dependent stochastic dynamics, allowing for complex congestion models. A diagram of the system is shown in Figure 1. We make only mild assumptions on the system structure and assume that the system is stabilizable. This rather basic assumption of stabilizability enables us to derive a wealth of interesting new results:

- A stationary, average-cost optimal policy exists, and under that policy the system is geometrically ergodic.
- The value iteration algorithm (VI) converges. In addition, after finitely many steps, the sub-optimal policies calculated via the VI render the system geometrically ergodic, and the induced average cost converges geometrically to the optimal average cost.
- Additionally, we show that a special case of our results generalizes the original stabilizability results of [1] to the case of multiple scheduled sensors with unique loss rates.

Section II describes the system structure, our key assumptions, and some basic results on the Kalman filtering part of the problem. The optimal control problems and results are presented in Section III, and Section IV contains the results on the convergence of the value iteration algorithm. An important special case is discussed in Section V. The proofs could not be included due to the limitation in the number of pages. For these please consult the unabridged version at http://users.ece.utexas.edu/∼ari/Papers/Sensors.

A. Notation

The letter d refers to the dimension of the state space. We let $\mathcal{M}_0^+$ ($\mathcal{M}^+$) denote the cone of real symmetric, positive semi-definite (positive definite) $d \times d$ matrices. For a matrix $G \in \mathcal{M}^+$, $\underline{\sigma}(G)$ and $\overline{\sigma}(G)$ denote the smallest and largest eigenvalues of G, respectively. Recall that the trace of a matrix, denoted by $\mathrm{tr}(\cdot)$, acts as a norm on $\mathcal{M}_0^+$. For $\Sigma_1, \Sigma_2 \in \mathbb{R}^{d\times d}$, we write $\Sigma_1 \preceq \Sigma_2$ when $\Sigma_2 - \Sigma_1 \in \mathcal{M}_0^+$ or $\Sigma_1 \prec \Sigma_2$ when $\Sigma_2 - \Sigma_1 \in \mathcal{M}^+$. For two real vectors λ, ϕ indexed by some set I, we say $\lambda \le \phi$ or $\lambda < \phi$ if for each $i \in I$, $\lambda_i \le \phi_i$ or $\lambda_i < \phi_i$, respectively. A function $f\colon \mathcal{M}_0^+ \to \mathbb{R}$ is concave if for $\Sigma_1, \Sigma_2 \in \mathcal{M}_0^+$,

$$f((1-\beta)\Sigma_1 + \beta\Sigma_2) \;\ge\; (1-\beta)f(\Sigma_1) + \beta f(\Sigma_2) \tag{1}$$

for all $\beta \in [0,1]$. Concavity for functions $f\colon \mathcal{M}_0^+ \to \mathcal{M}_0^+$ is defined in the same way, but replacing the inequality in (1) with the ordering $\succeq$. We also denote a normal distribution with mean x and covariance matrix Σ as $\mathcal{N}(x,\Sigma)$. Given a strictly positive real function f on $\mathbb{S} \times \mathcal{M}_0^+$, where $\mathbb{S}$ is a finite set, the f-norm of a function $g\colon \mathbb{S}\times\mathcal{M}_0^+ \to \mathbb{R}$ is given by

$$\|g\|_f \coloneqq \sup_{(s,\Sigma)\in\mathbb{S}\times\mathcal{M}_0^+} \frac{|g(s,\Sigma)|}{f(s,\Sigma)}\,.$$

We denote by $\mathcal{O}(f)$ the set of real-valued functions on $\mathbb{S} \times \mathcal{M}_0^+$ which have finite f-norm and are continuous, concave, and non-decreasing in the second argument.

II. System, Sensor, and Network Model

We consider a linear quadratic Gaussian (LQG) system

$$\begin{aligned} X_{t+1} \;&=\; AX_t + BU_t + DW_t\,, \quad t \ge 0 \\ X_0 \;&\sim\; \mathcal{N}(x_0, \Sigma_0)\,, \end{aligned} \tag{2}$$

where $X_t \in \mathbb{R}^d$ is the system state, $U_t \in \mathbb{R}^{d_u}$ is the control, and $W_t \in \mathbb{R}^{d_w}$ is a white noise process. We assume that each $W_t \sim \mathcal{N}(0, I_{d_w})$ is i.i.d. and independent of X_0, and that (A,B) is stabilizable. The system is observed via a finite number of sensors scheduled or queried by the controller at each time step. Let $\{\gamma_t\}$ be a Bernoulli process indicating if the data is lost in the network: each observation is either received ($\gamma_t = 1$) or lost ($\gamma_t = 0$). A scheduled sensor attempts to send information to the controller through the network; depending on the state of the network, the information may be received or lost. This behavior is modeled as

$$Y_t \;=\; C_{Q_{t-1}} X_t + F_{Q_{t-1}} W_t\,, \quad t \ge 1, \tag{3}$$

if $\gamma_t = 1$, otherwise no observation is received. The dimension of Y_t may be variable, and naturally equals the number of rows of C_q for $q = Q_{t-1}$. The query process $\{Q_t\}$ takes values in the finite set of allowable sensor queries denoted by $\mathbb{Q}$. For each query $q \in \mathbb{Q}$, we assume that $\det(F_q F_q^{\mathsf{T}}) \neq 0$ and (primarily to simplify the analysis) that $DF_q^{\mathsf{T}} = 0$. Also without loss of generality, we assume that $\mathrm{rank}(B) = N_u$; if not, we restrict control actions to the row space of B.

The network congestion is modeled as a random process S_t, also controlled by Q_t, taking values on a finite set $\mathbb{S}$ of network states:

$$\mathbb{P}\big(S_{t+1} = s' \mid S_t = s, Q_t = q\big) \;=\; p_q(s,s')\,, \tag{4}$$

for $s, s' \in \mathbb{S}$, $t \ge 0$, and a known initial state $S_0 = s_0 \in \mathbb{S}$. The observed information is lost with a probability that depends on the network state and the query, i.e.,

$$\mathbb{P}(\gamma_{t+1} = 0) = \lambda(S_t, Q_t)\,, \tag{5}$$

where the loss rate $\lambda\colon \mathbb{S}\times\mathbb{Q} \to [0,1]$. The network state S_t and the value of γ_t are assumed to be known to the controller at every time step.

The running cost is the sum of a positive network cost $\mathcal{R}\colon \mathbb{S}\times\mathbb{Q} \to \mathbb{R}$ and a quadratic plant cost $\mathcal{R}_{\mathsf{p}}\colon \mathbb{R}^d \times \mathbb{R}^{N_u} \to \mathbb{R}$ given by

$$\mathcal{R}_{\mathsf{p}}(x,u) = x^{\mathsf{T}} R x + u^{\mathsf{T}} M u\,,$$

where $R, M \in \mathcal{M}^+$. To help with later analysis, we choose some distinguished network state denoted as $\theta \in \mathbb{S}$, which satisfies

$$\theta \in \arg\min_{s\in\mathbb{S}} \Big(\min_{q\in\mathbb{Q}} \mathcal{R}(s,q)\Big)\,,$$

and without loss of generality assume $\min_{q\in\mathbb{Q}} \mathcal{R}(\theta, q) = 1$.

At each time t, the controller takes an action $v_t = (U_t, Q_t)$, the system state evolves as in (2), and the network state transitions according to (4). Then the observation at $t+1$ is either lost or received, determined by (3) and (5). The decision v_t is non-anticipative, i.e., should depend only on the history $\mathcal{F}_t$ of observations up to time t defined by $\mathcal{F}_t \coloneqq \sigma(s_0, x_0, \Sigma_0, S_1, Y_1, \gamma_1, \ldots, S_t, Y_t, \gamma_t)$. Such a sequence of decisions $v = \{v_t\colon t \ge 0\}$ is called a policy, and we denote the set of admissible policies by $\mathcal{V}$. As customary, a policy is called Markov if v_t depends only on the current state.

For an initial condition (s_0, X_0) and a policy $v \in \mathcal{V}$, let $\mathbb{P}^v$ be the unique probability measure on the trajectory space, and $\mathbb{E}^v$ the corresponding expectation operator. When necessary, the explicit dependence on (the law of) the initial conditions or their parameters will be indicated in a subscript, such as $\mathbb{P}^v_{s_0,X_0}$ or $\mathbb{E}^v_{s_0,x_0,\Sigma_0}$.

A. Kalman Filter and Update Properties

We have thus far described a system given by partially observed controlled Markov chain, which we now convert to an equivalent completely observed model. Standard linear estimation theory tells us that the expected value of the state $\widehat{X}_t \coloneqq \mathbb{E}[X_t \,|\, \mathcal{F}_t]$ is a sufficient statistic. Let $\widehat{\Pi}_t$ denote the error covariance matrix given by

$$\widehat{\Pi}_t = \mathrm{cov}(X_t - \widehat{X}_t) = \mathbb{E}\big[(X_t - \widehat{X}_t)(X_t - \widehat{X}_t)^{\mathsf{T}}\big].$$

The state estimate $\widehat{X}_t$ and the error covariance matrix $\widehat{\Pi}_t$ can be dynamically calculated via the Kalman filter

$$\begin{aligned} \widehat{X}_{t+1} &= A\widehat{X}_t + BU_t \\ &\quad + \widehat{K}_{Q_t,\gamma_{t+1}}(\widehat{\Pi}_t)\big(Y_{t+1} - C_{Q_t}(A\widehat{X}_t + BU_t)\big)\,, \end{aligned} \tag{6}$$

with $\widehat{X}_0 = x_0$. The Kalman gain $\widehat{K}_{q,\gamma}$ is given by

$$\widehat{K}_{q,\gamma}(\widehat{\Pi}) \coloneqq \Xi(\widehat{\Pi})\gamma C_q^{\mathsf{T}}\big(\gamma^2 C_q \Xi(\widehat{\Pi}) C_q^{\mathsf{T}} + F_q F_q^{\mathsf{T}}\big)^{-1},$$
$$\Xi(\widehat{\Pi}) \coloneqq DD^{\mathsf{T}} + A\widehat{\Pi}A^{\mathsf{T}},$$

and the error covariance evolves on $\mathcal{M}_0^+$ as

$$\widehat{\Pi}_{t+1} = \Xi(\widehat{\Pi}_t) - \widehat{K}_{Q_t,\gamma_{t+1}}(\widehat{\Pi}_t) C_{Q_t} \Xi(\widehat{\Pi}_t)\,, \tag{7}$$

with $\widehat{\Pi}_0 = \Sigma_0$. When an observation is lost ($\gamma_t = 0$), the gain $\widehat{K}_{q,\gamma_t} = 0$ and the observer (6) simply evolves without any correction factor.

For a sensor query $q \in \mathbb{Q}$, define $\mathcal{T}_q \colon \mathcal{M}_0^+ \to \mathcal{M}_0^+$ by

$$\mathcal{T}_q(\widehat{\Pi}) \coloneqq \Xi(\widehat{\Pi}) - \widehat{K}_{q,1}(\widehat{\Pi}) C_q \Xi(\widehat{\Pi})$$

and an operator $\widehat{\mathcal{T}}_q$ on functions $f \colon \mathbb{S} \times \mathcal{M}_0^+ \to \mathbb{R}$,

$$\widehat{\mathcal{T}}_q f(s,\widehat{\Pi}) = \sum_{s' \in \boldsymbol{S}} p_q(s,s')\big((1-\lambda(s,q)) f(s', \mathcal{T}_q(\widehat{\Pi})) + \lambda(s,q) f(s', \Xi(\widehat{\Pi}))\big)\,.$$

It is clear then that $(S_t, \widehat{\Pi}_t)$ forms a completely observed controlled Markov chain on $\mathbb{S} \times \mathcal{M}_0^+$, with action space $\mathbb{Q}$, and kernel $\widehat{\mathcal{T}}_q$. Admissible and Markov policies are defined just as previously but with $v_t = Q_t$, since the evolution of $\widehat{\Pi}_t$ does not depend on the state control U_t. Thus

$$\widehat{\mathcal{T}}_q f(s,\widehat{\Pi}) = \mathbb{E}^q_{s,\widehat{\Pi}}\big[f(S_1, \widehat{\Pi}_1)\big] \coloneqq \mathbb{E}^q\big[f(S_{t+1}, \widehat{\Pi}_{t+1}) \,\big|\, S_t = s\,, \widehat{\Pi}_t = \widehat{\Pi}\big]\,.$$

We sightly abuse terminology by calling a function f on $\mathbb{S} \times \mathcal{M}_0^+$ concave/continuous/monotone if $f(s,\cdot)$ is concave/continuous/monotone for all $s \in \mathbb{S}$. Note that a function on $\mathcal{M}_0^+$, such as $\mathrm{tr}(\cdot)$, can be naturally extended to $\mathbb{S} \times \mathcal{M}_0^+$, but that $\widehat{\mathcal{T}}_q \mathrm{tr}(\cdot)$ depends implicitly on s. The following lemma follows easily from the definition of $\widehat{\mathcal{T}}_q$ using standard results from, for example, [8, Lemmas 1–2].

Lemma 2.1: $\widehat{\mathcal{T}}_q$ preserves continuity and lower semi-continuity of all functions, and preserves concavity and monotonicity of non-decreasing functions (w.r.t. $\preceq$).

Using the fact that the trace function is concave and non-decreasing, one can show that

$$\widehat{\mathcal{T}}_{q_k} \circ \cdots \circ \widehat{\mathcal{T}}_{q_0} \mathrm{tr}(m\widehat{\Pi}) \;\le\; m\widehat{\mathcal{T}}_{q_k} \circ \cdots \circ \widehat{\mathcal{T}}_{q_0} \mathrm{tr}(\widehat{\Pi}) \tag{8}$$

for any sequence of sensor queries $\{q_0, \dots, q_k\}$.

Note that there is no strict separation principle between estimation and control for the LQG model with sensor scheduling, but the partial separation result in [20] makes optimal control synthesis possible, and renders the completely observed controlled Markov chain $(S_t, \widehat{\Pi}_t)$ equivalent to the partially observed one for control purposes.

B. Stability

A well-known necessary condition for stability is that (A,B) is stabilizable and $(\overline{C}, A)$ is detectable, where $\overline{C} \coloneqq [C_{q_1}^{\mathsf{T}} \mid \cdots \mid C_{q_{|\mathbb{Q}|}}^{\mathsf{T}}]^{\mathsf{T}}$. In the absence of intermittency it has been shown in [20] that these conditions are also sufficient. However, with intermittency these conditions are clearly not sufficient, and simple algebraic sufficient conditions for stability with intermittent observations do not seem possible, even for a system without sensor scheduling [1]. In this work we will simply assume that the estimation is stabilizable under some scheduling policy, and then investigate the optimal control problem under the running cost $\mathcal{R} + \mathcal{R}_{\mathsf{p}}$.

Suppose that a particular query process $\overline{Q}$, together with some state estimation scheme are known to result in a bounded trajectory of the error covariance matrix. It is then clear, by the optimality of the Kalman filter, that $\overline{Q}$ together with the Kalman filter estimator in (7) will also keep the error covariance bounded. Moreover, since (A,B) is stabilizable then a feedback controller can be designed so that the variance of X stays bounded.

Assumption 2.1: The following hold:

(i) The pair (A,D) is controllable.

(ii) The controlled Markov chain governing the network dynamics given in (4) is aperiodic (over any admissible querying policy) and uniformly irreducible in the following sense: there exists $n_\circ \in \mathbb{N}$ such that for any pair of states $s, s' \in \mathbb{S}$, and any sequence of $n_\circ$ queries $\{q_i\}_{i \in \{1,\dots,n_\circ\}}$ in $\mathbb{Q}^{n_\circ}$, there exists a sequence of states $s_1, \dots, s_n$ with $n < n_\circ$ such that

$$p_{q_1}(s,s_1)\, p_{q_2}(s_1,s_2) \cdots p_{q_n}(s_n,s') \;>\; 0\,.$$

(iii) There exists $s_\circ \in \mathbb{S}$, $\Sigma_\circ \in \mathcal{M}_0^+$, and an admissible query process $\overline{Q} = \{\overline{Q}_t : t \ge 0\}$ such that

$$\sup_{t>0} \mathbb{E}^{\overline{Q}}_{s_\circ,\Sigma_\circ}\big[\mathrm{tr}(\widehat{\Pi}_t)\big] \;<\; \infty\,. \tag{9}$$

One can show that Assumption 2.1(iii) generalizes to all initial state combinations. This leads to the following lemma.

Lemma 2.2: Under Assumption 2.1, for any $s \in \mathbb{S}$, $\Sigma \in \mathcal{M}_0^+$, there exists an admissible query process $\overline{Q} = \{\overline{Q}_t : t \ge 0\}$ such that

$$\sup_{t>0} \mathbb{E}^{\overline{Q}}_{s,\Sigma}\big[\mathrm{tr}(\widehat{\Pi}_t)\big] \;\le\; c_0 + c_1 \mathrm{tr}(\Sigma)\,. \tag{10}$$

for some positive constants c_0 and c_1 which does not depend on (s,Σ).

III. Optimal Control

We are interested in finding admissible policies that minimize the long-term average cost,

$$J^v \;\coloneqq\; \limsup_{T\to\infty} \frac{1}{T}\, \mathbb{E}^v\left[\sum_{t=0}^{T-1} \big(\mathcal{R}(S_t,Q_t) + \mathcal{R}_{\mathsf{p}}(X_t,U_t)\big)\right].$$

In approaching the average cost problem, we also consider the α-discounted finite horizon cost for $\alpha \in (0,1)$, given by

$$J^v_{\alpha,n} \;\coloneqq\; \mathbb{E}^v\left[\sum_{t=0}^{n-1} \alpha^t \big(\mathcal{R}(S_t,Q_t) + \mathcal{R}_{\mathsf{p}}(X_t,U_t)\big) + \alpha^n X_n^{\mathsf{T}} \Pi_{\mathsf{fin}} X_n\right], \tag{11}$$

where $\Pi_{\mathsf{fin}} \in \mathcal{M}_0^+$ is a terminal cost, and the α-discounted cost,

$$J_\alpha^v \coloneqq \mathbb{E}^v\left[\sum_{t=0}^{\infty} \alpha^t \big(\mathcal{R}(S_t, Q_t) + \mathcal{R}_{\mathsf{p}}(X_t, U_t)\big)\right].$$

In each of these problems and throughout the analysis, we assume that $S_0 = s_0 \in \mathbb{S}$ and $X_0 \sim \mathcal{N}(x_0, \Sigma_0)$ unless otherwise specified.

A. Optimal Control for the Finite Horizon Problem

The optimal feedback control for the finite horizon problem is well understood; detailed derivations can be found in, for example, [21, Sec. 5.2]. For the finite horizon α-discounted problem, given any particular sequence of n sensor queries, the optimal control policy can be derived directly from (11), and takes the form of the linear feedback control $U_{\alpha,t} = -K_{\alpha,t}\,\mathbb{E}[X_t \mid \mathcal{F}_t]$, where the feedback gain $K_{\alpha,t}$ is determined by the backward recursion

$$\begin{aligned} K_{\alpha,t} &= \alpha(M + \alpha B^{\mathsf{T}}\Pi_{\alpha,t+1}B)^{-1}B^{\mathsf{T}}\Pi_{\alpha,t+1}A\,, \\ \Pi_{\alpha,t} &= R + \alpha A^{\mathsf{T}}\Pi_{\alpha,t+1}A - \alpha A^{\mathsf{T}}\Pi_{\alpha,t+1}BK_{\alpha,t}\,, \end{aligned} \tag{12}$$

with $\Pi_{\alpha,N} = \Pi_{\mathsf{fin}}$. However, to facilitate the study of the infinite horizon case, we note that since (A, B) is stabilizable, there exists a unique matrix $\Pi_\alpha^* \in \mathcal{M}^+$ that solves the algebraic Riccati equation

$$\begin{aligned} \Pi_\alpha^* = R &+ \alpha A^{\mathsf{T}}\Pi_\alpha^* A \\ &- \alpha^2 A^{\mathsf{T}}\Pi_\alpha^* B(M + \alpha B^{\mathsf{T}}\Pi_\alpha^* B)^{-1}B^{\mathsf{T}}\Pi_\alpha^* A\,. \end{aligned}$$

By setting $\Pi_{\mathsf{fin}} = \Pi_\alpha^*$, the backward recursion in (12) is t-invariant and, as noted in Section II, the expected value of the state can be dynamically calculated via the Kalman filter estimate $\widehat{X}$ in (6). So the optimal control for the plant takes the form of a linear feedback given by

$$\begin{aligned} U_{\alpha,t}^* &= -K_\alpha^* \hat{X}_t\,, \\ K_\alpha^* &= (M + \alpha B^{\mathsf{T}}\Pi_\alpha^* B)^{-1}\alpha B^{\mathsf{T}}\Pi_\alpha^* A\,. \end{aligned} \tag{13}$$

Define

$$\tilde{\Pi}_\alpha \coloneqq R - \Pi_\alpha^* + \alpha A^{\mathsf{T}}\Pi_\alpha^* A\,. \tag{14}$$

The following result recasts the finite horizon optimal control problem in terms of the error covariance rather than the system state and control.

Theorem 3.1: Let $v_{\alpha,n}^* = \{U_{\alpha,t}^*, Q_{\alpha,t}^*\}_{0\le t\le n-1}$, where $U_{\alpha,t}^*$ is the linear feedback defined in (13) and $\{Q_{\alpha,t}^*\}$ is a selector from the minimizer in the n-step dynamic programming equation. Define

$$f_t^{(n)}(s, \widehat{\Pi}) = \min_{q\in\mathbb{Q}} \big\{\mathcal{R}(s, q) + \alpha\widehat{\mathcal{T}}_q f_{t+1}^{(n)}(s, \widehat{\Pi})\big\} + \mathrm{tr}(\tilde{\Pi}_\alpha\widehat{\Pi})$$

for $t = 0, \dots, n-1$, with $f_n^{(n)} = 0$. Then $v_{\alpha,n}^*$ is optimal for the finite horizon control problem with $\Pi_{\mathsf{fin}} = \Pi_\alpha^*$, and we have

$$\begin{aligned} J_{\alpha,n}^{v_{\alpha,n}^*} = \inf_{v\in\mathcal{V}} J_{\alpha,n}^v &= f_0^{(n)}(s_0, \Sigma_0) + x_0^{\mathsf{T}}\Pi_\alpha^* x_0 + \mathrm{tr}(\tilde{\Pi}_\alpha\Sigma_0) \\ &\quad + \sum_{k=1}^{n} \alpha^k \mathrm{tr}(\Pi_\alpha^* DD^{\mathsf{T}})\,. \end{aligned}$$

Before proceeding to the infinite horizon results, we show an essential application of the bound in (10).

Lemma 3.1: There exists a positive constant c_s such that with the stabilizing query process $\bar{Q}$ from Assumption 2.1, for any $n > 0$ and $\alpha \in (0, 1)$

$$J_{\alpha,n}^{v_{\alpha,n}^*} \le J_{\alpha,n}^{U_\alpha^*, \bar{Q}} \le c_s\left(\|x_0\|^2 + \frac{1}{1-\alpha} + \frac{\mathrm{tr}(\Sigma_0)}{1-\alpha}\right). \tag{15}$$

Bounds of this form, relating optimal costs to trace, will prove repeatedly useful as the analysis proceeds.

B. Optimal Control for the α-Discounted Problem

Once again, we can recast the optimal control problem in terms of the error covariance rather than the state and control processes. In the infinite horizon case, this leads to a modified discounted optimality equation.

Theorem 3.2: For $\alpha \in (0, 1)$, there exists a unique lower semicontinuous function $f_\alpha^*\colon \mathbb{S}\times\mathcal{M}_0^+ \to \mathbb{R}_+$ that satisfies

$$f_\alpha^*(s, \widehat{\Pi}) = \min_{q\in\mathbb{Q}} \big\{\mathcal{R}(s, q) + \alpha\widehat{\mathcal{T}}_q f_\alpha^*(s, \widehat{\Pi})\big\} + \mathrm{tr}(\tilde{\Pi}_\alpha\widehat{\Pi})\,, \tag{16}$$

with $\tilde{\Pi}_\alpha$ as in (14). If $q_\alpha^*\colon \mathbb{S}\times\mathcal{M}_0^+ \to \mathbb{Q}$ is a selector of the minimizer in (16), then the Markov policy given by $v_\alpha^* = \{q_\alpha^*(S_t, \widehat{\Pi}_t), U_{\alpha,t}^*\}_{t\ge 0}$ is optimal for the α-discounted infinite horizon problem, and

$$\begin{aligned} J_\alpha^{v_\alpha^*}(s_0, x_0, \Sigma_0) &= \inf_{v\in\mathcal{V}} J_\alpha^v(s_0, x_0, \Sigma_0) \\ &= f_\alpha^*(s_0, \Sigma_0) + x_0^{\mathsf{T}}\Pi_\alpha^* x_0 + \mathrm{tr}(\tilde{\Pi}_\alpha\Sigma_0) \\ &\quad + \frac{\alpha}{1-\alpha}\mathrm{tr}(\Pi_\alpha^* DD^{\mathsf{T}})\,. \end{aligned}$$

Further, the querying component of any optimal stationary Markov policy is an a.e. selector of the minimizer in (16).

C. Optimal Control for the Average Cost Problem

We use the vanishing discount approach to establish a solution to the average cost problem. A critical result enabling the vanishing discount approach is the following:

Lemma 3.2: The differential discounted value function $\bar{f}_\alpha \coloneqq f_\alpha^* - f_\alpha^*(\theta, 0)$ is locally bounded, uniformly in $\alpha \in (0, 1)$, and the set $\{\bar{f}_\alpha : \alpha \in (0, 1)\}$ is locally Lipschitz equicontinuous on compact subsets of $\mathcal{M}_0^+$.

In the course of proving Lemma 3.2, another upper bound with trace is shown: for some positive constant κ_0 we have

$$\bar{f}_\alpha(s, \Sigma) \le \kappa_0\big(1 + \mathrm{tr}(\Sigma)\big)\,. \tag{17}$$

Using this bound and the properties of trace, we characterize solutions of the average cost problem and show that an optimal stationary policy exists.

Theorem 3.3: There exist a constant ϱ^* and a continuous function $f^*\colon \mathbb{S}\times\mathcal{M}_0^+ \to \mathbb{R}_+$ that satisfy

$$f^*(s, \widehat{\Pi}) + \varrho^* = \min_{q\in\mathbb{Q}} \big\{\mathcal{R}(s, q) + \mathrm{tr}(\tilde{\Pi}^*\widehat{\Pi}) + \widehat{\mathcal{T}}_q f^*(s, \widehat{\Pi})\big\}\,, \tag{18}$$

with $\tilde{\Pi}^* \coloneqq R - \Pi^* + A^{\mathsf{T}}\Pi^* A$, and $\Pi^* \in \mathcal{M}^+$ the unique solution of the algebraic Riccati equation

$$\Pi^* = R + A^{\mathsf{T}}\Pi^* A - A^{\mathsf{T}}\Pi^* B(M + B^{\mathsf{T}}\Pi^* B)^{-1}B^{\mathsf{T}}\Pi^* A.$$

If $q^*\colon \mathbb{S}\times\mathcal{M}_0^+\to\mathbb{Q}$ is a selector of the minimizer in (18), then the policy given by $v^*=\{U_t^*, q^*(S_t,\widehat{\Pi}_t\}_{t\ge0}$, with

$$\begin{aligned} U_t^* &\coloneqq -K^*\widehat{X}_t\,,\\ K^* &\coloneqq (M+B^{\mathsf{T}}\Pi^*B)^{-1}B^{\mathsf{T}}\Pi^*A\,, \end{aligned} \tag{19}$$

and $\{\widehat{X}_t\}$ as in (6), is optimal, and satisfies

$$J^{v^*}=\inf_{v\in\mathcal{V}}J^v=\varrho^*+\operatorname{tr}(\Pi^*DD^{\mathsf{T}})\,.$$

In addition, the querying part of any optimal stationary Markov policy is an a.e. selector of the minimizer in (18).

It is worth noting that f^* is concave and non-decreasing in $\mathcal{M}_0^+$, and that using (17) and the vanishing discount construction of f^*, there exist constants $m_1^*>0$ and $m_0^*\in\mathbb{R}$ such that

$$f^*(s,\Sigma)\;\le\; m_1^*\operatorname{tr}(\Sigma)+m_0^*\,. \tag{20}$$

Furthermore, directly from (18),

$$f^*(s,\Sigma)\;\ge\;\underline{\sigma}(\tilde{\Pi}^*)\operatorname{tr}(\Sigma)-\varrho^*\,,$$

so f^* must be strictly increasing in Σ.

Noting the definition of f^* in (18), for the remainder of the paper we consider the equivalent average cost optimization problem with the cost function $r^q(s,\Sigma)\coloneqq\mathcal{R}(s,q)+\operatorname{tr}(\tilde{\Pi}^*\Sigma)$.

Remark 3.1: For computational purposes, the unbounded cone $\mathcal{M}_0^+$ is clearly impractical. However, we can approximate the process on the bounded subset

$$\mathcal{B}_r\;\coloneqq\mathbb{S}\times\{\Sigma\in\mathcal{M}_0^+:\operatorname{tr}(\Sigma)\le r\}\,,\quad r>0\,.$$

First, we choose any stable control $\bar{q}$. Then we construct a function $f^r\colon\mathbb{S}\times\mathcal{M}_0^+\to\mathbb{R}_+$ by solving the dynamic programming equation

$$f^r(s,\Sigma)+\varrho^r\;=\;\min_{q\in\mathbb{Q}}\{r^q(s,\Sigma)+\widehat{\mathcal{T}}_q f^r(s,\Sigma)\}\,, \tag{21}$$

for $(s,\Sigma)\in\mathcal{B}_r$, while for $(s,\Sigma)\in\mathcal{B}_r^c$ we solve the Poisson equation corresponding to (21) with $q=\bar{q}$. We let q^r denote the concatenation of the control $\bar{q}$ with a measurable selector from (21). Note that f^r satisfies the geometric drift condition in the proof of Theorem 3.3. As a result the process under q^r is stable. We leave it to the reader to verify that as $r\to\infty$, $\varrho^r\to\varrho^*$, and so the truncated system is a good approximation of the complete system.

IV. Relative Value Iteration

The relative value iteration (RVI) and value iteration (VI) algorithms generate a sequence of real-valued functions on $\mathbb{S}\times\mathcal{M}_0^+$ and associated constants that, as we will show, approach solutions (f^*,ϱ^*) of (18). For a stationary Markov policy $\bar{q}\colon\mathbb{S}\times\mathcal{M}_0^+\to\mathbb{Q}$, we adopt the notation

$$r^{\bar{q}}(s,\Sigma)\coloneqq\mathcal{R}(s,\bar{q}(s,\Sigma))+\operatorname{tr}(\tilde{\Pi}^*\Sigma).$$

Respectively, the RVI and VI are given by

$$\varphi_{n+1}=\min_{q\in\mathbb{Q}}\left\{r^q+\widehat{\mathcal{T}}_q\varphi_n\right\}-\varphi_n(\theta,0)\,, \tag{22}$$

$$\overline{\varphi}_{n+1}=\min_{q\in\mathbb{Q}}\left\{r^q+\widehat{\mathcal{T}}_q\,\overline{\varphi}_n\right\}-\varrho^*\,,\quad\overline{\varphi}_0=\varphi_0\,, \tag{23}$$

where both algorithms are initialized with the same function $\varphi_0\colon(\mathbb{S}\times\mathcal{M}_0^+)\to\mathbb{R}_+$.

Using the bound in (20), we can find positive constants θ_1 and θ_2 such that

$$\min_{q\in\mathbb{Q}}r^q(s,\Sigma)\ge\theta_1 f^*(s,\Sigma)-\theta_2\,.$$

Without loss of generality we can assume $\theta_1<1$ to facilitate some later estimates.

The next theorem proves that both the RVI and VI algorithms converge. Note that the initialization requirements are easily satisfied by, for example, $\varphi_0=0$.

Theorem 4.1: If $\varphi_0\in\mathcal{O}_{f^*}$, then $\overline{\varphi}_n$ converges to c_0+f^* for some $c_0\in\mathbb{R}$ satisfying

$$-\frac{\varrho^*+\theta_2}{\theta_1}\;\le\;c_0\;\le\;\frac{\varrho^*+\theta_2}{\theta_1}\,\|\varphi_0\|_{f^*}\,, \tag{24}$$

and φ_n converges to $f^*-f^*(\theta,0)+\varrho^*$.

Stability of the policies generated by the VI/RVI algorithms is usually not guaranteed. One would hope that the Markov policy computed at the n^{th} stage of the value iteration is a stable Markov policy and its performance converges to the optimal performance as $n\to\infty$. This topic is commonly referred to as rolling horizon, and is well understood for finite state MDPs [22] but it is decidedly unexplored for nonfinite state models. Among the very few results in the literature is the study in [23] for bounded running cost and under a simultaneous Doeblin hypothesis, and the results in [24] under strong blanket stability assumptions. For the model considered here there is no blanket stability; instead, the inf-compactness of the running cost penalizes unstable behavior. Exploiting the constructive steps of the value iteration convergence proofs allows us to show that the rolling horizon policies are indeed stable, as follows.

Theorem 4.2: For large n, the policy $\widehat{q}^n$ generated by the n^{th} stage of the VI or RVI algorithm is geometrically stable, and the average cost obtained under $\widehat{q}^n$ converges geometrically to ϱ^* as $n\to\infty$.

V. Sensor-Dependent Loss Rates

We now turn our attention to a special case of the previous results, with a single network state. In this case, the network cost is simply a function of the query process $\{Q_t\}$, taking values in the finite set of allowable sensor queries $\mathbb{Q}$. The loss rate depends only on the query, so can be treated as a vector λ in $[0,1]^{|\mathbb{Q}|}$, indexed by the corresponding query:

$$\mathbb{P}(\gamma=1)=(1-\lambda_q),\qquad\mathbb{P}(\gamma=0)=\lambda_q, \tag{25}$$

for $q\in\mathbb{Q}$. We are interested in characterizing the set of loss rates $\Lambda_s\subset[0,1]^{|\mathbb{Q}|}$ for which the system is stabilizable. Our formulation generalizes the problem in [1], which analyzes the system (2)–(3) without sensor scheduling ($C_q=C$) and therefore with a single loss rate.

Recalling the discussion around Assumption 2.1, $\Lambda_s=\varnothing$ unless (A,B) is stabilizable and $(\overline{C},A)$ is detectable. Hence, without loss of generality, we assume (A,B) is stabilizable and $(\overline{C},A)$ is detectable and therefore, by the results in [20], $0\in\Lambda_s$.

Theorem 5.1: If the system (2)–(3) with (25) is stabilizable for a loss rate $\lambda' \in [0,1]^{|\mathbb{Q}|}$, then it is also stabilizable for any other loss rate $\lambda \leq \lambda'$. In other words, the set Λ_s is order-convex with respect to the natural ordering of positive vectors in $\mathbb{R}^{|\mathbb{Q}|}$.

Moreover, a lower loss rate leads to a smaller error covariance at every time step. We continue with another important result.

Theorem 5.2: If the system (2)–(3) with (25) is stabilizable for a loss rate $\lambda \in [0,1]^{|\mathbb{Q}|}$, there exists an open neighborhood $\mathcal{B} \subset [0,1]^{|\mathbb{Q}|}$ around λ such that the system is stabilizable for $\lambda' \in \mathcal{B}$.

Combining these results we obtain the following corollary concerning the structure of Λ_s.

Corollary 5.1: Suppose that (A, B) is stabilizable and $(\overline{C}, A)$ is detectable. Then, there exists a critical surface $\mathcal{W}$ in $(0,1]^{|\mathbb{Q}|}$ such that the system is stabilizable with loss rate λ if and only if $\lambda < \lambda' \in \mathcal{W}$. More precisely, there exists a function $\mathcal{F} : \mathbb{R}^{|\mathbb{Q}|-1} \to [0,1]$ which is nonincreasing in each argument such that the system is stabilizable with loss rate λ if and only if $\lambda_{|\mathbb{Q}|} < \mathcal{F}(\lambda_1, \ldots, \lambda_{|\mathbb{Q}|-1})$. In other words, Λ_s is the strict hypograph of $\mathcal{F}$.

We call the set of sensor queries $\mathbb{Q}$ *non-redundant* if the system is not detectable with any proper subset of the sensor queries. That is, the system using only $\mathbb{Q} \setminus \{q\}$ for any $q \in \mathbb{Q}$ is not stabilizable for any admissible query sequence. When $\mathbb{Q}$ is non-redundant and $\mathbf{q}$ is a stabilizing stationary Markov policy, the set of states where any particular query q is chosen, $\mathbf{S}_q = \{\Sigma \in \mathcal{M}_0^+ : \mathbf{q}(\Sigma) = q\}$, satisfies $\mu_{\mathbf{q}}(\mathbf{S}_q) > 0$ for each $q \in \mathbb{Q}$ where $\mu_{\mathbf{q}}$ is the invariant probability measure. Furthermore, there must be a subset $\widehat{\mathbf{S}}_q \subset \mathbf{S}_q$ with $\mu_{\mathbf{q}}(\widehat{\mathbf{S}}_q) > 0$ such that $T_q(\widehat{\Sigma}) < \Xi(\widehat{\Sigma})$ for all $\widehat{\Sigma} \in \widehat{\mathbf{S}}_q$; if not, then a different sensor could be queried instead of q and the system would still be stable.

Theorem 5.3: Suppose that the set of sensors is non-redundant and that $\lambda, \lambda' \in \Lambda_s$ such that $\lambda \leq \lambda'$ and $\lambda \neq \lambda'$. Then $\varrho_\lambda^* < \varrho_{\lambda'}^*$.

Noting that the average cost $\varrho_\lambda^* \to \infty$ as the system parameters approach the boundary of the stability region the set $\Lambda(\kappa) \coloneqq \{\lambda \colon \varrho_\lambda^* < \kappa\}$ is a ray-connected neighborhood of 0 for all $\kappa > 0$. Clearly, $\bigcup_{\kappa>0} \Lambda(\kappa) = \Lambda_s$.

Remark 5.1: Suppose that the loss rates depend only on the query, as in (25), but are unknown. Then the implications of Theorem 5.2 are remarkable. Since stability is shown to be an open property, if one can find an estimator sequence $\widehat{\lambda}_t \to \lambda$ a.s., then the system will retain stability and the long-term average performance would be the same as the if the rates were known beforehand. Since the channel is Bernoulli, recursive estimation of the loss rates leading to a.s. convergence to the true value is rather straightforward. For example, a maximum likelihood estimator can be employed, as in [25].

References

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

[2] Y. Mo and B. Sinopoli, "Kalman filtering with intermittent observations: Critical value for second order system," IFAC Proceedings Volumes, vol. 44, no. 1, pp. 6592 – 6597, 2011.

[3] K. Plarre and F. Bullo, "On Kalman filtering for detectable systems with intermittent observations," IEEE Trans. Automat. Control, vol. 54, no. 2, pp. 386–390, 2009.

[4] S. Kar, B. Sinopoli, and J. M. F. Moura, "Kalman filtering with intermittent observations: weak convergence to a stationary distribution," IEEE Trans. Automat. Control, vol. 57, no. 2, pp. 405–420, 2012.

[5] Y. Mo and B. Sinopoli, "Kalman filtering with intermittent observations: tail distribution and critical value," IEEE Trans. Automat. Control, vol. 57, no. 3, pp. 677–689, 2012.

[6] D. E. Quevedo, A. Ahlén, and K. H. Johansson, "State estimation over sensor networks with correlated wireless fading channels," IEEE Trans. Automat. Control, vol. 58, no. 3, pp. 581–593, 2013.

[7] T. Sui, K. You, and M. Fu, "Stability conditions for multi-sensor state estimation over a lossy network," Automatica J. IFAC, vol. 53, pp. 1–9, 2015.

[8] V. Gupta, T. H. Chung, B. Hassibi, and R. M. Murray, "On a stochastic sensor selection algorithm with applications in sensor scheduling and sensor coverage," Automatica J. IFAC, vol. 42, no. 2, pp. 251–260, 2006.

[9] G. Battistelli, A. Benavoli, and L. Chisci, "State estimation with remote sensors and intermittent transmissions," Systems Control Lett., vol. 61, no. 1, pp. 155–164, 2012.

[10] Y. Mo, E. Garone, A. Casavola, and B. Sinopoli, "Stochastic sensor scheduling for energy constrained estimation in multi-hop wireless sensor networks," IEEE Trans. Automat. Control, vol. 56, no. 10, pp. 2489–2495, 2011.

[11] S. Trimpe and R. D'Andrea, "Event-based state estimation with variance-based triggering," IEEE Trans. Automat. Control, vol. 59, no. 12, pp. 3266–3281, 2014.

[12] D. Han, Y. Mo, J. Wu, S. Weerakkody, B. Sinopoli, and L. Shi, "Stochastic event-triggered sensor schedule for remote state estimation," IEEE Trans. Automat. Control, vol. 60, no. 10, pp. 2661–2675, 2015.

[13] S. Weerakkody, Y. Mo, B. Sinopoli, D. Han, and L. Shi, "Multi-sensor scheduling for state estimation with event-based, stochastic triggers," IEEE Trans. Automat. Control, vol. 61, no. 9, pp. 2695–2701, 2016.

[14] D. Han, J. Wu, H. Zhang, and L. Shi, "Optimal sensor scheduling for multiple linear dynamical systems," Automatica J. IFAC, vol. 75, pp. 260–270, 2017.

[15] R. Ambrosino, B. Sinopoli, and K. Poolla, Optimal sensor scheduling for remote estimation over wireless sensor networks, ser. Underst. Complex Syst. Springer, Berlin, 2009, pp. 127–142.

[16] Q.-S. Jia, L. Shi, Y. Mo, and B. Sinopoli, "On optimal partial broadcasting of wireless sensor networks for Kalman filtering," IEEE Trans. Automat. Control, vol. 57, no. 3, pp. 715–721, 2012.

[17] Y. Mo, E. Garone, and B. Sinopoli, "On infinite-horizon sensor scheduling," Systems Control Lett., vol. 67, pp. 65–70, 2014.

[18] L. Zhao, W. Zhang, J. Hu, A. Abate, and C. J. Tomlin, "On the optimal solutions of the infinite-horizon linear sensor scheduling problem," IEEE Trans. Automat. Control, vol. 59, no. 10, pp. 2825–2830, 2014.

[19] A. S. Leong, S. Dey, and D. E. Quevedo, "Sensor scheduling in variance based event triggered estimation with packet drops," IEEE Trans. Automat. Control, vol. 62, no. 4, pp. 1880–1895, 2017.

[20] W. Wu and A. Arapostathis, "Optimal sensor querying: general Markovian and LQG models with controlled observations," IEEE Trans. Automat. Control, vol. 53, no. 6, pp. 1392–1405, 2008.

[21] D. P. Bertsekas, Dynamic programming and optimal control. Vol. I, 3rd ed. Athena Scientific, Belmont, MA, 2005.

[22] E. Della Vecchia, S. Di Marco, and A. Jean-Marie, "Illustrated review of convergence conditions of the value iteration algorithm and the rolling horizon procedure for average-cost MDPs," Ann. Oper. Res., vol. 199, pp. 193–214, 2012.

[23] R. Cavazos-Cadena, "A note on the convergence rate of the value iteration scheme in controlled Markov chains," Systems Control Lett., vol. 33, no. 4, pp. 221–230, 1998.

[24] O. Hernández-Lerma and J. B. Lasserre, "Error bounds for rolling horizon policies in discrete-time Markov control processes," IEEE Trans. Automat. Control, vol. 35, no. 10, pp. 1118–1124, 1990.

[25] E. Fernández-Gaucherand, A. Arapostathis, and S. I. Marcus, "Analysis of an adaptive control scheme for a partially observed controlled Markov chain," IEEE Trans. Automat. Control, vol. 38, no. 6, pp. 987–993, 1993.