A Stochastic Online Sensor Scheduler for Remote State Estimation With Time-Out Condition

Junfeng Wu, Karl Henrik Johansson, and Ling Shi

***Abstract*—This technical note considers remote state estimation subject to limited sensor-estimator communication rate. We propose a stochastic online sensor scheduler for remote state estimation with time-out condition. The decision rule under which the sensor sends data is based on its measurements and a finite-state holding time between the present and the most recent sensor-to-estimator communication instance. This decision process is formulated as an optimization problem, relaxed and solved using generalized geometric programming optimization techniques with a low computational complexity. Moreover, the proposed scheduler is easy to execute, and provides a guaranteed performance which is shown to outperform the optimal offline scheduler. Numerical examples are provided to illustrate the proposed scheduler.**

***Index Terms*— Generalized geometric programming (GGP), networked control systems (NCSs).**

I. Introduction

In the last decade, the research on networked control systems (NCSs) has received a boom of interest. State estimation is inherent in many applications of NCSs such as civil structure maintenance, emergency rescue and environmental monitoring, where sensor measurements are sent to a remote state estimator (e.g., a base station or a central computation unit) for real-time processing. As sensors are mostly battery-powered and consume most energy in data communication, it is critical to reduce the sensor-estimator communication rate to conserve energy. On the other hand, reduced communication rate may lead to a poor estimation quality. Therefore it is important to understand the limitation on the remote estimation quality when the communication rate is constrained, so as to help reduce the (expensive) communication resources while still guarantee a desired estimation quality. For example, Trimpe and D'Andrea [1] demonstrated via the Balancing Cube experimental platform that by properly designing a sensor transmission scheduler, significant communication reduction is achieved without affecting system stability.

Related research on state estimation under communication constraints and sensor scheduling are introduced herein. Sandberg *et al.* [2] considered estimation using two types of sensors: the first type has low-quality measurement but small processing delay, while the second type has high-quality measurement but large delay. Using a time-periodic Kalman filter, they found an optimal schedule of the sensor communication. Savage and La Scala [3] provided the optimal schedule to minimize the terminal estimation error covariance under the constraint that within a time horizon N, only $n < N$ measurements could be taken. The above works focus on offline schedulers, i.e., the communication schedule is made independent of the sensor's measurements, before the system runs. In [4], an event-triggering rule was designed to determine the data transmission from a sensor to a remote observer, subject to a constraint on transmission frequency. Computing the optimal event-triggering rule was shown to be computationally intractable when the state dimension is greater than two or when the considered time-horizon is large. Wu *et al.* [5] presented an event-based remote state estimator and illustrated how to achieve a tradeoff between the sensor-estimator communication rate and the estimation quality. Sijs *et al.* [6] proposed an event-based state estimator according to a hybrid update rule, which attains a bounded error covariance. In both [5] and [6] Gaussian approximations for some truncated Gaussian random variables are adopted to facilitate the analysis. More related works can be found in recent works [7], [8] and references therein.

In this paper, we consider a remote state estimation problem where a sensor obtains the output of a linear system, pre-processes the measurement and calculates a local state estimate, then decides whether or not to send its local estimate to a remote estimator. Unlike [4], where this decision-making process is formulated as a Markov decision problem, we present a simple and easily computable strategy. The main contributions are summarized as follows.

1) We propose a new online sensor scheduler and provide a closed-form expression for the estimation error covariance matrix, computed by the remote minimum mean-squared error (MMSE) estimator.
2) We relax a scheduler optimization problem and show that the scheduler outperforms the optimal offline one. This sub-optimal scheduler can be solved using generalized geometric programming (GGP) optimization techniques with a low computational complexity.

The decision for the sensor transmission is based on an innovation process and a finite-state holding time between two consecutive events. Compared to [4] and [9], our method has relatively light computational complexity for high-dimensional systems. Note that [8] also considered remote estimation, but with a different scheduling algorithm. We show by simulations that the scheduler in [8] is slightly better than our scheduler when the communication rate is low, and the difference is negligible when the communication rate increases. Using our scheduler, however, we provide with a closed-form expression of the relationship between the threshold and the estimation error covariance; while it relies on numerical computations to analyze the scheduler in [8]. The remainder of this paper is organized as follows. In Section II, we provide the mathematical model and problem setup. In Sections III and IV, we derive the MMSE estimator under the proposed scheduler and find a sub-optimal scheduler by solving a GGP optimization problem. Simulation is provided in Section V and some concluding remarks are given in Section VI.

Manuscript received July 14, 2013; revised February 5, 2014; accepted April 24, 2014. Date of publication May 6, 2014; date of current version October 21, 2014. This work was supported by a Hong Kong Research Grant Council GRF grant 618612, the Knut and Alice Wallenberg Foundation, and the Swedish Research Council. Recommended by Associate Editor S. Zampieri.

J. Wu and L. Shi are with the Department of Electronic and Computer Engineering, Hong Kong University of Science and Technology, Hong Kong (e-mail: jfwu@ust.hk; eesling@ust.hk).

K. H. Johansson is with the ACCESS Linnaeus Center, School of Electrical Engineering, Royal Institute of Technology, Stockholm, Sweden (e-mail: kallej@ee.kth.se).

Color versions of one or more of the figures in this paper are available online at http://ieeexplore.ieee.org.

Digital Object Identifier 10.1109/TAC.2014.2322153

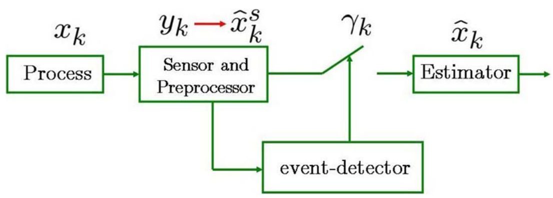


Fig. 1. Finite-state online scheduler for remote state estimation.

Notations: $\mathbb{Z}$ (or $\mathbb{Z}_+$) is the set of non-negative (or positive) integers. $\mathbb{S}_+^n$ is the set of n by n positive semi-definite matrices. $\sigma(\cdot)$ is the smallest σ-algebra generated by random variables. $\|x\|_\infty$ is the infinity norm of a vector x. $\mathcal{S}$ denotes the state space of a Markov chain. Define the function $h : \mathbb{S}_+^n \to \mathbb{S}_+^n$ as $h(X) \triangleq AXA' + Q$.

II. Problem Setup

Consider a linear time-invariant process (Fig. 1)

$$x_{k+1} = Ax_k + w_k \tag{1}$$

$$y_k = Cx_k + v_k \tag{2}$$

where $x_k \in \mathbb{R}^n$ is the process state vector, $y_k \in \mathbb{R}^m$ is the observation vector, $w_k \in \mathbb{R}^n$ and $v_k \in \mathbb{R}^m$ are zero-mean Gaussian random vectors with $\mathbb{E}[w_k w_j'] = \delta_{kj} Q (Q \geq 0)$, $\mathbb{E}[v_k v_j'] = \delta_{kj} R (R > 0)$, $\mathbb{E}[w_k v_j'] = 0 \; \forall j, k$. The δ_{kj} is the Kronecker delta function with $\delta_{kj} = 1$ if $k = j$ and $\delta_{kj} = 0$ otherwise. The initial state x_0 is a zero-mean Gaussian random vector that is uncorrelated with w_k and v_k and has covariance $\Sigma_0 \geq 0$. The pair (C, A) is assumed to be observable and $(A, \sqrt{Q})$ is controllable.

Define $\mathcal{F}_k^s$ as the filtration generated by all the measurements collected by the sensor up to time k, i.e., $\mathcal{F}_k^s \triangleq \sigma(y_t, 0 \leq t \leq k)$. We will use a triple $(\Omega, \mathcal{F}, \mathbb{P})$ to denote the common probability space for all random variables in this paper, where $\mathcal{F} = \sigma(\cup_{k=1}^\infty \mathcal{F}_k)$ and $\mathbb{P}$ is the probability measure on $\mathcal{F}$. The sensor computes $\hat{x}_k^s$, the MMSE estimate of x_k in (1) based on $\mathcal{F}_k^s$, i.e., $\hat{x}_k^s = \mathbb{E}[x_k|\mathcal{F}_k^s] \in \mathcal{F}_k^s$. Let e_k^s and P_k^s be the corresponding estimation error and error covariance matrix, i.e.

$$e_k^s = x_k - \hat{x}_k^s \tag{3}$$

$$P_k^s = \mathbb{E}\left[(e_k^s)(e_k^s)' | \mathcal{F}_k^s\right] \tag{4}$$

which are computed recursively via a Kalman filter [10]. The recursion starts from $\hat{x}_0^s = 0$. As P_k^s converges to its steady-state value, $\overline{P}$, exponentially fast [10] and we consider an infinite-time horizon, we omit the transient estimation process at the sensor side and assume $\Sigma_0 = \overline{P}$ in the sequel. As a result, we make a standing assumption throughout the paper, which is $P_k^s = \overline{P}$, $\forall k \in \mathbb{Z}$. The following result is a useful properties of $\overline{P}$. See [11] for a proof.

Lemma 2.1: For $0 \leq t_1 < t_2$, the following inequality holds:

$$\mathrm{Tr}\left(h^{t_1}(\overline{P})\right) < \mathrm{Tr}\left(h^{t_2}(\overline{P})\right). \tag{5}$$

After $\hat{x}_k^s$ is computed, the sensor decides whether it will send $\hat{x}_k^s$ to the remote estimator. Let $\gamma_k \in \{0, 1\}$ be its decision variable at time k, i.e., if $\gamma_k = 1$, $\hat{x}_k^s$ is sent; otherwise $\hat{x}_k^s$ is not sent. Define the holding time $\tau_k \in \mathbb{Z}$ as follows:

$$\tau_k \triangleq k - \max_{1 \leq t \leq k} \{t : \gamma_t = 1\} \tag{6}$$

which denotes the time between k and the most recent instance when the sensor communicated with the estimator. When the time index is clear from the context, we will write τ_k as τ for simplicity. We define I_k as the information pattern available to the remote estimator up to time k, i.e., $\mathrm{I}_k = \{\hat{x}_1^s, \hat{x}_2^s, \ldots, \hat{x}_{k-\tau}^s\} \cup \{\gamma_1, \gamma_2, \ldots, \gamma_k\}$.

Define an infinite time-horizon schedule θ as $\theta = \{\gamma_1, \ldots, \gamma_k, \ldots\} \in \{0, 1\}^\infty$. Under a given θ, the remote estimator calculates $\hat{x}_k$ and P_k, its own MMSE estimate of x_k and the corresponding estimation error covariance, based on I_k

$$\hat{x}_k = \mathbb{E}[x_k|\mathrm{I}_k] \text{ and } P_k = \mathbb{E}\left[(x_k - \hat{x}_k)(\cdot)'|\mathrm{I}_k\right].$$

Define $J(\theta)$ as the trace of the average expected estimation error covariance, i.e., $J(\theta) \triangleq \limsup_{T \to +\infty} (1/T) \sum_{k=0}^{T-1} \mathrm{Tr}(P_k(\theta))$. We are interested in finding a schedule θ which solves the following problem:

$$\min_\theta J(\theta), \quad \text{s.t.} \limsup_{T \to +\infty} \frac{1}{T} \sum_{k=0}^{T-1} \gamma_k(\theta) = \Psi \tag{7}$$

where $\Psi \in [0, 1]$ denotes the maximum rate that the sensor communicates with the estimator. For simplicity, Ψ is assumed to be rational and can be written as $\Psi = v/u$ for two co-prime integers u and v. Note that there must exist $q \in \mathbb{Z}_+$ which satisfies

$$qv < u \leq (q+1)v. \tag{8}$$

The following proposition presents an optimal offline schedule to (7), which is periodic and easy to implement in practice. The proof, which is similar to the proof of Theorem 5.2 in [12], is omitted. As shown in the next few sections, by properly utilizing the online information and constructing the corresponding online schedule, the estimation quality can be improved.

Proposition 2.2: An optimal offline schedule $\theta_{\text{off}}^\star$ over one period u in (8) depicted as

$$\underbrace{\left(\underbrace{1, 0, \ldots, 0}_{q \text{ times}}\right), \ldots, \left(\underbrace{1, 0, \ldots, 0}_{q \text{ times}}\right)}_{(vq+v-u) \text{ times}} \quad \underbrace{\left(\underbrace{1, 0, \ldots, 0}_{(q+1) \text{ times}}\right), \ldots, \left(\underbrace{1, 0, \ldots, 0}_{(q+1) \text{ times}}\right)}_{(u-vq) \text{ times}}$$

provides an optimal offline solution to (7). The corresponding cost function J is given by

$$J\left(\theta_{\text{off}}^\star\right) = \Psi \sum_{i=0}^{q-1} \mathrm{Tr}\left(h^i(\overline{P})\right) + (1 - \Psi q)\mathrm{Tr}\left(h^q(\overline{P})\right). \tag{9}$$

Remark 2.3: Let us treat each $(\underbrace{1, 0, \ldots, 0}_{q \text{ times}})$ or $(\underbrace{1, 0, \ldots, 0}_{q+1 \text{ times}})$ as one unit. Any permutation of every unit inside one period generates an optimal offline schedule $\theta_{\text{off}}^\star$.

III. A Stochastic Online Sensor Scheduler

In control of engineering systems, actions are often taken only after certain events occur. These events may contain useful information about the system. Wu *et al.* [13] proposed two simple online schedulers which, however, only apply when communication rate is in the range of (0.5, 1). In this section we construct an online sensor scheduler with time-out condition to overcome the limitations of the hybrid scheduler in [13], where the decision of γ_k is made based on how the measurement history is generated.

A. Preliminaries

Define ε_k as the incremental innovative information in $\hat{x}_k^s$ compared to $\hat{x}_{k-1}^s$

$$\varepsilon_k \triangleq \hat{x}_k^s - A\hat{x}_{k-1}^s. \tag{10}$$

It has the properties given in the following lemma.

Lemma 3.1: The following statements on ε_k hold:

1) ε_k is zero-mean Gaussian with $\mathbb{E}[\varepsilon_k \varepsilon_k'] = h(\overline{P}) - \overline{P}$. For any $d \in \mathbb{Z}$, ε_{k-d} and e_k^s are independent and $\mathbb{E}[e_k^s \varepsilon_{k-d}'] = 0$.
2) ε_j and ε_k are independent for any $j \neq k$.

Proof: From (3) and the previous two parts, one obtains that $\mathbb{E}[x_k(e_k^s)'] = \mathbb{E}[(e_k^s + \hat{x}_k^s)(e_k^s)'] = \overline{P}$. Therefore $\mathbb{E}[(x_k - A\hat{x}_{k-1}^s)(e_k^s)'] = \mathbb{E}[x_k(e_k^s)'] - A\mathbb{E}[\hat{x}_{k-1}^s(e_k^s)'] = \overline{P}$, which results in

$$\begin{aligned}
\mathbb{E}\left[\varepsilon_k \varepsilon_k'\right] &= \mathbb{E}\left[\left(x_k - A\hat{x}_{k-1}^s - \hat{e}_k^s\right)\left(x_k - A\hat{x}_{k-1}^s - \hat{e}_k^s\right)'\right] \\
&= \mathbb{E}\left[\left(x_k - A\hat{x}_{k-1}^s\right)(\cdot)'\right] - \mathbb{E}\left[\left(x_k - A\hat{x}_{k-1}^s\right)\left(e_k^s\right)'\right] \\
&\quad - \mathbb{E}\left[e_k^s\left(x_k - A\hat{x}_{k-1}^s\right)'\right] + \mathbb{E}\left[\left(e_k^s\right)\left(e_k^s\right)'\right] \\
&= h(\overline{P}) - \overline{P}.
\end{aligned}$$

The other parts are from properties of innovations sequence [10]. ∎

Let the rank of $h(\overline{P}) - \overline{P}$ be r. Since $h(\overline{P}) - \overline{P} \geq 0$, there exists an orthonogal matrix $U \in \mathbb{R}^{n \times n}$ such that $U'(h(\overline{P}) - \overline{P})U = \mathrm{diag}(\Lambda, \mathbf{0}_{n-r})$, where $\Lambda = \mathrm{diag}(\lambda_1, \ldots, \lambda_r)$ and $\lambda_1, \ldots, \lambda_r \in \mathbb{R}$ are the r positive eigenvalues of $h(\overline{P}) - \overline{P}$. Define $F \in \mathbb{R}^{n \times n}$ as $F \triangleq U\mathrm{diag}(\Lambda^{-(1/2)}, I_{n-r})$. Then $F'(h(\overline{P}) - \overline{P})F = \mathrm{diag}(I_r, \mathbf{0}_{n-r})$. Define ϵ_k as $\epsilon_k \triangleq F'\varepsilon_k$. Through this linear transformation, the coordinates of ε_k are decorrelated, which enables us to analyze the performance of the proposed scheduler in subsequent sections. Note that ϵ_k has the following property.

Corollary 3.2: The ϵ_k's are mutually independent and have zero-mean. For any $d \in \mathbb{Z}$, ϵ_{k-d} and e_k^s are independent. Consequently, $\mathbb{E}[e_k^s \epsilon_{k-d}'] = 0$ and $\mathbb{E}[\epsilon_k \epsilon_j'] = 0$ for all $k \neq j$.

Proof: The result follows easily from Lemma 3.1. ∎

In order to illustrate the idea of our proposed stochastic online scheduler, let us consider the special case when $\gamma_{k-1} = 1$ and $\epsilon_k = 0$. If the remote estimator is aware of the fact that $\epsilon_k = 0$, then even without receiving any data from the sensor, its state estimate $\hat{x}_k = A\hat{x}_{k-1}$ will have error covariance $\overline{P}$ instead of $h(\overline{P})$, which is greater than $\overline{P}$. This motivates us, similar to [5], to consider the following simple scheduler θ_{s} for a properly chosen $\delta > 0$:

$$\gamma_k = \begin{cases} 0, & \text{if } \|\epsilon_k\|_\infty < \delta, \\ 1, & \text{otherwise.} \end{cases} \tag{11}$$

It can be shown that when A is unstable and $\Psi \leq 1 - (1/\max_i |\lambda_i(A)|^2)$, we have $\lim_{k \to +\infty} \mathbb{E}[P_k] = +\infty$. The simple scheduler θ_{s} defined by (11) performs worse than $\theta_{\mathrm{off}}^\star$. Simulation result in Section VII also demonstrates this observation. To properly utilize the online information and avoid the divergence issue of θ_{s}, we propose the following stochastic scheduler $\theta_{\mathrm{e}}(N)$ with a time-out condition:

$$\gamma_{k+1} = \begin{cases} 0, & \text{if } \tau_k \leq N-1 \text{ and } \|\epsilon_{k+1}\|_\infty < \delta_{\tau_k+1}, \\ 1, & \text{otherwise} \end{cases} \tag{12}$$

where $N \in \mathbb{Z}_+$ and $\delta_i \in [0, +\infty]$, $i = 1, \ldots, N$, are design parameters. Note that under $\theta_{\mathrm{e}}(N)$, we restrict τ_k to take only finite values with a maximum value N, i.e., if the sensor has not communicated with the estimator for N consecutive time steps, then the sensor is forced to send its estimate to the estimator. The intuition behind this is that although the incremental innovation can be very small, the cumulative estimation error may increase. The time-out condition then guarantees that the cumulative effect is eliminated after at most N steps.

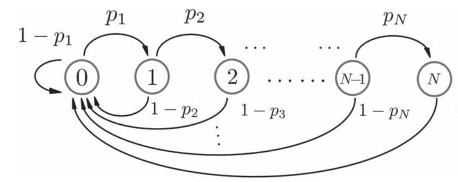


Fig. 2. A Markov chain modeling the evolution of τ_k.

It is easy to see that the τ_k's form a Markov chain with state space $\mathcal{S} = \{0, 1, \ldots, N\}$ as illustrated by Fig. 2, where $p_i \in [0, 1]$ is the state transition probability from $\tau_{k-1} = i - 1$ to $\tau_k = i$, and given by $p_i = \Pr(\|\epsilon_k\|_\infty < \delta_i)$. Note that for $\delta_i = +\infty$, we have $p_i \equiv 1$. The transition probability matrix $\mathbf{P}$ is given by

$$\mathbf{P} = \begin{bmatrix} 1-p_1 & p_1 & \cdots & 0 \\ \vdots & & \ddots & \vdots \\ 1-p_N & 0 & \cdots & p_N \\ 1 & 0 & \cdots & 0 \end{bmatrix}.$$

Since the Markov chain has finite states and is irreducible, it approaches the unique stationary distribution $\Pi \triangleq [\pi_0, \ldots, \pi_N]$ with an arbitrary initial state. Simple calculation yields that

$$\pi_j = \frac{\prod_{i=1}^j p_i}{\sum_{l=0}^N \prod_{i=1}^l p_i}, \quad j = 0, 1, \ldots, N \tag{13}$$

where $\prod_{i=1}^0 p_j \triangleq 1$. We have $p_i = \pi_i / \pi_{i-1}$ and $\pi_j \leq \pi_{j-1}$.

B. Estimation and Communication Under $\theta_{\mathrm{e}}(N)$

The following two theorems show how the remote estimator calculates $\hat{x}_k$ and the corresponding estimation error covariance P_k under the online scheduler $\theta_{\mathrm{e}}(N)$.

Theorem 3.3: The MMSE estimate of x_k is

$$\hat{x}_k = \begin{cases} \hat{x}_k^s, & \text{if } \gamma_k = 1, \\ A^\tau \hat{x}_{k-\tau}^s, & \text{if } \gamma_k = 0. \end{cases} \tag{14}$$

Proof: To make the presentation simple and clear, we only consider $\tau_k = 2$ in this proof as other cases can be proved in a similar way.

When $\gamma_k = 1$, it is obvious that $\hat{x}_k = \hat{x}_k^s$ since $\hat{x}_k^s$ is the MMSE estimate of x_k conditioning on $\mathcal{F}_k^s$. Now consider $\gamma_k = 0$. It suffices to show that $\mathbb{E}[x_k | \mathcal{Y}_{k-2}^s, \gamma_{k-1} = \gamma_k = 0] = A^2 \hat{x}_{k-2}^s$. At time $k-2$, the remote estimator receives $\hat{x}_{k-2}^s$ from the sensor, and $\gamma_{k-1} = \gamma_k = 0$ implies that $\|\epsilon_{k-1}\|_\infty < \delta_1$ and $\|\epsilon_k\|_\infty < \delta_2$. The Tower property gives

$$\begin{aligned}
&\mathbb{E}\left[x_k | \mathcal{Y}_{k-2}^s, \gamma_{k-1} = 0, \gamma_k = 0\right] \\
&= \mathbb{E}\left[\mathbb{E}\left[x_k | \mathcal{F}_k^s\right] | \mathcal{Y}_{k-2}^s, \gamma_{k-1} = 0, \gamma_k = 0\right] \\
&= \mathbb{E}\left[A^2 \hat{x}_{k-2}^s + AF'^{-1}\epsilon_{k-1} + F'^{-1}\epsilon_k | \mathcal{Y}_{k-2}^s, \gamma_{k-1} = \gamma_k = 0\right] \\
&= A^2 \hat{x}_{k-2}^s
\end{aligned}$$

where the last equality is due to that ϵ_{k-1} and ϵ_k are independent (Corollary 3.2) and their means are both zero. ∎

Theorem 3.4: Under $\theta_{\mathrm{e}}(N)$, when $\gamma_k = 1$, $P_k = \overline{P}$. And when $\gamma_k = 0$, P_k is given by

$$P_k = \overline{P} + \sum_{i=0}^{\tau_k - 1} \left[1 - \beta\left(\delta_{\tau_k - i}\right)\right]\left[h^{i+1}(\overline{P}) - h^i(\overline{P})\right] \tag{15}$$

where β is given by

$$\beta(\delta) \triangleq \frac{2}{\sqrt{2\pi}}\delta e^{-\frac{\delta^2}{2}}\left[1-2Q(\delta)\right]^{-1} \tag{16}$$

and $Q(\delta) \triangleq \int_\delta^{+\infty}(1/\sqrt{2\pi})e^{-(x^2/2)}\mathrm{d}x$ is the Q-function.

Proof: We only prove the case when $\gamma_k=0$ as the other case is straightforward to see. At the estimator, if no data are received at time k, then $\gamma_k=0$ and $\|\epsilon_{k-\tau+1}\|_\infty<\delta_1,\ldots,\|\epsilon_k\|_\infty<\delta_\tau$. Since $\hat{x}_k=A^\tau x^s_{k-\tau}$, we have

$$\begin{aligned}
&\mathbb{E}\left[(x_k-\hat{x}_k)(\cdot)'|\mathcal{Y}^s_{k-\tau},\gamma_{k-\tau+1}=0,\ldots,\gamma_k=0\right]\\
&=\mathbb{E}\left[\left(e^s_k+\varepsilon_k+A\varepsilon_{k-1}+\ldots+A^{\tau-1}\varepsilon_{k-\tau+1}\right)(\cdot)'|\right.\\
&\qquad\left.\mathcal{Y}^s_{k-\tau},\|\epsilon_{k-\tau+1}\|_\infty<\delta_1,\ldots,\|\epsilon_k\|_\infty<\delta_\tau\right]\\
&=\mathbb{E}\left[\mathbb{E}\left[\left(e^s_k+\varepsilon_k+A\varepsilon_{k-1}+\ldots+A^{\tau-1}\varepsilon_{k-\tau+1}\right)(\cdot)'|\mathcal{F}^s_k\right]\right.\\
&\qquad\left.|\mathcal{Y}^s_{k-\tau},\|\epsilon_{k-\tau+1}\|_\infty<\delta_1,\ldots,\|\epsilon_k\|_\infty<\delta_\tau\right]\\
&=\mathbb{E}\left[\left(e^s_k\right)\left(e^s_k\right)'|\mathcal{F}^s_k\right]+\mathbb{E}\left[\left(\varepsilon_k\right)\left(\varepsilon_k\right)'|\|\epsilon_k\|_\infty<\delta_\tau\right]\\
&\quad+A\mathbb{E}\left[(\varepsilon_{k-1})(\varepsilon_{k-1})'|\|\epsilon_{k-1}\|_\infty<\delta_{\tau-1}\right]A'+\ldots\\
&\quad+A^{\tau-1}\mathbb{E}\left[(\varepsilon_{k-\tau+1})(\varepsilon_{k-\tau+1})'|\|\epsilon_{k-\tau+1}\|_\infty<\delta_1\right]A'^{\tau-1}\\
&=\overline{P}+\sum_{i=0}^{\tau_k-1}\left[1-\beta(\delta_{\tau_k-i})\right]\left[h^{i+1}(\overline{P})-h^i(\overline{P})\right]\triangleq P_k
\end{aligned}$$

where the second equality is from the Tower property and the third one is due to the fact that $\varepsilon_{k-\tau+1},\ldots,\varepsilon_k$ are independent of $\mathcal{Y}^s_{k-\tau}$. ∎

Remark 3.5: In [5], a commonly used Gaussian approximation from nonlinear filtering is adopted to derive a simple recursive form of the remote state estimator. Theorem 3.4 gives a closed-form expression of the estimation error covariance without using any approximation.

The following lemma states the relations between $\lim_{k\to+\infty}\mathbb{E}[P_k]$ and $J(\theta_\mathrm{e}(N))$, π_0 and $\lim_{T\to+\infty}(1/T)\sum_{k=1}^T\gamma_k(\theta_\mathrm{e}(N))$, allowing us to replace the time average using the ensemble average.

Lemma 3.6: Under $\theta_\mathrm{e}(N)$, we have $\lim_{k\to+\infty}\mathrm{Tr}(\mathbb{E}[P_k])=J(\theta_\mathrm{e}(N))$ and $\lim_{T\to+\infty}(1/T)\sum_{k=0}^{T-1}\gamma_k(\theta_\mathrm{e}(N))=\pi_0$ $\mathbb{P}$-almost surely.

Proof: Define $f:\{0,\ldots,N\}\to\mathbb{R}$ as

$$f(j)=\begin{cases}\mathrm{Tr}(\overline{P}), & \text{if } j=0,\\ \mathrm{Tr}\left(\overline{P}+\sum\limits_{i=0}^{j-1}\left[1-\beta(\delta_{j-i})\right]\right.\\ \qquad\left.\left[h^{i+1}(\overline{P})-h^i(\overline{P})\right]\right), & \text{if } j\neq 0.\end{cases}$$

Since the Morkov chain $\{\tau_k\}$ has finite states and is irreducible and f is bounded, by ergodic theorem (see [14]) we immediately have the first assertion. If we define $f(j)=\begin{cases}1, & \text{if } j=0,\\ 0, & \text{if } j=1,\ldots,N\end{cases}$, the last assertion follows. ∎

Since τ_k approaches a limiting distribution Π, Theorem 3.4 and Lemma 3.6 together lead to that, $\mathbb{P}$-almost surely,

$$J\left(\theta_\mathrm{e}(N)\right)=\overline{P}+\sum_{j=1}^N\pi_j\sum_{i=0}^{j-1}\left[1-\beta(\delta_{j-i})\right]\left(h^{i+1}(\overline{P})-h^i(\overline{P})\right). \tag{17}$$

To make sure (7) holds, we have $\pi_0\le\Psi$. According to (13), N must satisfy $N\ge q$. If $N\le q-1$, we have $\pi_0=(\sum_{l=0}^N\prod_{i=1}^l p_i)^{-1}\ge q^{-1}$, which contradicts the definition of q in (8). We now construct a $\theta_\mathrm{e}(N)$ with a fixed $N\ge q$ such that $J(\theta_\mathrm{e}(N))$ is minimized while $\pi_0\le\Psi$, i.e, we consider the following problem:

$$\begin{aligned}
\min_{\theta_\mathrm{e}(N)}\quad & \mathrm{Tr}\left(\overline{P}+\sum_{j=1}^N\pi_j\sum_{k=0}^{j-1}\left[1-\beta(\delta_{j-k})\right]\left[h^{k+1}(\overline{P})-h^k(\overline{P})\right]\right)\\
\text{s.t.}\quad & p_i=\left[1-2Q(\delta_i)\right]^r,\ i=1,\ldots,N,\\
& \pi_j=\frac{\prod_{i=1}^j p_i}{\sum_{l=0}^N\prod_{i=1}^l p_i},\ j=0,\ldots,N,\quad \pi_0\le\Psi.
\end{aligned}$$

By selecting δ_i, $i=1,\ldots,N$, one can find an optimal schedule $\theta^\star_\mathrm{e}(N)$ which minimizes $J(\theta_\mathrm{e}(N))$. However, the objective and inequality constraints are non-convex functions and are difficult to be converted into convex ones. In the following section, we will replace the objective function by an upper bound and provide a suboptimal solution.

IV. Relaxation and Implementation

A. An Upper Bound for $J(\theta_\mathrm{e}(N))$

Lemma 4.1:

$$1-\beta(\delta)\le\left[1-2Q(\delta)\right]^2.$$

Proof: See Appendix. ∎

Replacing $1-\beta(\delta_i)$ by its upper bound $[1-2Q(\delta_i)]^2$, one can define one upper bound of $J(\theta_\mathrm{e}(N))$ as

$$\begin{aligned}
&J_u\left(\theta_\mathrm{e}(N)\right)\\
&\triangleq\mathrm{Tr}\left(\overline{P}+\sum_{j=1}^N\pi_j\sum_{k=0}^{j-1}\left(\frac{\pi_{j-k}}{\pi_{j-k-1}}\right)^{\frac{2}{r}}\left[h^{k+1}(\overline{P})-h^k(\overline{P})\right]\right).
\end{aligned}$$

The following proposition claims the existence of a $\theta_\mathrm{e}(N)$ such that $J_u(\theta_\mathrm{e}(N))\le J(\theta^\star_\mathrm{off})$.

Proposition 4.2: Let $N\ge q$. Then we always have that $\min_{\theta_\mathrm{e}(N)} J_u(\theta_\mathrm{e}(N))\le J(\theta^\star_\mathrm{off})$.

Proof: Consider a $\theta_\mathrm{e}(N)$ with the following parameters: $\delta_1=\delta_2=\cdots=\delta_{q-1}=+\infty$, $\delta_q=Q^{-1}((1-(u/(v-q))^{1/r})/2)$ and $\delta_{q+1}=\cdots=\delta_N=0$. When $u/v=q+1$, $J_u(\theta_\mathrm{e}(N))=J(\theta^\star_\mathrm{off})=(1/q)\sum_{i=0}^{q-1}\mathrm{Tr}(h^i(\overline{P}))$. Observe that when $q<u/v<q+1$, using the upper bound of $\beta(\delta)$, we obtain

$$\begin{aligned}
J_u\left(\theta_\mathrm{e}(N)\right)&=(1-\Psi q)\mathrm{Tr}\left(h^q(\overline{P})\right)+\Psi\sum_{i=0}^{q-1}\mathrm{Tr}\left(h^i(\overline{P})\right)\\
&\quad-(1-\Psi q)\left(1-\left[1-2Q(\delta)\right]^2\right)\mathrm{Tr}\left(h(\overline{P})-\overline{P}\right).
\end{aligned}$$

Then

$$\begin{aligned}
&J\left(\theta^\star_\mathrm{off}\right)-J_u\left(\theta_\mathrm{e}(N)\right)\\
&\qquad=(1-\Psi q)\left(1-\left[1-2Q(\delta)\right]^2\right)\mathrm{Tr}\left(h(\overline{P})-\overline{P}\right)
\end{aligned}$$

which shows $J_u(\theta_\mathrm{e}(N))\le J(\theta^\star_\mathrm{off})$ by Lemma 2.1 and completes the proof. ∎

B. Relaxation and Implementation

As shown in (13), Π is uniquely determined by $\mathbf{P}$ (hence by the p_i's). p_i's are uniquely determined by π_i's. Thus, $\theta_\mathrm{e}(N)$ can be alternatively represented from $\{\pi_0,\ldots,\pi_N\}$. By replacing the objective function in the previous problem with its upper bound and letting Π be the optimization variable, we obtain the following problem:

$$\min_{\Pi} \mathrm{Tr}\left(\overline{P}+\sum_{j=1}^{N}\pi_j\sum_{k=0}^{j-1}\left(\frac{\pi_{j-k}}{\pi_{j-k-1}}\right)^{\frac{2}{r}}\left[h^{k+1}(\overline{P})-h^{k}(\overline{P})\right]\right) \quad (18)$$

$$\text{s.t.} \sum_{j=0}^{N}\pi_j = 1, \quad \Psi \geq \pi_0 \geq \ldots \geq \pi_N \geq 0. \quad (19)$$

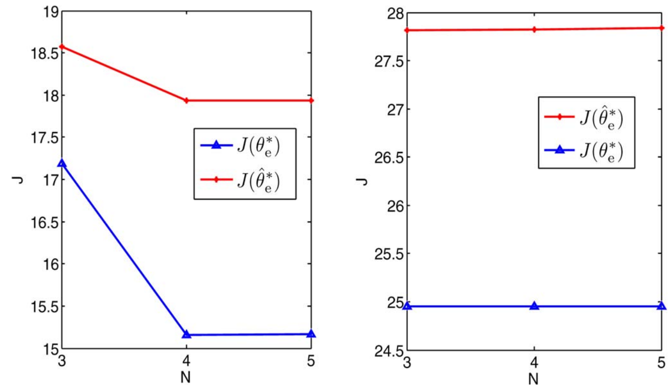
Denote by $\hat{\theta}_{\mathrm{e}}^{\star}(N)$ the optimal schedule for the above relaxed optimization problem. It can be shown that the schedule $\hat{\theta}_{\mathrm{e}}^{\star}(N)=\hat{\Pi}^{\star}$ must satisfy $\hat{\pi}_0^{\star}=\Psi$ and $\sum_{j=1}^{N}\hat{\pi}_j^{\star}=1-\Psi$. Otherwise, there must exist an integer l such that $l=\min_{1\leq j\leq N}\{j:\sum_{i=1}^{j}\hat{\pi}_i^{\star}\geq 1-\Psi\}$. Then choose $\varphi\in(0,1)$ satisfying $\sum_{i=1}^{N}\hat{\pi}_i^{\star}-(1-\Psi)=(1-\varphi)\sum_{i=l}^{N}\hat{\pi}_i^{\star}$. Consider a $\theta_{\mathrm{e}}(N)$ given by $\pi_0=\Psi$, $\pi_i=\varphi\hat{\pi}_i^{\star}$, $i=l,\ldots,N$ and other π_i's equal to those of $\hat{\Pi}^{\star}$. One has that $\pi_0>\hat{\pi}_0^{\star}$ and $\pi_j<\hat{\pi}_j^{\star}\forall j=l,\ldots,N$. Note that $J_u(\theta_{\mathrm{e}}(N))<J_u(\hat{\theta}_{\mathrm{e}}^{\star}(N))$, which contradicts the property of $\hat{\Pi}^{\star}$. Hence π_0 in (19) can be assumed to be equal to Ψ, i.e., $\pi_0=\Psi$, and (19) can be relaxed to an inequality constraint, i.e., $\sum_{j=1}^{N}\pi_j\geq 1-\Psi$. According to the above reasoning, $\hat{\Pi}^{\star}$ is still feasible for the relaxed problem. Since π_j is a proper fraction, we let p_j be the optimization variable to avoid numerical error made by inversion of π_j. Then $\pi_j=\pi_0\prod_{i=1}^{j}p_i$, and (18) and (19) are transformed into the following ones:

$$\min_{\{p_1,\ldots,p_N\}} \mathrm{Tr}\left(\overline{P}+\Psi\sum_{j=1}^{N}\prod_{i=1}^{j}p_i\sum_{k=0}^{j-1}p_{j-k}^{\frac{2}{r}}\left[h^{k+1}(\overline{P})-h^{k}(\overline{P})\right]\right),$$

$$\text{s.t.} \quad 1-\Psi-\Psi\sum_{j=1}^{N}\prod_{i=1}^{j}p_i\leq 0,\ p_j\in[0,1],\ j=1,\ldots,N.$$

This is a GGP problem characterized by posynomial objective and constraints that are the difference of monomials, which has a non-convex feasible region. There are a few optimization approaches attempting to attain the global optimum of GGP problems, such as [15], [16]. In [16], Maranas and Floudas provided a deterministic global optimization algorithm for GGP, which was applied to a number of small to medium size engineering design problems and was shown to ϵ-converge to the global minimal solution. To solve this problem efficiently, we adopt this algorithm. At each iteration, GGP is relaxed into a geometric program, which is solved in polynomial time complexity of N. The iteration time depends on convergence tolerance. Once p_j is obtained, we can compute δ_1 to δ_N by invoking (13) and (20). In this way $\hat{\theta}_{\mathrm{e}}^{\star}(N)$ is obtained. Note that in [9] and [4], this rate-error trade-off problem was posted as a Markov decision problem. As the cardinalities of state space and action space both grow with the system dimension, the computational complexity of solving these Markov decision problems becomes exponential in the system dimension. Our proposed scheduler is simple and can be easily computed, having a computational complexity that is independent of the system dimension.

Remark 4.3: As we mentioned previously, $N\geq q$ is necessary and sufficient for $\theta_e(N)$ to outperform $\theta_{\mathrm{off}}^{\star}$. How to choose the optimal N is an interesting question. In practice, one may simply choose $N=q$ for a guaranteed estimation quality (as we do in the simulations). In general, the larger N is, the better estimation quality we may expect. This, however, increases the computation complexity. Simulation examples show that once N is sufficiently large, the incremental estimation quality it improves is marginal.



(a) (b)

Fig. 3. The relation between N and $J(\hat{\theta}_{\mathrm{e}}^{\star}(N))$ for different A's.

V. Examples

Consider the following parameters for system (1), (2):

$$A=\begin{bmatrix}0.5 & 0.5\\ 0 & 1.5\end{bmatrix}, \quad C=[1 \quad 1], \quad Q=\begin{bmatrix}5 & 0\\ 0 & 5\end{bmatrix}, \quad R=5.$$

For $\Psi>0.35$, we have $q\leq 2$. To abide by the communication rate constraint, N should satisfy $N\geq 2$. In principle, $J(\theta_{\mathrm{e}}(N_1))\leq J(\theta_{\mathrm{e}}(N_2))$ for $N_1>N_2$ as we can always choose $\delta_i=0$, $i=N_2+1,\ldots,N_1$. The influence of N on $J(\theta_{\mathrm{e}}(N))$ depends on different system parameters, especially on how stable/unstable A is. The more unstable A is, the faster P_k will grow. Even with a large N, the scheduler $\theta_{\mathrm{e}}(N)$ tends not to let τ_k be large and the influence of N on $J(\theta_{\mathrm{e}}(N))$ is therefore marginal.

Let us take $\Psi=0.35$ as an example. In Fig. 3(a), when we consider a more stable A, e.g., $A=\begin{bmatrix}0.5 & 0.5\\ 0 & 1.2\end{bmatrix}$, a larger N introduces an obvious improvement to $J(\hat{\theta}_{\mathrm{e}}^{\star}(N))$. In Fig. 3(b), $J(\hat{\theta}_{\mathrm{e}}^{\star}(N))$ is almost constant as N increases. For the remainder, we choose $N=2$ by default to illustrate the main results. After solving a GGP problem, we obtain that $p_1=0.9442$ and $p_2=0.9668$. Therefore, $\delta_1=Q^{-1}((1-p_1^{1/r})/2)=1.9129$ and $\delta_2=Q^{-1}((1-p_2^{1/r})/2)=2.1299$. We compare this $\hat{\theta}_{e}^{\star}(N)$ with four other schedules:

1) $\theta_{\mathrm{off}}^{\star}$: the optimal offline sensor schedule given in Section II.
2) θ_{r}: a stochastic sensor schedule where γ_k's is an i.i.d sequence and $\mathbb{E}[\gamma_k]=\Psi$;
3) θ_{c}: the scheduler for a single sensor case when the sensor has an embedded local MMSE estimator in [8].[1]
4) The optimal $\theta_{\mathrm{e}}^{\star}(N)$ which minimizes $J(\theta_{\mathrm{e}}(N))$.

We compare $J(\theta_{\mathrm{off}}^{\star})$, $J(\hat{\theta}_{\mathrm{e}}^{\star}(N))$, $J(\theta_{\mathrm{r}})$, $J(\theta_{\mathrm{e}}^{\star}(N))$ and $J(\theta_{\mathrm{c}})$ for Ψ ranged from 0.35 to 1. As shown in Fig. 4, $J(\theta_{\mathrm{off}}^{\star})\geq J(\hat{\theta}_{\mathrm{e}}^{\star}(N))$. Under an equal communication rate using $\hat{\theta}_{\mathrm{e}}^{\star}(N)$ can reduce 43% estimation error compared to offline scheduling. Note that in this example $J(\hat{\theta}_{\mathrm{e}}^{\star}(N))$ agrees with $J(\theta_{\mathrm{e}}^{\star}(N))$, although it does not happen in general since $\hat{\Pi}^{\star}$ does not necessarily minimize $J_u(\theta_{\mathrm{e}}(N))$ and $J(\theta_{\mathrm{e}}(N))$ simultaneously, but at any rate, one can guarantee that $J(\theta_{\mathrm{r}}^{\star})\geq J(\theta_{\mathrm{off}}^{\star})\geq J(\hat{\theta}_{\mathrm{e}}^{\star}(N))$. $\hat{\theta}_{\mathrm{e}}^{\star}(N)$ offers a much better tradeoff between the sensor-estimator communication rate and the estimation quality. Note that Fig. 4 illustrates that $J(\theta_{\mathrm{off}}^{\star})$ is a piecewise affine function with respect to Ψ whose slope increases at each $\Psi=1/q$. Fig. 4 also shows that, for the parameters specified in this paper, $J(\theta_{\mathrm{c}})$

[1]The state of the system is directly observed in [8]. In this paper, the local event is defined as $\|\hat{x}_k^s-A\hat{x}_{k-1}\|_{\infty}\geq\delta$. We assume that the synchronization is operated every $T=2$ steps.

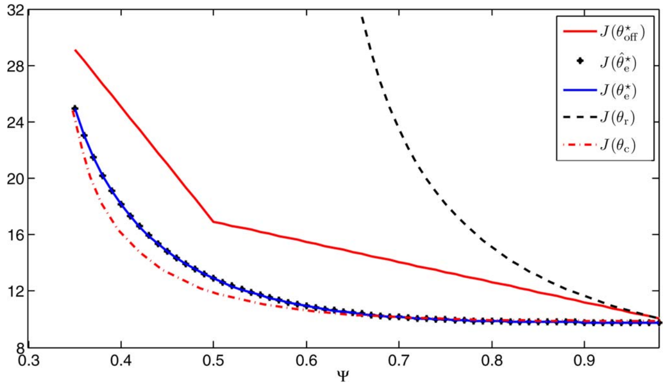


Fig. 4. Performance comparison.

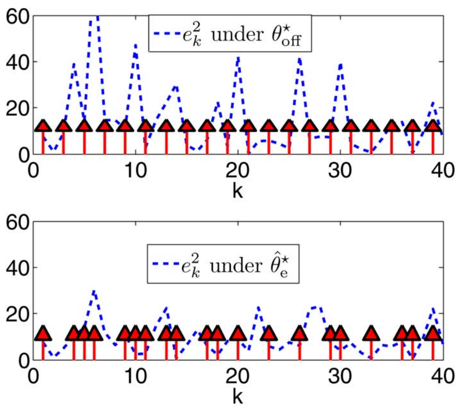


Fig. 5. Realization of $\theta^{\star}_{\mathrm{off}}$ and $\hat{\theta}_{\mathrm{e}}(N)^{\star}$ for $\Psi = 0.5$.

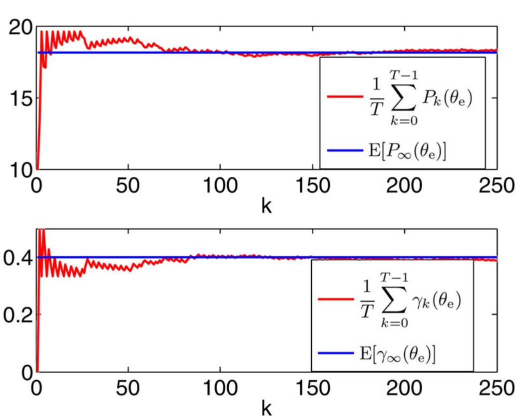


Fig. 6. Sample average vs time average.

is slightly better than $J(\hat{\theta}^{\star}_{\mathrm{e}}(N))$ when the communication rate is low, and the difference becomes negligible when the communication rate increase. An important thing to note is that in this work we give a closed-form expression of the relationship between the communication rate and the estimation error covariance; while to analyze the strategy θ_c in [8], it relies on numerical computations.

Fig. 6 shows that $(1/T)\sum_{k=0}^{k=T-1} J(\hat{\theta}^{\star}_{\mathrm{e}}(N))$ and $\lim_{k\to\infty}\mathbb{E}[P_k(\hat{\theta}^{\star}_{\mathrm{e}}(N)]$, $(1/T)\sum_{k=0}^{k=T-1}\gamma_k(\hat{\theta}^{\star}_{\mathrm{e}}(N))$ and π_0 are close after 100 steps, which agree with Lemma 3.6. We also plot in Fig. 5 a sample path of e_k^2 under $\theta^{\star}_{\mathrm{off}}$ and $\hat{\theta}^{\star}_{\mathrm{e}}(N)$ for $\Psi = 0.5$, respectively. In Fig. 5, a red arrow indicates a time instance k when $\gamma_k = 1$. Note that under the offline schedule, as $\Psi = 0.5$, the sensor transmission times are scheduled alternatively (and periodically). On the contrary, the sensor transmission times become random under $\hat{\theta}^{\star}_{\mathrm{e}}(N)$ which considers the importance of the sensor data and thereby reduces the estimation error and improves the estimation quality.

VI. Conclusion

We present an online sensor scheduler with time-out condition and show that it leads to a significant improved estimation quality when compared with the optimal offline sensor scheduler, a periodic data transmission strategy. This scheduler is simple and can be easily computed. Compared to [4] and [9], it has relatively light computational complexity for high-dimensional systems. In addition, under our proposed scheduler, the remote estimate is MMSE estimate and both the estimation error covariance matrix and communication rate are given in closed-form. Future work includes extensions to multi-sensor scheduling and close-loop control.

Appendix

Lemma A.1: $\beta(\delta)$ in (16) satisfies the following

1) $\beta(\delta)$ is strictly decreasing in δ.
2) $\lim_{\delta\to+\infty}\beta(\delta) = 0$ and $\lim_{\delta\to 0}\beta(\delta) = 1$.

The proof is straightforward from the definition and is omitted.

Lemma A.2: Let $\mathbf{x} \in \mathbb{R}$ be a standard Gaussian random variable. For $\delta > 0$, we have $\mathbb{E}[\mathbf{x^2}||\mathbf{x}| < \delta] = 1 - \beta(\delta)$.

Proof: The proof is straightforward from the conditional expectation and the definition of $\beta(\delta)$. ■

Proof to Lemma 4.1: It suffices to show $\varphi(\delta) \triangleq (1-\beta(\delta))/[1-2Q(\delta)]^2 \le 1, \forall\delta > 0$. Take derivative of $\varphi(\delta)$, i.e.

$$\frac{\mathrm{d}\varphi(\delta)}{\mathrm{d}\delta} = \frac{\frac{2}{\sqrt{2\pi}}e^{-\frac{\delta^2}{2}}}{[1-2Q(\delta)]^4}\left\{\delta^2\left[1-2Q(\delta)\right] - 3\int_{-\delta}^{\delta}\frac{t^2 e^{-\frac{t^2}{2}}}{\sqrt{2\pi}}\mathrm{d}t\right\}.$$

Let $\delta^2[1-2Q(\delta)] - 3\int_{-\delta}^{\delta}(t^2/\sqrt{2\pi})e^{-(t^2/2)}\mathrm{d}t \triangleq \phi(\delta)$. One obtains

$$\frac{\mathrm{d}\phi(\delta)}{\mathrm{d}\delta} = 2\delta\left(\int_{-\delta}^{\delta}\frac{1}{\sqrt{2\pi}}e^{-\frac{t^2}{2}}\mathrm{d}t - \frac{2\delta}{\sqrt{2\pi}}e^{-\frac{\delta^2}{2}}\right) > 0,$$

since $\int_{-\delta}^{\delta}(1/\sqrt{2\pi})e^{-(t^2/2)}\mathrm{d}t - (2\delta/\sqrt{2\pi})e^{-(\delta^2/2)} > 0$ for any $\delta > 0$. Therefore $\phi(\delta) > 0$ observing that $\phi(0) = 0$. Note that $(2/\sqrt{2\pi})e^{-(\delta^2/2)}/[1-2Q(\delta)]^4 > 0$ leads to $\mathrm{d}\varphi(\delta)/\mathrm{d}\delta > 0$. Altogether with the fact that

$$\lim_{\delta\to+\infty}\varphi(\delta) = \frac{\lim_{\delta\to+\infty}[1-\beta(\delta)]}{\lim_{\delta\to+\infty}[1-2Q(\delta)]^2} = 1$$

one has $\varphi(\delta) \le 1$. ■

Lemma A.3: Let $\xi \in \mathbb{R}^r$ be a Gaussian random vector with zero mean and $\mathbb{E}[\xi\xi'] = I_r$. For $\delta > 0$, we have $\Pr(\|\xi\|_\infty < \delta) = [1-2Q(\delta)]^r$ and $\mathbb{E}[\xi\xi'|\|\xi\|_\infty < \delta] = [1-\beta(\delta)]I_r$.

Proof: Denote ξ_i as the ith element of $\xi (0 \le i \le r)$. Each ξ_i is then a standard Gaussian random variable with zero mean and unit variance, and ξ_i, ξ_j are mutually independent if $i \ne j$. Note that $\|\xi\|_\infty < \delta$ iff $|\xi_i| < \delta$. We get $\Pr(\|\xi\|_\infty < \delta) = \prod_{i=0}^{r}\Pr(|\xi_i| \le \delta) = [1-2Q(\delta)]^r$. The second equation follows from Lemma A.2. ■

Lemma A.4:

$$\mathbb{P}\left(\|\epsilon_k\|_\infty \ge \delta\right) = 1 - [1-2Q(\delta)]^r,$$

$$\mathbb{E}\left[\varepsilon_k\varepsilon_k'|\|\epsilon_k\|_\infty < \delta\right] = [1-\beta(\delta)]\left[h(\overline{P}) - \overline{P}\right]. \qquad (20)$$

Proof: Let $\epsilon_k = \begin{bmatrix} \xi_k^1 \\ \xi_k^2 \end{bmatrix}$ where $\xi_k^1 \in \mathbb{R}^r$ and $\xi_k^2 \in \mathbb{R}^{n-r}$. Since $\mathbb{E}[\varepsilon_k \varepsilon_k'] = h(\overline{P}) - \overline{P}$, we have

$$\mathbb{E}\left[\epsilon_k \epsilon_k'\right] = F' \mathbb{E}\left[\varepsilon_k \varepsilon_k'\right] F = \mathrm{diag}(I_r, \mathbf{0}_{n-r}).$$

Hence ξ_k^1 is a zero mean Gaussian random vector with unit variance, and $\xi_k^2 = 0$ $\mathbb{P}$-almost surely. Lemma A.3 leads to (20) and

$$\mathbb{E}\left[\epsilon_k \epsilon_k' | \|\epsilon_k\|_\infty < \delta\right] = [1 - \beta(\delta)]\, \mathrm{diag}(I_r, \mathbf{0}_{n-r}).$$

Since $U' = U^{-1}$, one has

$$\begin{aligned} \mathbb{E}\left[\varepsilon_k \varepsilon_k' | \|\epsilon_k\|_\infty < \delta\right] &= (F')^{-1} \mathbb{E}\left[\epsilon_k \epsilon_k' | \|\epsilon_k\|_\infty < \delta\right] F^{-1} \\ &= [1 - \beta(\delta)]\, U \mathrm{diag}(\Lambda, \mathbf{0}_{n-r}) U' \\ &= [1 - \beta(\delta)] \left[h(\overline{P}) - \overline{P}\right] \end{aligned}$$

which completes the proof. ∎

Acknowledgment

The authors would like to thank L. Schenato for helpful discussions.

References

[1] S. Trimpe and R. D. Andrea, "Event-based state estimation with variance-based triggering," in *Proc. 51st IEEE Conf. Decision Control*, 2012, pp. 6583–6590.

[2] H. Sandberg, M. Rabi, M. Skoglund, and K. H. Johansson, "Estimation over heterogeneous sensor networks," in *Proc. 47th IEEE Conf. Decision Control*, 2008, pp. 4898–4903.

[3] C. O. Savage and B. F. L. Scala, "Optimal scheduling of scalar Gauss-Markov systems with a terminal cost function," *IEEE Trans. Autom. Control*, vol. 54, no. 5, pp. 1100–1105, 2009.

[4] L. Li, M. Lemmon, and X. Wang, "Event-triggered state estimation in vector linear processes," in *Proc. Amer. Control Conf.*, 2010, pp. 2138–2143.

[5] J. Wu, Q.-S. Jia, K. Johansson, and L. Shi, "Event-based sensor data scheduling: Trade-off between communication rate and estimation quality," *IEEE Trans. Autom. Control*, vol. 58, no. 4, pp. 1041–1046, 2013.

[6] J. Sijs and M. Lazar, "Event based state estimation with time synchronous updates," *IEEE Trans. Autom. Control*, vol. 57, no. 10, pp. 2650–2655, 2012.

[7] B. Demirel, V. Gupta, and M. Johansson, "On the trade-off between control performance and communication cost for event-triggered control over lossy networks," in *Proc. Eur. Control Conf.*, Zurich, Switzerland, Jul. 2013.

[8] M. Xia, V. Gupta, and P. Antsaklis, "Networked state estimation over a shared communication medium," in *Proc. Amer. Control Conf.*, 2013, pp. 4128–4133.

[9] Y. Xu and J. Hespanha, "Optimal communication logics in networked control systems," in *Proc. 43rd IEEE Conf. Decision Control*, 2004, vol. 4, pp. 3527–3532.

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

[11] L. Shi, K. H. Johansson, and L. Qiu, "Time and event-based sensor scheduling for networks with limited communication resources," in *Proc. 18th World Congress Int. Fed. Autom. Control*, 2011, pp. 13 263–13 268.

[12] L. Shi, P. Cheng, and J. Chen, "Sensor data scheduling for optimal state estimation with communication energy constraint," *Automatica*, vol. 47, no. 8, pp. 1693–1698, 2011.

[13] J. Wu, Y. Yuan, H. Zhang, and L. Shi, "How can online schedules improve communication and estimation tradeoff?" *IEEE Trans. Signal Processing*, vol. 61, no. 7, pp. 1625–1631, 2013.

[14] L. R. Bellet, "Ergodic properties of markov processes," in *Open Quantum Syst. II*. New York: Springer, 2006, pp. 1–39.

[15] C. A. Floudas and V. Visweswaran, "A global optimization algorithm (GOP) for certain classes of nonconvex NLPs-I. Theory," *Comp. Chem. Eng.*, vol. 14, no. 12, pp. 1397–1417, 1990.

[16] C. D. Maranas and C. A. Floudas, "Global optimization in generalized geometric programming," *Comp. Chem. Eng.*, vol. 21, no. 4, pp. 351–569, 1997.