State estimation with remote sensors under limited communication rate

G. Battistelli[a], A. Benavoli[b], L. Chisci[a]

[a] *Dipartimento di Sistemi e Informatica, Università di Firenze 50139 Firenze, Italy*
[b] *Istituto "Dalle Molle" di Studi sull'Intelligenza Artificiale, 6928 Manno-Lugano, Switzerland*

Abstract

This paper deals with the problem of estimating the state of a discrete-time linear stochastic dynamical system on the basis of data collected from multiple sensors subject to a limitation on the communication rate from the remote sensor units. The optimal probabilistic measurement-independent strategy for deciding when to transmit estimates from each sensor is derived. Simulation results show that the derived strategy yields certain advantages in terms of worst-case time-averaged performance with respect to periodic ones when coordination among sensors is not possible.

Keywords: Networked estimation, sensor fusion, optimal filtering, Markov chains.

1. Introduction

This paper deals with the problem of estimating the state of a discrete-time linear stochastic dynamical system

$$\mathbf{x}_{k+1} = \mathbf{A}\mathbf{x}_k + \mathbf{w}_k \tag{1}$$

on the basis of measurements collected from multiple sensors

$$\mathbf{y}_k^i = \mathbf{C}^i\mathbf{x}_k + \mathbf{v}_k^i \tag{2}$$

for $i \in \mathcal{S} \triangleq \{1, 2, \ldots, S\}$ subject to a limitation on the communication rate from each remote sensor unit to the state estimation unit. More specifically, the attention will be focused on the use of a sensor network consisting of: S remote *sensing nodes* which collect noisy measurements of the given system, can process them to find filtered estimates and transmit such estimates at a reduced communication rate; a *fusion node* F which receives data from the S sensors and, based on such data, should estimate, in the best possible way, the system's state.

The objective is to devise a *transmission strategy* (TS) with fixed rate that guarantees bounded state covariance and possibly optimal estimation performance, in terms of minimum *Mean Square Error* (MSE), at the node F.

The above scenario reflects the practical situation in which the sensing units and the monitoring unit are remotely dislocated one with respect to each other and the communication rate between them is severely limited by energy, band and/or security concerns. This happens, for example, in wireless sensor networks wherein every transmission typically reduces the lifetime of the sensor devices, wireless communication being the major source of energy consumption [1]. Further, a reduction in the sensors data transmission rate can be crucial in networked control systems in order to reduce the network traffic and hopefully avoid congestion [2].

State estimation under finite communication bandwidth has been thoroughly investigated, see e.g. [3, 4, 5]. In the above cited references, the emphasis is on the analysis of the quantization effects due to the encoding of transmitted data into a finite alphabet of symbols as well as on the design of efficient, possibly optimal, coding algorithms. Conversely, following [2, 6, 7, 8], the present work tackles the issue of communication bandwidth finiteness from a completely different viewpoint. Specifically, it is assumed that infinite-precision data are transmitted over the communication channel[1] while the bandwidth limitation is accomplished by imposing a suitable value of the transmission rate. In this context, the focus will be on the choice of a decentralized TS for deciding which data transmit from the remote sensors to F. It is pointed out that the idea of controlling data transmission in a networked control system so as to achieve a tradeoff between communication and estimation performance is not novel in the literature, see [2, 6, 7, 9, 10, 11] and the references therein. Further, a somehow related problem is the so called sensor scheduling problem, wherein, at every time step, the fusion node selects only a subset of the available sensors to receive the information [12, 13, 14].

In this paper, probabilistic decentralized measurement-independent strategies are considered and issues related to the stability and optimality of such strategies are investigated. Specifically, attention is devoted to TS's wherein the time intervals between consecutive transmissions are random variables governed by a finite-state Markov chain. Notice that a similar TS has been analyzed in [15] in the context of a model-based networked control system. Further, in the signal processing literature, the idea of randomly varying the time between consecutive measurements, commonly referred to as *additive random sampling* [16], has been extensively studied in order to overcome the aliasing problem. Finally, additive random sampling is a common practice in Internet flow monitoring to avoid synchronization problems [17].

In the present context, the main motivation for considering this kind of probabilistic TS's stems for the observation that in the multisensor case, supposing that the information provided by different sensors be mutually correlated, the performance of periodic TS's can depend in a crucial way on the phase shift among the sensors and coordination would be needed so as to ensure that sensor transmissions be well distributed over time. For example, when all sensors are identical, performance is optimized when sensor transmissions are uniformly distributed over time, whereas a degradation of the estimation performance at the fusion node is expected when all sensors periodically transmit at the very same time instants. Thus, in a purely decentralized setting wherein sensor coordination is not allowed, the performance of periodic TS's can vary significantly for different realizations. In other words, periodic multisensor TS's induce an undesired "non ergodic" behaviour. As will be shown, this effect can be counteracted by considering probabilistic TS's generated via aperiodic Markov chains so as to ensure that ergodicity holds and, thus, the long-run average performance be always realization-independent. In this connection, the main contribution of the present paper lies in showing that probabilistic strategies yield certain advantages with respect to periodic ones when coordination

[1]This assumption, though incompatible with the finite bandwidth, holds in practice provided that quantization errors are negligible with respect to measurement errors.

among sensors is not possible and provided that the transmission probabilities be adequately chosen.

Preliminary work has been carried out in [9] which addressed the single-sensor case and in [10] which formulated the optimal multisensor transmission problem with reference to a generic fusion algorithm and discussed possible suboptimal solutions. The main contributions of the present paper with respect to the above cited references are as follows: the multisensor TS problem is formulated considering *covariance intersection* [18, 19] as data fusion rule and exploiting local standard detectability forms so as to deal with possible local undetectability issues; a closed-form is derived for the multisensor TS that minimizes the average MSE at the fusion node; issues concerning the misbehavior of non-ergodic (periodic) multisensor TS's are thoroughly worked out, also via simulation experiments, and provably optimal aperiodic perturbations of the periodic strategy are developed which avoid such a misbehavior with a minimal increase of the average cost.

The notations are quite standard: $\mathbb{Z}_+$ is the set of nonnegative integers; given a square matrix $\mathbf{M}$, $\mathrm{tr}(\mathbf{M})$ and $\mathbf{M}'$ denote its trace and, respectively, transpose; $\mathbb{E}\{\cdot\}$ and $\mathbb{P}\{\cdot\}$ denote the expectation and, respectively, probability operators; finally, given a vector-valued sequence $\{\mathbf{z}_k; k = 0, 1, \ldots\}$, $\mathbf{z}_{t_1:t_2}$ stands for its restriction to the time interval $\{t_1, t_1 + 1, \ldots, t_2\}$.

2. Communication strategy

The aim of this section is to formalize the concept of *transmission strategy (TS)* with fixed rate α^i for sensor i. To this end, let us introduce for each sensor i binary variables c_k^i such that

$$c_k^i = \begin{cases} 1, & \text{if sensor } i \text{ transmits at time } k \\ 0, & \text{if sensor } i \text{ does not transmit at time } k \end{cases}$$

The attention is restricted to estimate transmission, assuming that each sensor node i has enough processing capability to update on-line the optimal state estimate $\hat{\mathbf{x}}_{k|k}^i$. Note that, with such a choice, the loss of information due to the finite bandwidth is mitigated as the estimate $\hat{\mathbf{x}}_{k|k}^i$ somehow summarizes the information collected by sensor i up to time k being the center of the local posterior PDF. As far as the decision mechanism is concerned, this can be formally defined as follows.

Definition 1. *A decision mechanism with rate $\alpha^i \in (0, 1)$ is any, deterministic or stochastic, mechanism of generating c_k^i such that*

$$\lim_{t\to\infty} \frac{1}{t} \sum_{k=1}^{t} c_k^i = \alpha^i \tag{3}$$

Several decision mechanisms can clearly be devised. In this work, the attention will be restricted to *measurement-independent* strategies that, at time k, decide whether to transmit or not independently of the measurement sequence $\mathbf{y}_{0:k}^i$. More specifically, it is supposed that each c_k^i is chosen according to some probabilistic criterion, possibly adapted only on the basis of the past transmission pattern, so as to ensure the desired communication rate. In this respect, let $n_k^i \geq 0$ denote, at a generic time k, the number of time instants elapsed from the last transmission of sensor i, i.e.

$$c^i_{k-n^i_k} = 1 \text{ and } c^i_k = c^i_{k-1} = \cdots = c^i_{k-n^i_k+1} = 0.$$

Note that n_k^i is a function of $c^i_{0:k}$ and can be recursively computed as

$$n_k^i = (1 - c_k^i)\,(n^i_{k-1} + 1)\,.$$

It should be evident that, in the considered framework, the greater n_k^i the more outdated is the last estimate $\hat{x}_{k-n^i_k|k-n^i_k}$ received at the fusion node F from sensor i. Then it seems reasonable to take into account the value of n^i_{k-1} when choosing whether to transmit or not at the generic time k. In this connection, the conditional probabilities

$$\mathbb{P}\{c_k^i = 1 | n^i_{k-1} = j\}\,, \quad j = 0, 1, \ldots$$

can be considered as design parameters in the communication strategy that can be suitably tuned to improve performance (e.g., to reduce the MSE).

Specifically, at the generic time k, c_k^i is chosen to be a Bernoulli random variable with parameter $\varphi^i(n^i_{k-1})$, i.e., c_k^i takes value 1 with probability $\varphi^i(n^i_{k-1})$ and value 0 with probability $1 - \varphi^i(n^i_{k-1})$. Clearly, this corresponds to

$$\mathbb{P}\{c_k^i = 1 | n^i_{k-1} = j\} = \varphi^i(j)\,,$$

for $j = 0, 1, \ldots$ and $i = 1, \ldots, S$. The functions $\varphi^i : \mathbb{Z}_+ \to [0, 1]$ must be chosen so that the transmission rate constraint is met for each sensor i.

Throughout the paper, the following notation will be adopted

$$\mathbf{c}_k \triangleq \text{col}\left(c_k^1, \ldots, c_k^S\right)\,, \quad \mathbf{n}_k \triangleq \text{col}\left(n_k^1, \ldots, n_k^S\right)\,, \quad \mathbf{y}_k \triangleq \text{col}\left(\mathbf{y}_k^1, \ldots, \mathbf{y}_k^S\right)\,.$$

3. Fusion algorithm

This section is devoted to the description of the operations that are performed in the sensor nodes $1, \ldots, S$ as well as in the fusion node F in order to recover a fused estimate.

Let $\mathbf{Q}$ and $\mathbf{R}^i$ be the covariances of $\mathbf{w}_k$ and, respectively, $\mathbf{v}_k^i$. Further, let $\mathbf{C} \triangleq \text{col}\,(\mathbf{C}_1, \ldots, \mathbf{C}_S)$. In order to ensure the well-posedness of the state estimation problem, the following preliminary assumption is needed.

A1. $(\mathbf{A}, \mathbf{Q}^{1/2})$ is stabilizable, $(\mathbf{A}, \mathbf{C})$ is detectable, and $\mathbf{R}^i > 0$ for $i = 1, \ldots, S$.

3.1. Sensor nodes operation

To enable the estimate transmission option, it is assumed that each sensor i has enough processing capabilities to compute the optimal state estimate $\hat{\mathbf{x}}^i_{k|k}$ on the basis of the measurements $\mathbf{y}^i_{0:k}$ gathered at time k. In this connection, a first important observation is that assumption A1 only ensures *collective* detectability from the whole sensor network, while detectability of each pair $(\mathbf{A}, \mathbf{C}_i)$ need not necessarily hold. Thus, when for sensor i the latter property does not hold, it is not possible to reliably estimate all the state vector $\mathbf{x}_k$ only on the grounds of the locally collected measurements $\mathbf{y}^i_{0:k}$. In order to avoid possible estimation error divergence phenomena, it is therefore convenient to put the system (1)-(2) into standard detectability form by separating the part of the system that is both unobservable from the local measurements $\mathbf{y}^i_{0:k}$ and unstable. By invoking classical results on linear system theory, this can always be done by means of a suitable nonsingular transformation matrix $\mathbf{T}^i$ so that the transformed system matrices take the form

$$(\mathbf{T}^i)^{-1}\mathbf{A}\mathbf{T}^i = \begin{bmatrix} \mathbf{A}^i_1 & 0 \\ \mathbf{A}^i_{21} & \mathbf{A}^i_2 \end{bmatrix}, \quad \mathbf{C}^i\mathbf{T}^i = \begin{bmatrix}\mathbf{C}^i_1 & 0\end{bmatrix}, \quad (\mathbf{T}^i)^{-1} = \begin{bmatrix} \mathbf{D}^i_1 \\ \mathbf{D}^i_2 \end{bmatrix}$$

with $(\mathbf{A}^i_1, \mathbf{C}^i_1)$ detectable by construction. Then, one can restrict the attention to the detectable subsystem with state $\mathbf{x}^i_k \stackrel{\triangle}{=} \mathbf{D}^i_1\mathbf{x}_k$ and equations

$$\begin{aligned} \mathbf{x}^i_{k+1} &= \mathbf{A}^i_1\mathbf{x}^i_k + \mathbf{D}^i_1\mathbf{w}_k && (4) \\ \mathbf{y}^i_k &= \mathbf{C}^i_1\mathbf{x}^i_k + \mathbf{v}^i_k\,. && (5) \end{aligned}$$

As for estimation of $\mathbf{x}^i_k$, it is well known that for a linear system the optimal estimates and covariances in the minimum mean square sense are recursively provided by the Kalman Filter

$$\left\{ \begin{aligned} \mathbf{P}^i_{k|k-1} &= \mathbf{A}^i_1\mathbf{P}^i_{k-1|k-1}(\mathbf{A}^i_1)' \ + \ \mathbf{D}^i_1\mathbf{Q}(\mathbf{D}^i_1)' \\ \mathbf{K}^i_k &= \mathbf{P}^i_{k|k-1}(\mathbf{C}^i_1)'\left(\mathbf{R}^i + \mathbf{C}^i_1\mathbf{P}^i_{k|k-1}(\mathbf{C}^i_1)'\right)^{-1} \\ \hat{\mathbf{x}}^i_{k|k} &= \mathbf{A}^i_1\,\hat{\mathbf{x}}^i_{k-1|k-1} + \mathbf{K}^i_k(\mathbf{y}^i_k - \mathbf{C}^i_1\,\mathbf{A}^i_1\,\hat{\mathbf{x}}^i_{k-1|k-1}) \\ \mathbf{P}^i_{k|k} &= \mathbf{P}^i_{k|k-1} - \mathbf{P}^i_{k|k-1}(\mathbf{C}^i_1)'\left(\mathbf{R}^i + \mathbf{C}^i_1\mathbf{P}^i_{k|k-1}(\mathbf{C}^i_1)'\right)^{-1}\mathbf{C}^i_1\mathbf{P}^i_{k|k-1} \end{aligned} \right. \qquad (6)$$

As well known the covariance matrices of the Kalman Filter do not depend on the sequence of measurements and, hence, can be autonomously computed by the node F so that only transmission of estimates but no transmission of covariances from each sensor node to F is needed.

Further, since the pair $(\mathbf{A}^i_1, \mathbf{C}^i_1)$ is detectable by construction, the a-priori state covariance $\mathbf{P}^i_{k|k-1}$ and the a-posteriori covariance $\mathbf{P}^i_{k|k}$ exponentially converge to the steady-state values $\mathbf{P}^i_b$ and $\mathbf{P}^i_a$, respectively, where $\mathbf{P}^i_b$ is the unique positive definite solution of the algebraic Riccati equation

$$\mathbf{P}^i_b = \mathbf{A}^i_1\left[\mathbf{P}^i_b - \mathbf{P}^i_b(\mathbf{C}^i_1)'\left(\mathbf{R}^i + \mathbf{C}^i_1\mathbf{P}^i_b(\mathbf{C}^i_1)'\right)^{-1}\mathbf{C}^i_1\mathbf{P}^i_b\right](\mathbf{A}^i_1)' + \mathbf{D}^i_1\mathbf{Q}(\mathbf{D}^i_1)' \qquad (7)$$

and $\mathbf{P}^i_a$ is given by

$$\mathbf{P}^i_a = \mathbf{P}^i_b - \mathbf{P}^i_b(\mathbf{C}^i_1)'\left(\mathbf{R}^i + \mathbf{C}^i_1\mathbf{P}^i_b(\mathbf{C}^i_1)'\right)^{-1}\mathbf{C}^i_1\mathbf{P}^i_b\,. \qquad (8)$$

3.2. Fusion node operation

For the fusion node F, the available information at time k is given by the vector of elapsed times $\mathbf{n}_k$ and the estimates

$$\hat{\mathbf{x}}^i_{k|k-n^i_k} = (\mathbf{A}^i_1)^{n^i_k}\,\hat{\mathbf{x}}^i_{k-n^i_k|k-n^i_k}$$

for $i = 1, \ldots, S$.

It is worth noting that, for non-unitary transmission rates, an explicit formula for recursively computing the optimal fused estimate from the sequence of local received estimates is not available. This means that the determination of the optimal fused estimate would entail a growing memory (since, at each time k, all the received local estimates, from time 0 up to the current time k, should be considered) and clearly would eventually become computationally unfeasible. Thus, different suboptimal fusion algorithms have been devised to combine data collected from multiple sensors and related information in order to derive a fused estimate $\hat{\mathbf{x}}_{k|k}$ of the true state $\mathbf{x}_k$ (see [20] and the references therein). In this paper, we shall focus on the *Covariance Intersection* (CI) algorithm [18, 19] since it usually provides a good tradeoff between performance, practicality, and cost [21]. A definite advantage of CI with respect to other conceivable fusion rules is that it does not require to know the correlations between the measurement noises of different sensors.

In CI, the fused estimate $\hat{\mathbf{x}}_{k|k}$ and covariance $\mathbf{P}_{k|k}$ are obtained by means of a convex combination of the local estimates $\hat{\mathbf{x}}^i_{k|k-n^i_k}$ and covariances $\mathbf{P}^i_{k|k-n^i_k}$ as follows[2]

$$\hat{\mathbf{x}}_{k|k} = \mathbf{P}_{k|k} \sum_{i=1}^{S} \omega^i_k (\mathbf{D}^i_1)' \, (\mathbf{P}^i_{k|k-n^i_k})^{-1} \, \hat{\mathbf{x}}^i_{k|k-n^i_k} \tag{9}$$

$$\mathbf{P}^{-1}_{k|k} = \sum_{i=1}^{S} \omega^i_k (\mathbf{D}^i_1)' \, (\mathbf{P}^i_{k|k-n^i_k})^{-1} \mathbf{D}^i_1 \tag{10}$$

where the scalar weights ω^i_k satisfy the constraints

$$\omega^i_k \geq 0, \; i = 1, \ldots, S, \quad \text{and} \quad \sum_{i=1}^{S} \omega^i_k = 1\,. \tag{11}$$

Notice that the scalar weights ω^i_k represent degrees of freedom which allow the estimate to be optimised with respect to different cost functions. For example such parameters can be chosen by minimizing the trace of the combined state covariance, i.e., by solving the optimization problem

$$\min_{\omega^1_k, \ldots, \omega^S_k} \operatorname{tr} \left\{ \mathbf{P}_{k|k} \mathbf{W} \right\} \tag{12}$$

where $\mathbf{W}$ is a positive definite weight matrix. The resulting fusion rule will be referred to as *Optimal CI* (OCI). Hereafter, in order to avoid unnecessary complications, the solution of (12) will always be assumed to be unique. Otherwise, one of the solutions can be arbitrarily chosen.

In case solving (12) at every sampling time is not computationally feasible, one can adopt a suboptimal choice like

$$\omega^i_k = 1/S. \tag{13}$$

This choice will be referred to as *Approximate CI* (ACI).

[2]Notice that equations (9)-(10) account also for the fact that each local estimate $\hat{\mathbf{x}}^i_{k|k-n^i_k}$ is referred to the locally detectable state components $\mathbf{x}^i_k$ which are related to the state vector $\mathbf{x}_k$ through equation $\mathbf{x}^i_k = \mathbf{D}^i_1 \mathbf{x}_k$.

4. Problem statement

The main positive feature of CI is that it yields a *consistent* estimate [22], regardless of the choice of the scalar weights ω_k^i satisfying (11), in that, for any given transmission pattern $\mathbf{c}_{0:k}$, the estimated covariance $\mathbf{P}_{k|k}$ is always an upper bound (in the positive definite sense) of the true error covariance, i.e.,

$$\mathbb{E}\left\{(\mathbf{x}_k - \hat{\mathbf{x}}_{k|k})(\mathbf{x}_k - \hat{\mathbf{x}}_{k|k})' | \mathbf{c}_{0:k}\right\} \leq \mathbf{P}_{k|k} \tag{14}$$

where the expectation is taken with respect to $\mathbf{x}_0, \mathbf{w}_{0:k}$, and $\mathbf{v}_{0:k}$. Notice that, in the considered framework, thanks to the linearity of the system the fused covariance $\mathbf{P}_{k|k}$ depends only on the transmission pattern $\mathbf{c}_{0:k}$. In this connection, in what follows we shall address minimization of the long-run average cost functional

$$J(\mathbf{c}_{0:\infty}) = \lim_{t\to\infty} \frac{1}{t} \sum_{k=1}^{t} \mathrm{tr}\left\{\mathbf{P}_{k|k}\mathbf{W}\right\} \tag{15}$$

with the weight matrix $\mathbf{W}$ as in (12). In fact, by (14), (15) provides an upper bound for

$$\lim_{t\to\infty} \frac{1}{t} \sum_{k=1}^{t} \mathbb{E}\left\{(\mathbf{x}_k - \hat{\mathbf{x}}_{k|k})' \,\mathbf{W}\, (\mathbf{x}_k - \hat{\mathbf{x}}_{k|k}) | \mathbf{c}_{0:k}\right\} .$$

More specifically, taking into account the transmission rate constraint (3), the following constrained optimization problem can be stated.

Problem 1. *Find the optimal transmission probabilities $\varphi^i(j)$, $i = 1, \ldots, S$, $j = 0, 1, \ldots$ that minimize the worst-case long-run average cost*

$$\sup_{\mathbf{c}_{0:\infty}} J(\mathbf{c}_{0:\infty}) \tag{16}$$

under the constraints (3), $i = 1, \ldots, S$, and

$$0 \leq \varphi^i(j) \leq 1\,, \quad j = 0, 1, \ldots, \quad i = 1, \ldots, S\,.$$

It is worth noting that the supremum in (16) accounts for the fact that, given the transmission probabilities, the long-run average in (15) may vary with the realization of the transmission pattern $\mathbf{c}_{0:\infty}$ (for instance, it may depend on the phase shifts among the sensors). This issue will be discussed in some detail later on.

Remark 1. *A first important observation on cost (15) is that the dependence of the transmission probabilities on the elapsed times n_k^i can be crucial to ensure the boundedness of cost (16). To see this, consider the case of a purely random TS wherein $\varphi^i(j) = \alpha^i$ for all $j = 0, 1, \ldots$ and for all $i = 1, \ldots, S$. Clearly, applying such a TS would be formally equivalent to introducing a lossy transmission channel with packet-drop rate α^i from each sensor i to the fusion node F. Then, as well known, in this case the MSE at node F may diverge if the transmission rates α^i do not exceed a certain critical value (the interested reader is referred to [23, 24] for the case of measurement transmission and to [6, 8, 9] for the case of estimate transmission).*

Nevertheless, it turns out that boundedness of (15) can be always ensured provided that an upper bound on the time between consecutive transmissions is imposed. Then, attention will be devoted to TS's satisfying the following assumption.

A2. The TS is such that for each sensor $i \in \{1, \dots, S\}$, one has $\varphi^i(N^i - 1) = 1$ for some finite $N^i > 1$.

This amounts to assuming that the time interval between two consecutive transmissions of sensor i never exceeds N^i. Clearly, assumption A2 implies that the vector of elapsed times $\mathbf{n}_k$ always takes value in the discrete finite set

$$\mathcal{N} \stackrel{\triangle}{=} \left\{ \mathbf{n} = \mathrm{col}(n^1, \dots, n^S) \in \mathbb{Z}_+^S : n^i \le N^i - 1\,,\ i = 1, \dots, S \right\} .$$

As it will be shown in the next section, TS's of this kind satisfying the communication rate constraint (3) always exist provided that $N^i \ge 1/\alpha^i$. It is also worth noting that such a class of strategies includes periodic strategies with transmission rate equal to $1/N^i$ (in this case $\varphi^i(j) = 0$ for $j = 0, \dots, N^i - 2$).

Remark 2. *Clearly, assumption A2 is not strictly necessary. For instance, a more general framework would amount to considering finite-memory strategies governed by finite-state Markov chains without requiring transmission after a bounded interval. However, it turns out that, in the considered setting, such a generalization can be avoided. In fact, as will be shown in the next sections, the optimal TS's do not depend on the upper bounds* N^i, $i = 1, \dots, S$ *(provided that they are large enough to satisfy the communication rate constraint) and so no advantages can be obtained in increasing (or removing) such limits.*

In order to analyze the asymptotic behavior of the fused covariance $\mathbf{P}_{k|k}$, first note that, for each sensor node i, we have

$$\lim_{k \to \infty} \mathbf{P}^i_{k|k-j} = \mathbf{P}^i(j)\,, \quad j = 0, \dots, N^i - 1$$

where the covariances $\mathbf{P}^i(j)$ can be obtained by repeatedly applying the Lyapunov difference equation:

$$\left\{ \begin{array}{lcl} \mathbf{P}^i(0) & = & \mathbf{P}^i_a \\ \mathbf{P}^i(j) & = & \mathbf{A}^i_1 \mathbf{P}^i(j-1)(\mathbf{A}^i_1)' + \mathbf{D}^i_1 \mathbf{Q} (\mathbf{D}^i_1)'\,, \quad j = 1, \dots, N^i - 1\,. \end{array} \right. \tag{17}$$

Then, it is immediate to see that, for both the OCI and ACI algorithms described in the previous section, the following properties hold.

P1. At each time $k = 0, 1, \dots$, the fused estimate $\hat{\mathbf{x}}_{k|k}$ is unbiased;

P2. At each time $k = 0, 1, \dots$ and for any given transmission pattern $\mathbf{c_{0:k}}$, the fused covariance $\mathbf{P}_{k|k}$ depends only on k and on the last transmission instant of each sensor, i.e.,

$$\mathbf{P}_{k|k} = \mathbf{P}_k(\mathbf{n}_k)\,.$$

Further, the following result can be stated.

Proposition 1. *Let either the OCI or the ACI algorithm be used at the fusion node. Then, under assumption A1, for any* $\mathbf{n} \in \mathcal{N}$*,* $\mathbf{P}_k(\mathbf{n})$ *converges exponentially to a bounded steady-state value* $\mathbf{P}(\mathbf{n})$*.*

Proof: see the Appendix. □

Turning back the attention to cost (15), let us denote by $p : \mathcal{N} \to [0, 1]$ the long-run average distribution of the sequence $\mathbf{n}_0, \mathbf{n}_1, \ldots$ defined as

$$p(\mathbf{n}) = \lim_{t\to\infty} \frac{1}{t} \sum_{k=1}^{t} \delta_{\mathbf{n}\mathbf{n}_k}\,, \quad \mathbf{n} \in \mathcal{N}$$

where $\delta_{\mathbf{n}\mathbf{n}_k}$ is the (multidimensional) Kronecker delta function. In view of assumption A2 and property P2, it is an easy matter to see that

$$J(\mathbf{c}_{0:\infty}) = \sum_{\mathbf{n}\in\mathcal{N}} p(\mathbf{n})\ \mathrm{tr}\left\{\mathbf{P}(\mathbf{n})\mathbf{W}\right\}\ . \tag{18}$$

Notice that, thanks to the finiteness of the set $\mathcal{N}$, such a cost turns out to be always bounded regardless of the transmission rates α^i.

5. Problem reformulation under ergodicity

In this section, a more convenient form for cost (15) will be derived. To this end, notice first that each sequence $n_0^i, n_1^i, \ldots$ can be described by a discrete-time Markov chain characterized by the state space $\mathcal{N}^i = \{0, 1, \ldots, N^i - 1\}$ and the transition matrix

$$\mathbf{\Phi^i} = \begin{bmatrix} \varphi^i(0) & \varphi^i(1) & \ldots & \varphi^i(N^i-2) & 1 \\ 1-\varphi^i(0) & 0 & \ldots & 0 & 0 \\ 0 & 1-\varphi^i(1) & \ldots & 0 & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \ldots & 1-\varphi^i(N^i-2) & 0 \end{bmatrix}.$$

For the reader's convenience, the associated state transition diagram is provided in Fig. 1. It is easy to check that such a Markov chain $n_0^i, n_1^i, \ldots$ is irreducible if and only if $\varphi^i(j) < 1$ for $j = 0, \ldots, N^i - 2$. In this case[3], there exists a unique invariant distribution $\pi^i(j)\,, \quad j = 0, \ldots, N^i - 1$ for the probabilities $\mathbb{P}\{n_k = j\}\,, \quad j = 0, \ldots, N^i - 1$. As well known, such invariant distribution $\boldsymbol{\pi}^i = \mathrm{col}\left(\pi^i(0) \quad \ldots \quad \pi^i(N^i - 1)]\right)$ corresponds to the Perron-Frobenius eigenvector of the transition matrix $\mathbf{\Phi}^i$ and can be computed by imposing the balance equations

$$\mathbf{\Phi}^i \boldsymbol{\pi}^i = \boldsymbol{\pi}^i \tag{19}$$

[3] Notice that, when $\varphi^i(j) = 1$ for some $j < N^i - 1$, the considered Markov chain is not irreducible. However, it is straightforward to see that the developments of this section still hold since, also in this case, there exists a unique closed communicating class corresponding to the set of states $\{0, 1, \ldots, M^i - 1\}$, where M^i is the smallest integer such that $\varphi(M^i) = 1$.

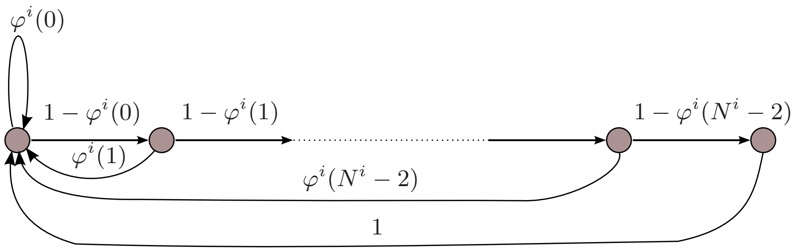

Figure 1: State transition diagram of the Markov chain $n_0^i, n_1^i, \ldots,$

and

$$(\boldsymbol{\pi}^i)'\mathbf{1} = 1 \tag{20}$$

where $\mathbf{1}$ is the column vector of suitable dimension with all entries equal to 1. With this respect, from (19) one can derive

$$\pi^i(j) = \pi^i(0) \prod_{h=0}^{j-1} (1 - \varphi^i(h)), \quad j = 0, \ldots, N^i - 1$$

where, for the sake of compactness, we define $\prod_{h=0}^{-1}(\cdot) \triangleq 1$. Since (20) can be rewritten as

$$\pi^i(0) \sum_{l=0}^{N^i-1} \prod_{h=0}^{l-1} (1 - \varphi^i(h)) = 1,$$

the invariant distribution is obtained as

$$\pi^i(j) = \left(\sum_{l=0}^{N^i-1} \prod_{h=0}^{l-1} (1 - \varphi^i(h)) \right)^{-1} \prod_{h=0}^{j-1} (1 - \varphi^i(h)), \qquad \text{for } j = 0, \ldots, N^i - 1. \tag{21}$$

Consider now the sequence $\mathbf{n}_0, \mathbf{n}_1, \ldots$. Clearly, supposing that the sensor TS's are independent, it can be described by a discrete-time Markov chain characterized by the state space $\mathcal{N}$ and the transition matrix $\boldsymbol{\Phi} = \bigotimes_{i=1}^{S} \boldsymbol{\Phi}^{\mathbf{i}}$ where $\otimes$ denotes the Kronecker product. Then, exploiting the mixed-product property (see, e.g, [25] p. 408), it is an easy matter to see that the unique invariant distribution of such a Markov chain is given by $\boldsymbol{\pi} = \bigotimes_{i=1}^{S} \boldsymbol{\pi}^{\boldsymbol{i}}$, that is, for any $\mathbf{n} = (n^1, \ldots, n^S)$, $\pi(\mathbf{n}) = \prod_{i=1}^{S} \pi^i(n^i)$.

As well known, two possibilities may arise:

(a) the Markov chain is aperiodic and ergodicity holds in that, for any state $\mathbf{n}$, the long-run average $p(\mathbf{n})$ converges to the unique invariant distribution $\pi(\mathbf{n})$; in this case cost (15) takes on the realization-independent form

$$J'(\mathbf{c}_{0:\infty}) = \sum_{\mathbf{n}\in\mathcal{N}} \pi(\mathbf{n}) \, \mathrm{tr}\,\{\mathbf{P}(\mathbf{n})\mathbf{W}\} \triangleq J'(\boldsymbol{\pi}); \tag{22}$$

(b) the Markov chain is periodic and the asymptotic behavior of the long-run average $p(\mathbf{n})$ depends on the initial conditions (in our case, the phase shift among the sensors).

In the context of our application, situation (b) does not correspond to a desired behavior since the fused covariance at node F would critically depend on the phase shift among the sensors and coordination would be needed in order to optimize performance. Then, we shall restrict our attention only to those choices of the transmission probabilites for which the Markov chain $n_0^i, n_1^i, \ldots$ is aperiodic for each sensor $i = 1, \ldots, S$.

Noting that, in this case,

$$\lim_{t\to\infty} \frac{1}{t} \sum_{k=1}^{t} c_k^i = \lim_{t\to\infty} \frac{1}{t} \sum_{k=1}^{t} \delta_{n_k^i 0} = \pi^i(0)$$

we are now in the position of reformulating Problem 1 as follows.

Problem 2. *Find the optimal transmission probabilities* $\varphi^i(j)$, $i = 1, \ldots, S$, $j = 0, \ldots, N^i - 2$ *such that*

(i) for each sensor $i = 1, \ldots, S$, *the Markov chain* $n_0^i, n_1^i, \ldots$ *is aperiodic; and*

(ii) the invariant distribution cost $J'(\boldsymbol{\pi})$ *is minimized under the constraints*

$$\begin{array}{ll} \pi^i(0) = \alpha^i & i = 1, \ldots, S, \\ 0 \le \varphi^i(j) \le 1\,, & j = 0, \ldots, N^i - 2, \\ & i = 1, \ldots, S\,. \end{array} \tag{23}$$

Note that the cost (22) depends on the invariant distribution probabilities $\pi^i(j)$ for $i \in \mathcal{S}$ and $j \in \mathcal{N}^i$ or, equivalently, on the transition probabilities $\varphi^i(j)$.

It must be pointed out that in general the cost (15) of Problem 1 is different from the cost (22) of Problem 2, but as discussed above the two costs coincide whenever the Markov chains turn out to be aperiodic (i.e., the condition in point (i) is satisfied). As will be shown in the next section, solution of the minimization problem in point (ii) may yield periodic Markov chains. In this case, the resulting periodic Markov chains should be somehow perturbed (in the respect of the transmission rate constraint) so as to ensure aperiodicity.

6. Optimal transmission strategy

In this section, the optimal solution of Problem 2 is derived and its properties are discussed. To this end, the following preliminary lemma is needed.

Lemma 1. *Suppose that assumptions A1-A2 hold and let either the OCI or the ACI algorithm be used at the fusion node. Then, for any* $\mathbf{W} > 0$, *one has*

$$\mathbf{n} \ge \bar{\mathbf{n}} \quad \Rightarrow \quad \mathrm{tr}\,\{\mathbf{P}(\mathbf{n})\mathbf{W}\} \ge \mathrm{tr}\,\{\mathbf{P}(\bar{\mathbf{n}})\mathbf{W}\} \tag{24}$$

where the inequality between vectors $\mathbf{n}, \bar{\mathbf{n}} \in \mathcal{N}$ *is taken componentwise.*

Proof: see the Appendix. □

In view of Lemma 1, the following result can be stated.

Theorem 1. *Suppose that assumptions A1-A2 hold and let either the OCI or the ACI algorithm be used at the fusion node. Then, the minimization in point (ii) of Problem 2 admits solution if and only if* $N^i \geq 1/\alpha^i$, $i = 1, \ldots, S$.

Further, for any $\mathbf{W} > 0$, *the optimal solution of such a minimization problem is given by*

$$\begin{array}{ll} \bar{\varphi}^i(j) = 0\,, & j = 0, \ldots, M^i - 3\,, \\ \bar{\varphi}(M^i - 2) = M^i - 1/\alpha^i\,, & \\ \bar{\varphi}(M^i - 1) = 1\,, & \end{array} \tag{25}$$

for $i = 1, \ldots, S$, *where* M^i *is the smallest integer such that* $M^i \geq 1/\alpha^i$.

Proof: see the Appendix. □

The following considerations about Theorem 1 can be made.

(a) When, for any sensor $i = 1, \ldots, S$, $1/\alpha^i$ is not integer, Theorem 1 yields an aperiodic TS. In fact, the transition graphs of the resulting Markov chains are aperiodic having cycles of both lengths $M^i - 1$ and M^i [26]. This means that the derived TS is consistent with the condition in point (i) of Problem 2 and consequently it represents also a solution of Problem 2. In this case, the resulting TS (25) minimizes the time-averaged squared error $J(\mathbf{c}_{0:\infty})$ not only in the worst-case sense (i.e., in the sense of cost (16)) but also in the sense of its average over independent realizations (i.e., $\mathbb{E}\{J(\mathbf{c}_{0:\infty})\}$).

(b) If instead, for some sensor, $1/\alpha^i$ is integer, then the optimal solution of the minimization in point (ii) Problem 2 turns out to be periodic. Unfortunately, this means that it does not represent a solution of Problem 2, not being consistent with the condition in point (i). Notice that in this case the equivalence between costs (16) and (22) does not hold, the latter being just a lower bound on the former.
However, it turns out that such a strategy can be made aperiodic by means of a suitable perturbation at the expense of a little increase in cost (22).

For instance, taking into account the structure of cost (22) as well as the monotonicity property of Lemma 1, one can argue that the best aperiodic perturbation corresponds to setting

$$\begin{array}{ll} \tilde{\varphi}^i(j) = 0\,, & j = 0, \ldots, M_i - 3\,, \\ \tilde{\varphi}^i(M^i - 2) = \varepsilon & \\ \tilde{\varphi}^i(M^i - 1) = (1 - 2\varepsilon)/(1 - \varepsilon)\,, & \\ \tilde{\varphi}^i(M^i) = 1\,, & \end{array} \tag{26}$$

where ε is a positive real such that $\varepsilon < 1/2$. Notice that (26) amounts to transmitting $100\,\varepsilon\%$ times with period $M_i - 1$, $100\,(1-2\,\varepsilon)\%$ times with period M_i, and the remaining $100\,\varepsilon\%$ times with period $M_i + 1$. It is immediate to see that such a perturbed TS is aperiodic and satisfies the rate constraint (3). Further, by simple continuity arguments, it turns out that the increase in cost (22) can be made arbitrarily small by choosing suitably small values of ε. As for optimality, the following result holds.

Theorem 2. *Let* $\bar{\boldsymbol{\pi}}^i$ *denote the invariant distribution in node* i *associated with the solution (25). Then, the TS (26) is the one with smallest cost (22), among all the TS's satisfying constraints (23) and for which*[4]

$$\|\boldsymbol{\pi}^i - \bar{\boldsymbol{\pi}}^i\|_1 \geq 2\,\varepsilon\,\alpha^i\,, \quad i = 1, \ldots, S\,. \tag{27}$$

[4]here $\|\cdot\|_1$ denotes the L^1 norm of a vector.

Proof: see the Appendix. □

Theorem 2 shows that the aperiodic perturbation (26) is the best choice among all aperiodic TS's with distance not smaller than $2\,\varepsilon\,\alpha^i$ (in the sense of inequality (27)) from the optimal periodic one.

7. Simulation results

In this section, a performance evaluation of the proposed centralized multisensor estimation framework is carried out. In particular, the goal is to compare different measurement-independent communication strategies between the sensor nodes and the central fusion node.

To this end, the centralized multisensor network is used to estimate the state of an object whose motion is described by the kinematic constant velocity model:

$$\mathbf{x}_{k+1} = \begin{bmatrix} 1 & T & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & T \\ 0 & 0 & 0 & 1 \end{bmatrix} \mathbf{x}_k \;+\; \mathbf{w}_k \,.$$

where $T = 1$ is the sampling interval and the unknown state vector is given by the position and speed components along the coordinate axes i.e. $\mathbf{x} = [p_x, v_x, p_y, v_y]'$. It is assumed that each sensor provides measurements of the object's position in cartesian coordinates, i.e.,

$$\mathbf{y}_k^i = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix} \mathbf{x}_k \;+\; \mathbf{v}_k^i \,.$$

The covariance matrices of measurement and process noises have been assumed equal to $\mathbf{R}^{i,j} = \text{diag}(1,1)\ \delta_{ij}$ and $\mathbf{Q} = \text{diag}(0.01, 0.001, 0.01, 0.001)$, respectively, where δ_{ij} is the Kronecker delta function. Transmission rates that are reciprocal of integers, i.e., $\alpha^i = \alpha$ with $1/\alpha \in \mathbb{Z}_+$, have been considered. For the sake of simplicity, the ACI algorithm has been implemented as data fusion rule.

The following TS's have been compared:

- **ATE:** Aperiodic Transmission of Estimates in (26) wherein the positive real ε has been chosen equal to 0.05;
- **RTE:** each sensor i transmits randomly to the fusion node according to the given transmission rate α, i.e., $\varphi^i(j) = \alpha$ for $j = 0, 1, \ldots$.
- **PTE:** sensors transmit periodically, once every $1/\alpha$ time instants, the phase shift among the sensors being random.

The comparison among these strategies has been carried out via Monte Carlo simulations with independent runs obtained for randomly generated trajectories by varying the measurement and process noises realizations. The *time averaged square error* (TASE) at the fusion node F

$$TASE = \frac{1}{K} \sum_{k=1}^{K} (\mathbf{x}_k - \hat{\mathbf{x}}_{k|k})'(\mathbf{x}_k - \hat{\mathbf{x}}_{k|k})$$

has been computed for each simulation run and for each TS. Then, the mean, the standard deviation and the maximum value of the TASE over the Monte Carlo runs have been considered as performance indices. The time horizon K and the number of Monte Carlo runs have been chosen equal to 10^4 and 10^3, respectively. Finally, the transmission rate α has been fixed to 0.1.

Fig. 2 (left column) shows the behavior of the mean TASE, the relative standard deviation and maximum value for the considered estimate transmission strategies as a function of the number of sensors. It can be noticed that ATE provides a notable improvement with respect to PTE in terms of worst-case performance (maximum TASE) and standard deviation, while ensuring an almost identical mean value. Further, ATE considerably outperforms also RTE especially whenever a few sensors are available. As discussed above, the poor behavior of the PTE strategy in terms of both standard deviation and maximum value of the TASE stems from the fact that, whenever the phase shift among the sensors is not well distributed (i.e., all sensors are almost in-phase), the estimation performance at the fusion node undergoes a notable degradation.

In order to check the robustness of the proposed strategies with respect to packet dropouts, further simulations have been carried out by considering a communication channel with 10% drop rate. The outcome of such simulations (reported in the right column of Fig. 2) shows that the performance degradation, due to the reduction in the number of successful transmissions, is quite similar for all the considered strategies and the same conclusions as in the non-lossy case can be drawn. This result is somehow expected because packet drop affects the various strategies in the same way.

Finally, we have also considered additional strategies in which each sensor sends the last collected measurement (instead of the state estimate) to the fusion node. In this case, the fusion algorithm is simply a Kalman filter that uses these measurements to update its state estimate. We denote by ATM, RTM and PTM the equivalent of the above described strategies in the case in which raw measurements are sent instead of estimates. As it can be seen from Fig. 3, also in this case, aperiodic strategies provide some advantages with respect to periodic ones. However, comparing ATE, PTE and RTE versus ATM, PTM and RTM, it is evident that sending estimates is in this case preferable to sending measurements, especially when only a few sensors are available.

8. Conclusions

The paper has addressed the state estimation problem assuming that: estimates are required at a distant location from the sensors connected via a communication link; a limitation on the communication rate is imposed; the sensor nodes have enough processing capability to compute locally the optimal estimates. Probabilistic measurement-independent strategies for deciding which data transmit have been investigated. In particular, considering covariance intersection as data fusion rule, it has been shown that a transmission strategy with improved performance can be generated through an ergodic Markov chain whereby an upper bound on the transmission dwell time (here defined as the time elapsed between consecutive transmissions) is guaranteed. Both the theoretical analysis and the simulation results have indicated that: thanks to its ergodicty, the derived strategy yields certain advantages in terms of worst-case time-averaged performance with respect to periodic ones when coordination among sensors is not possible; thanks to the existence of a maximum transmission dwell time, it compares favorably

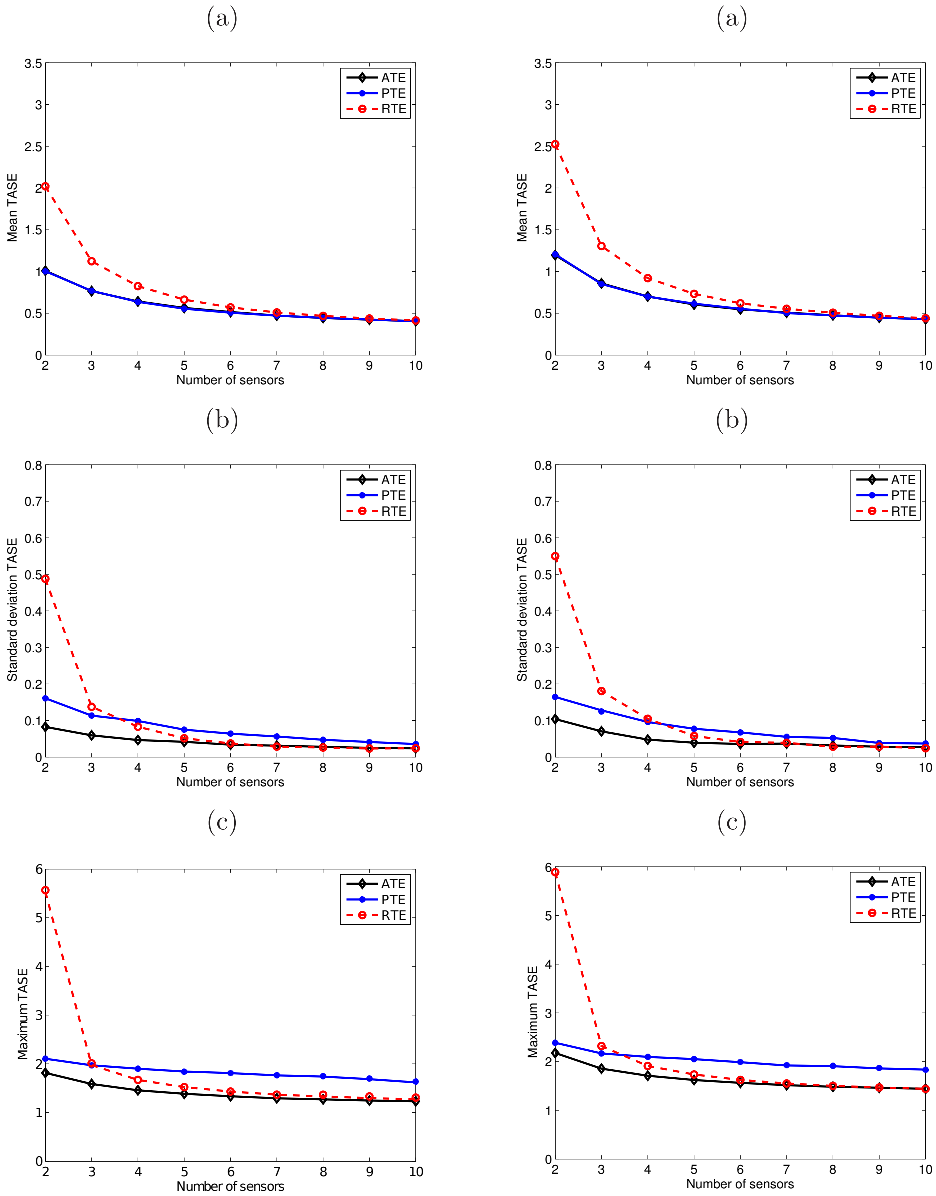

Figure 2: Mean TASE (a), standard deviation of TASE (b), maximum TASE (c) for the strategies sending estimates in the case of ideal transmission channel (left column) and lossy transmission channel (right column) with 10% packet drop rate.

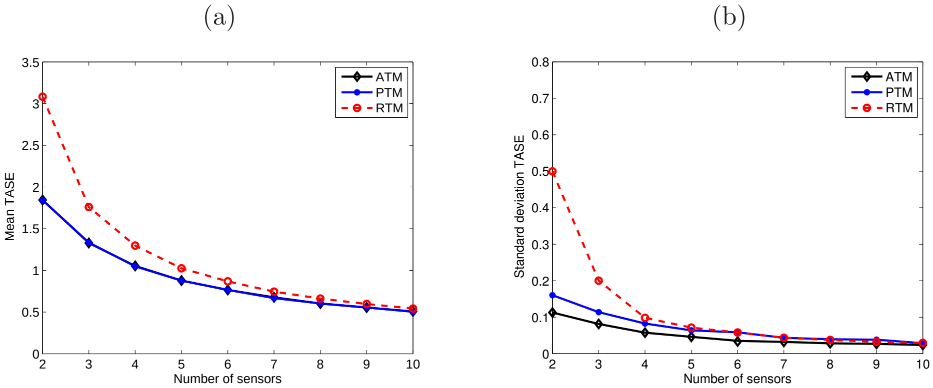

Figure 3: Mean TASE (a) and standard deviation of TASE (b) for the strategies sending measurements.

also to purely random strategies since it always ensures boundedness of the estimation error at the fusion node.

Appendix A.

Proof of Proposition 1

Consider first the ACI fusion rule. Clearly, for any $\mathbf{n} = \mathrm{col}(n^1, \dots, n^S)$, when the steady-state value $\mathbf{P}(\mathbf{n})$ (obtained from $\mathbf{P}^1(n^1), \dots, \mathbf{P}^S(n^S)$ by applying (10)) exists, it must satisfy the equation

$$\begin{aligned} (\mathbf{P}(\mathbf{n}))^{-1} &= (1/S) \sum_{i=1}^{S} (\mathbf{D}_1^i)' (\mathbf{P}^i(n^i))^{-1} \mathbf{D}_1^i \\ &= (1/S)\, \mathbf{D}'\, \mathbf{\Psi}(\mathbf{n})\, \mathbf{D} \end{aligned}$$

where $\mathbf{D} \triangleq \mathrm{col}\left(\mathbf{D}_1^1, \dots, \mathbf{D}_1^S\right)$ and $\mathbf{\Psi}(\mathbf{n}) \triangleq \mathrm{diag}\left((\mathbf{P}^1(n^1))^{-1}, \dots, (\mathbf{P}^S(n^S))^{-1}\right)$. Let now $\sigma_{\min}\{\cdot\}$ denote the smallest singular value of a matrix. Then, in view of the adopted local detectability decompositions, one has $\sigma_{\min}\left\{(\mathbf{P}^i(n^i))^{-1}\right\} > 0$ for any i and, consequently, $\sigma_{\min}\{\mathbf{\Psi}(\mathbf{n})\} > 0$. Further, under the collective detectability assumption A1, it is an easy matter to see that the matrix $\mathbf{D}$ is full-rank. This, in turn, implies that $\sigma_{\min}\{\mathbf{D}'\, \mathbf{\Psi}(\mathbf{n})\, \mathbf{D}\} > 0$. As a consequence, the steady-state covariance $\mathbf{P}(\mathbf{n})$ exists and can be obtained as

$$\mathbf{P}(\mathbf{n}) = \left((1/S)\, \mathbf{D}'\, \mathbf{\Psi}(\mathbf{n})\, \mathbf{D}\right)^{-1}.$$

Notice also that, by optimality, boundedness of the the steady-state covariance $\mathbf{P}(\mathbf{n})$ for the ACI implies also boundedness of the corresponding steady-state covariance for the OCI. Then, in order to conclude the proof, it is sufficient to note that the latter can be obtained by solving an optimization problem similar to (12) with $\mathbf{P}_{k|k}$ replaced by $\left(\sum_{i=1}^{S} \omega_i (\mathbf{D}_1^i)' (\mathbf{P}^i(n^i))^{-1} \mathbf{D}_1^i\right)^{-1}$.

Proof of Lemma 1

First of all, notice that for each sensor i one has

$$\mathbf{P}^i(0) \leq \mathbf{P}^i(1) \leq \ldots \leq \mathbf{P}^i(N^i - 1)\,. \tag{A.1}$$

In fact, equations (7)-(8) and (17) imply $\mathbf{P}^i(1) = \mathbf{P}^i_b$. Thus, from (8), it follows that $\mathbf{P}^i(0) \leq \mathbf{P}^i(1)$. Suppose now that, for a given j, we have $\mathbf{P}^i(j-1) \leq \mathbf{P}^i(j)$. Then

$$\begin{aligned}
&\mathbf{P}^i(j+1) - \mathbf{P}^i(j) \\
&\quad = \mathbf{A}^i_1 \mathbf{P}^i(j)(\mathbf{A}^i_1)' + \mathbf{D}^i_1 \mathbf{Q}(\mathbf{D}^i_1)' - \mathbf{A}^i_1 \mathbf{P}^i(j-1)(\mathbf{A}^i_1)' - \mathbf{D}^i_1 \mathbf{Q}(\mathbf{D}^i_1)' \\
&\quad = \mathbf{A}^i_1(\mathbf{P}^i(j) - \mathbf{P}^i(j-1))(\mathbf{A}^i_1)' \geq 0\,,
\end{aligned}$$

so that, by induction, (A.1) holds.

Consider now two vectors $\mathbf{n} = (n^1, \ldots, n^S)$ and $\bar{\mathbf{n}} = (\bar{n}^1, \ldots, \bar{n}^S)$ and let

$$\mathbf{P}(\mathbf{n}) = \left(\sum_{i=1}^{S} \omega^i \, (\mathbf{D}^i_1)' \left(\mathbf{P}^i(n^i)\right)^{-1} \mathbf{D}^i_1 \right)^{-1}$$

and

$$\mathbf{P}(\bar{\mathbf{n}}) = \left(\sum_{i=1}^{S} \bar{\omega}^i \, (\mathbf{D}^i_1)' \left(\mathbf{P}^i(\bar{n}^i)\right)^{-1} \mathbf{D}^i_1 \right)^{-1}$$

be the corresponding steady-state fused covariances.

Focusing first on the OCI, by optimality of the weights $\bar{\omega}^i$, $i \in \mathcal{S}$, one has

$$\begin{aligned}
&\mathrm{tr}\left\{ \left(\sum_{i=1}^{S} \bar{\omega}^i \, (\mathbf{D}^i_1)' \left(\mathbf{P}^i(\bar{n}^i)\right)^{-1} \mathbf{D}^i_1 \right)^{-1} \mathbf{W} \right\} \\
&\qquad \leq \mathrm{tr}\left\{ \left(\sum_{i=1}^{S} \omega^i \, (\mathbf{D}^i_1)' \left(\mathbf{P}^i(\bar{n}^i)\right)^{-1} \mathbf{D}^i_1 \right)^{-1} \mathbf{W} \right\}.
\end{aligned} \tag{A.2}$$

Further, by virtue of (A.1), $\bar{\mathbf{n}} \leq \mathbf{n}$ implies $\mathbf{P}^i(\bar{n}^i) \leq \mathbf{P}^i(n^i)$, $i \in \mathcal{S}$, and consequently

$$\left(\sum_{i=1}^{S} \omega^i \, (\mathbf{D}^i_1)' \left(\mathbf{P}^i(\bar{n}^i)\right)^{-1} \mathbf{D}^i_1 \right)^{-1} \leq \left(\sum_{i=1}^{S} \omega^i \, (\mathbf{D}^i_1)' \left(\mathbf{P}^i(n^i)\right)^{-1} \mathbf{D}^i_1 \right)^{-1},$$

which, in turn, yields

$$\begin{aligned}
&\mathrm{tr}\left\{ \left(\sum_{i=1}^{S} \omega^i \, (\mathbf{D}^i_1)' \left(\mathbf{P}^i(\bar{n}^i)\right)^{-1} \mathbf{D}^i_1 \right)^{-1} \mathbf{W} \right\} \\
&\qquad \leq \mathrm{tr}\left\{ \left(\sum_{i=1}^{S} \omega^i \, (\mathbf{D}^i_1)' \left(\mathbf{P}^i(n^i)\right)^{-1} \mathbf{D}^i_1 \right)^{-1} \mathbf{W} \right\}.
\end{aligned} \tag{A.3}$$

Then, (24) follows at once by combining (A.2) with (A.3).

As to the ACI, since, in this case, $\omega^i = \bar{\omega}^i = 1/S$ for any $i \in \mathcal{S}$ it is immediate to see that (24) holds in view of (A.1).

Proof of Theorem 1

In order to address the solution of Problem 2, it is convenient to recast it in terms of the invariant distribution $\pi^i(j)$, $j \in \mathcal{N}^i$, $i \in \mathcal{S}$. To this end, note that if one refers to (21), the probabilities $\pi^i(j)$ can also be obtained recursively as

$$\pi^i(j) = \pi^i(j-1)\,(1 - \varphi^i(j-1))\,, \quad j = 1, \ldots, N^i - 1\,. \tag{A.4}$$

Then the feasibility constraints $0 \le \varphi^i(j) \le 1\,,\ j = 0, \ldots, N^i - 2\,,$ turn out to be equivalent to

$$\begin{array}{ll} \pi^i(j) \ge 0\,, & j = 1, \ldots, N^i - 1\,, \\ \pi^i(j+1) \le \pi^i(j)\,, & j = 0, \ldots, N^i - 2\,. \end{array} \tag{A.5}$$

Further, notice that $\varphi^i(N^i - 1) = 1$ implies that the probabilities $\pi^i(j)$, $j = 0, \ldots, N^i - 1$, sum up to one, i.e.,

$$\sum_{j=0}^{N^i-1} \pi^i(j) \;=\; 1\,. \tag{A.6}$$

Consider now a generic sensor i and let the transmission strategies of the other $S-1$ sensors be fixed. Then, cost (22) can be rewritten as $J'(\boldsymbol{\pi}^i) = \sum_{j \in \mathcal{N}^i} \pi^i(j)\, w^i(j)$ where

$$w^i(j) = \sum_{(n^1, n^2, \ldots, n^S) \in \mathcal{N},\, n^i = j} \operatorname{tr}\left\{\mathbf{P}(n^1, n^2, \ldots, n^S)\mathbf{W}\right\} \prod_{\ell \in \mathcal{S} \setminus \{i\}} \pi^\ell(n^\ell)\,.$$

In view of Lemma 1, it is immediate to see that, regardless of how the transmission strategies of the other $S-1$ sensors are chosen, the weights $w^i(j)$ satisfy the monotonicity condition

$$w^i(0) \le w^i(1) \le \ldots \le w^i(N^i - 1)\,. \tag{A.7}$$

Since $\pi^i(j) \le \alpha^i$ for $j = 0, \ldots, N^i - 1$ (see (23) and (A.5)), then one can have (A.6) only if $N^i \ge 1/\alpha^i$. Conversely, when $N^i \ge 1/\alpha^i$, it is immediate to see that the choice

$$\begin{array}{ll} \pi^i(j) = \alpha^i\,, & j = 0, \ldots, M^i - 2\,, \\ \pi^i(M^i - 1) = 1 - (M^i - 1)\,\alpha^i\,, & \\ \pi^i(j) = 0\,, & j = M^i, \ldots, N^i - 1\,, \end{array} \tag{A.8}$$

where M^i is the smallest integer such that $M^i \ge 1/\alpha^i$ satisfies all the constraints . Such a choice is also optimal because the weights $w^i(j)$ are monotonically nondecreasing. Notice that this represents always the optimal TS for sensor i regardless of how the transmission strategies of the other $S-1$ sensors are chosen. Then, in order to conclude the proof, it is sufficient to note that (A.8) is equivalent to (25) by virtue of relationships (A.4).

Proof of Theorem 2

For the sake of simplicity, we assume that $1/\alpha^i$ is integer for all sensors (if this condition does not hold, the proof can be given along similar lines). In this case, as shown in the proof of Theorem 1, the invariant distribution $\bar{\boldsymbol{\pi}}^i$ for a generic sensor i turns out to be $\bar{\pi}^i(j) = \alpha^i$ for $i = 0, \ldots, M^i - 1$ and 0 otherwise.

Then a generic perturbed strategy can be written as (recall that $\pi^i(0)$ must always be equal to α^i)

$$\begin{aligned}
&\pi^i(j) = \alpha^i - \varepsilon^i(j)\,, && j = 1, \dots, M^i - 1\,, \\
&\pi^i(j) = \varepsilon^i(j)\,, && j \geq M^i\,,
\end{aligned} \tag{A.9}$$

where $\varepsilon^i(1), \dots, \varepsilon^i(N^i - 1)$ are non-negative reals. For such a perturbed strategy, the L^1 distance from the optimal periodic one is given by

$$\|\boldsymbol{\pi}^i - \bar{\boldsymbol{\pi}}^i\|_1 = \sum_{j=1}^{N^i-1} \varepsilon^i(j)$$

Thus, coherently with (27), we suppose $\sum_{j=1}^{N^i-1} \varepsilon^i(j) \geq 2\,\varepsilon\,\alpha^i$. Further, in order to satisfy the constraint (A.6), it must be $\sum_{j=1}^{M^i-1} \varepsilon^i(j) - \sum_{j=M^i}^{N^i-1} \varepsilon^i(j) = 0$ which, in turn, implies

$$\sum_{j=1}^{M^i-1} \varepsilon^i(j) = \sum_{j=M^i}^{N^i-1} \varepsilon^i(j) \geq \varepsilon\,\alpha^i. \tag{A.10}$$

Let now the transmission strategies of the other $S-1$ sensors be fixed. Then, the cost associated with a generic perturbed strategy can be written as

$$J' = \bar{J}' - \sum_{j=1}^{M^i-1} w^i(j)\varepsilon^i(j) + \sum_{j=M^i}^{N^i-1} w^i(j)\varepsilon^i(j)$$

where $\bar{J}'$ is the cost associated with the optimal periodic strategy (25). Like in the proof of Theorem 1, the weights $w^i(j)$ satisfy the monotonicity condition (A.7), regardless of how the transmission strategies of the other $S-1$ sensors be chosen.

Consider now that, for the aperiodic perturbation (26), we have $\varepsilon^i(j) = \varepsilon\alpha^i$ for $j = M^i - 1, M^i$ and $\varepsilon^i(j) = 0$ otherwise. As a consequence, the difference between the cost $\tilde{J}'$ of such a strategy and the cost J' of a generic perturbed strategy can be readily obtained as

$$\tilde{J}' - J' = [w^i(M^i) - w^i(M^i - 1)]\,\varepsilon\alpha^i + \sum_{j=1}^{M^i-1} w^i(j)\varepsilon^i(j) - \sum_{j=M^i}^{N^i-1} w^i(j)\varepsilon^i(j)\,.$$

Therefore, the proof can be concluded by showing that, under the distance constraint (27), we have $\tilde{J}' - J' \leq 0$ for any possible perturbed strategy. To see this, we can exploit the monotonicity of the weights $w^i(j)$ and write

$$\begin{aligned}
\tilde{J}' - J' &\leq [w^i(M^i) - w^i(M^i - 1)]\,\varepsilon\alpha^i + w^i(M^i - 1) \sum_{j=1}^{M^i-1} \varepsilon^i(j) - w^i(M^i) \sum_{j=M^i}^{N^i-1} \varepsilon^i(j) \\
&= [w^i(M^i) - w^i(M^i - 1)] \left(\varepsilon\alpha^i - \sum_{j=M^i}^{N^i-1} \varepsilon^i(j) \right)
\end{aligned}$$

where the latter equality follows from (A.10). Since (A.10) implies also that $\varepsilon\alpha^i - \sum_{j=M^i}^{N^i-1} \varepsilon^i(j) \leq 0$, non-positiveness of $\tilde{J}' - J'$ follows at once.

References

[1] L. M. Feeney, M. Nilsson, Investigating the energy consumption of a wireless network interface in an ad hoc networking environment, in: Proc. IEEE Infocom, 2001, pp. 1548–1557.

[2] J. K. Yook, D. M. Tilbury, N. R. Soparkar, Trading computation for bandwidth: Reducing communication in distributed control systems using state estimators, IEEE Transactions on Control Systems Technology 10 (2002) 503–518.

[3] X. Li, W. S. Wong, State estimation with communication constraints, Systems & Control Letters 28 (1996) 49–54.

[4] W. S. Wong, R. W. Brockett, Systems with finite communication bandwidth constraints - Part I: state estimation problem, IEEE Transactions on Automatic Control 42 (1997) 1294–1299.

[5] S. Tatikonda, S. Mitter, Control unde communication constraints, IEEE Transactions on Automatic Control 49 (7) (2004) 1056–1068.

[6] Y. Xu, J. P. Hespanha, Estimation under uncontrolled and controlled communications in networked control systems, in: Proc. of the 44th IEEE Conference on Decision and Control, 2005, pp. 842–847.

[7] J. P. Hespanha, P. Naghshtabrizi, Y. Xu, A survey of recent results in networked control systems, Proc. of IEEE Special Issue on Technology of Networked Control Systems 95 (2007) 138–162.

[8] V. Gupta, B. Hassibi, M. Murray, Optimal LQG control across packet-dropping links, Systems & Control Letters 56 (2007) 439–446.

[9] G. Battistelli, A. Benavoli, L. Chisci, State estimation with a remote sensor under limited communication rate, in: Proc. of the 3rd International Symposium on Communications, Control and Signal Processing, 2008, pp. 654–659.

[10] G. Battistelli, A. Benavoli, L. Chisci, State estimation with remote sensors under limited communication rate, in: A. Chiuso, L. Fortuna, M. Frasca, A. Rizzo, L. Schenato, S. Zampieri (Eds.), Modelling, Estimation and Control of Networked Complex Systems, Springer, 2009.

[11] L. Shi, P. Cheng, J. Chen, Sensor data scheduling for optimal state estimation with communication energy constraint, Automatica, to appear.

[12] J. S. Baras, A. Bensoussan, Optimal sensor querying: General Markovian and LQG models with controlled observations, SIAM Journal on Control and Optimization 27 (1989) 786–813.

[13] V. Gupta, T. H. Chung, B. Hassibi, R. M. Murray, On a stochastic sensor selection algorithm with applications in sensor scheduling and sensor coverage, Automatica 42 (2006) 251–260.

[14] W. Wu, A. Araposthatis, Optimal sensor querying: general Markovian and LQG models with controlled observations, IEEE Transactions on Automatic Control 53 (6) (2008) 1392–1405.

[15] L. A. Montestruque, P. Antsaklis, Stability of model-based networked control systems with time-varying transmission times, IEEE Transactions on Automatic Control 49 (2004) 1562–1572.

[16] I. Bilinskis, A. Mikelsons, Randomized Signal Processing, Prentice Hall International, 1992.

[17] V. Paxson, End-to-end routing behavior in the internet, ACM SIGCOMM Computer Communication Review 26 (1996) 25–38.

[18] S. Julier, J. Uhlmann, A non-divergent estimation algorithm in the presence of unknown correlations, in: Proceedings of the 1997 American Control Conference, Vol. 4, 1997, pp. 2369–2373.

[19] L. Chen, P. Arambel, R. Mehra, Estimation under unknown correlation: covariance intersection revisited, IEEE Transactions on Automatic Control 47 (11) (2002) 1879–1882.

[20] X. R. Li, Y. M. Zhu, J. Wang, C. Z. Han, Optimal linear estimation fusion-part I: unified fusion rules, IEEE Transactions on Information Theory 49 (2003) 2192–2208.

[21] N. G. Wah, Y. Rong, Comparison of decentralized tracking algorithms, in: Proceedings of the Sixth International Conference of Information Fusion, Vol. 1, 2003, pp. 107–113.

[22] A. Jazwinski, Stochastic Processes and Filtering Theory, Academic Press, 1970.

[23] B. Sinopoli, L. Schenato, M. Franceschetti, K. Poolla, M. Jordan, S. Sastry, Kalman filtering with intermittent observations, IEEE Transactions on Automatic Control 49 (2004) 1453–1464.

[24] X. Liu, A. Goldsmith, Kalman filtering with partial observation losses, in: Proceedings of the IEEE Conference on Decision and Control, Vol. 4, 2004, pp. 4180–4186.

[25] P. Lancaster, M. Tismenetsky, The Theory of Matrices, Academic Press, 1985.

[26] J. P. Jarvis, D. R. Shier, Graph-theoretic analysis of finite Markov chains, in: D. R. Shier, K. T. Wallenius (Eds.), Applied Mathematical Modeling: A Multidisciplinary Approach, CRC Press, 1996.