2017 IEEE 56th Annual Conference on Decision and Control (CDC)
December 12-15, 2017, Melbourne, Australia

Optimal sensor scheduling strategies in networked estimation

Marcos M. Vasconcelos, Ashutosh Nayyar and Urbashi Mitra

***Abstract*— Consider a remote sensing system where multiple pairs of stochastic subsystems and corresponding non-collocated estimators share a common wireless network. In order to avoid packet collisions, the access to the network is mediated by a manager that chooses, based on the realizations of the states, which one (if any) of the states will be transmitted over the network. Each estimator is interested solely on the state of its corresponding subsystem. This paper studies the jointly optimal design of scheduling and estimation strategies for a cost functional consisting of a mean-squared error and a communication cost in a one-shot problem formulation. The global optimality of a scheduling and estimation strategy pair for the unicast network model is established. A person-by-person optimal pair of strategies is obtained for a broadcast network model.**

I. INTRODUCTION

Reliable real-time wireless networking is an important requirement of modern cyberphysical systems [1]. Due to their large scale, these systems are typically formed by multiple physically distributed subsystems, that communicate over wireless networks of limited capacity. In addition to degrading the performance of overall system, the fact that the communication among the different agents in cyber-physical systems is imperfect often leads dynamic team-decision problems with nonclassical information structures. Such problems are usually nonconvex, and are in general difficult to solve [2].

We consider a class of *one-shot* remote estimation problems involving a large scale distributed system. The system consists of multiple subsystems paired with their corresponding estimators, communicating over a shared network as in Fig. 1. The states of the multiple subsystems are observed by a *network manager* that chooses a single observation to be transmitted in a packet over the wireless network. The role of the network manager is to implement a strategy that eliminates simultaneous transmissions that otherwise would result in packet collisions as in [3].

We assume two different models for the wireless network: the *unicast* and the *broadcast* models. Each model induces a distinct information structure in the team-decision problem. In the unicast network, a single estimator receives its intended packet and the remaining estimators receive "no-transmission" symbols; In the *broadcast* network, all of the estimators receive the same transmitted packet. Our goal is to find scheduling and estimation strategies that jointly minimize an objective cost functional consisting of a mean-squared error criterion and a communication cost. We note that both cases correspond to team-decision problems with nonclassical information structures. Obtaining globally optimal solutions for these problems is a challenging task, in general.

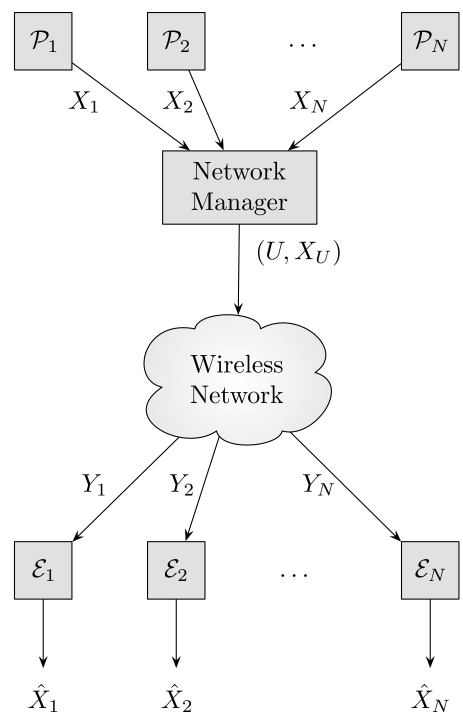

Fig. 1. Block diagram of the observation-driven sensor scheduling problem.

The problem we address in this paper is motivated by applications in networked control and estimation where there is a hard constraint on the number of packets that can be transmitted over the network. Due to the presence of interference in wireless networks, we are interested in systems where some form of collision avoidance algorithm is implemented. Such mechanisms are broadly classified into two categories: contention-free and contention-based medium access control (MAC) protocols [4]. Contention-free MAC protocols are characterized by the fact that each node either receives a *fixed* portion of the communication resources such as in time-division or frequency-division multiple access; or the resource is allocated *dynamically* such as in token and polling-based multiple access. Contention-based protocols are characterized by the use of random access and collision resolution mechanisms such as in the ALOHA and CSMA

This research has been funded in part by the following grants: ONR N00014-15-1-2550, NSF CNS-1213128, NSF CCF-1410009, AFOSR FA9550-12-1-0215, NSF CPS-1446901. M. M. Vasconcelos, A. Nayyar and U. Mitra are with the Ming Hsieh Department of Electrical Engineering, University of Southern California, Los Angeles, CA 90089 USA. E-mails: {mvasconc,ashutoshn,ubli}@usc.edu.

978-1-5090-2872-6/17/$31.00 ©2017 IEEE

protocols.

There is a growing body of work related to the class of systems considered here. One of the first references on scheduling and its impact on the stability of networked control systems is [5]. The performance of event-triggered control loops over a shared network for different scheduling protocols was studied in [6]. The idea of introducing a network manager to schedule the access to the network by multiple control loops was proposed in [7] as well as in [8]. The impact of contention-based MAC schemes in networked estimation was considered in [9]. The issue of scheduling in networked estimation systems is considered by [10], where the performances of static, randomized and dynamic scheduling strategies were compared using a long-run average cost criterion. Our problem formulation distinguishes itself from the aforementioned works by jointly designing the scheduling and estimation strategies in a one-shot problem formulation. Solving the one-shot problem is a fundamental step in solving more general sequential estimation problems such as in [11] and [12].

The main contribution of this paper is to establish two optimality results of a scheduling strategy with the following structure: the norms of the state vectors are compared to a threshold; if at least one of the norms exceeds this threshold, the vector with the largest norm is transmitted. The optimal estimation strategy associated with this choice of scheduling strategy outputs the mean of the observation when a no-transmission symbol is received, or the received vector, otherwise. For the unicast network, we show that this strategy is optimal despite the fact that the optimization problem is nonconvex. For the broadcast network, we show that this strategy is *person-by-person* optimal. Our results hold for mutually independent vectors, with unimodal circularly symmetric densities. These results provide insight and reinforce the structure of strategies used in the architecture of some networked control and estimation systems such as [10].

The remainder of the paper is organized as follows. The problem setup, preliminary definitions and a brief discussion on signaling are presented in Section II. Our two main optimality results are stated in Section III. The proofs of our main results are provided in Section IV. The paper ends with our conclusions and future research directions in Section V.

II. System model and problem formulation

Consider a stochastic system with $N \geq 2$ independent subsystems $\mathcal{P}_1, \cdots, \mathcal{P}_N$ and corresponding non-collocated estimators $\mathcal{E}_1, \cdots, \mathcal{E}_N$. Let the state of subsystem $\mathcal{P}_i$ be denoted by $X_i \in \mathbb{R}^{n_i}$ where

$$X_i \sim f_{X_i}, \quad i \in \{1, \cdots, N\}. \tag{1}$$

Each estimator $\mathcal{E}_i$ is interested in forming estimates of the state of its corresponding subsystem $\mathcal{P}_i$, $i \in \{1, \cdots, N\}$. The subsystems communicate with their respective estimators over a synchronous wireless network that can support a single packet per time slot. In order to avoid collisions, the access to the shared network is scheduled by a *network manager*, whose role is to decide which one of the subsystems will transmit a packet containing the measured state X_i and its index over the network.

Let the vector X denote the concatenation of $X_1, \cdots, X_N$ such that $X \in \mathbb{R}^n$ with $n = n_1 + \cdots + n_N$. The vector X is distributed according to f_X. The *scheduling strategy* implemented by the network manager is a map

$$\gamma : \mathbb{R}^n \to \{0, 1, \cdots, N\} \tag{2}$$

and a decision variable U is computed according to

$$U = \gamma(X). \tag{3}$$

The decision variable U represents the index of the subsystem chosen by the network manager to transmit its state over the network. If $U \in \{1, \cdots, N\}$, the network input consists of a packet containing the chosen index U and the state X_U. If $U = 0$, the network input is the "no-transmission" symbol denoted by $\varnothing$. The space of admissible scheduling strategies is denoted by Γ.

Remark 1: A communication packet is assumed to contain enough bits to represent a real vector x_i with negligible quantization error. However, in practice, a packet has a finite number of bits. Therefore, the network manager is ***not allowed*** to encapsulate the state of the entire system in a single packet and broadcast it over the network.

A. Information structures

The signals observed at the estimators are determined by the decision variable U and the wireless network. The structure of the network induces an *information structure* in our model. Let the network output signal observed by estimator $\mathcal{E}_i$ be denoted by Y_i, $i \in \{1, \cdots, N\}$. In this paper, we will focus on two network models.

1) Unicast information structure ($\mathcal{I}_1$): Information structure $\mathcal{I}_1$ corresponds to the case where a single transmitter-receiver pair is selected by the network manager at a time. The communication between the chosen pair is perfect, and the remaining pairs remain idle. The network outputs are specified as follows:

$$Y_i = \begin{cases} (U, X_U), & \text{if } \ U = i \\ \varnothing, & \text{otherwise} \end{cases}, \quad i \in \{1, \cdots, N\}. \tag{4}$$

Based on its observation, estimator $\mathcal{E}_i$ forms an estimate of $\hat{X}_i$ of the state of subsystem $\mathcal{P}_i$, according to a function ψ_i. An estimation strategy for estimator $\mathcal{E}_i$ is a map

$$\psi_i : \{\varnothing\} \cup \{i \times \mathbb{R}^{n_i}\} \to \mathbb{R}^{n_i}, \quad i \in \{1, \cdots, N\}. \tag{5}$$

2) Broadcast information structure ($\mathcal{I}_2$): Information structure $\mathcal{I}_2$ corresponds to the case where the network manager broadcasts its choice to all the receivers connected to the network, i.e., transmits the same information bearing signal over the wireless medium. The network outputs are specified as follows:

$$Y_i = \begin{cases} (U, X_U), & \text{if } \ U \neq 0 \\ \varnothing, & \text{otherwise} \end{cases}, \quad i \in \{1, \cdots, N\}. \tag{6}$$

Based on its observation, estimator $\mathcal{E}_i$ forms an estimate $\hat{X}_i$ of the state of subsystem $\mathcal{P}_i$, according to a function ψ_i. An estimation strategy for estimator $\mathcal{E}_i$ is a map

$$\psi_i : \{\varnothing\} \cup \bigcup_{j=1}^{N} \{j \times \mathbb{R}^{n_j}\} \to \mathbb{R}^{n_i}, \quad i \in \{1, \cdots, N\}. \tag{7}$$

The estimates $\hat{X}_1, \cdots, \hat{X}_N$ are computed according to

$$\hat{X}_i = \psi_i(Y_i), \quad i \in \{1, \cdots, N\}. \tag{8}$$

From now on, we will refer to an *estimation strategy profile* ψ as the vector of estimation strategies defined as

$$\psi \stackrel{\text{def}}{=} (\psi_1, \cdots, \psi_N). \tag{9}$$

The space of admissible estimation strategy profiles is denoted by Ψ.

B. Cost functional

We are interested in obtaining a pair of scheduling and estimation strategies $(\gamma, \psi) \in \Gamma \times \Psi$ such that the aggregate mean-squared error in estimating the random vectors X_i plus a communication cost for transmitting over the network is minimized. This corresponds to the cost functional $\gamma : \Gamma \times \Psi \to \mathbb{R}$ such that

$$\mathcal{J}(\gamma, \psi) \stackrel{\text{def}}{=} \mathbf{E}\left[\sum_{i=1}^{N} \|X_i - \hat{X}_i\|^2 + c \cdot \mathbf{1}(U \neq 0)\right], \tag{10}$$

where $c \geq 0$. The general problem considered in this paper is stated below.

Problem 1: Given the joint probability density function of X, the constant $c \geq 0$, and a fixed information structure $\mathcal{I}_1$ or $\mathcal{I}_2$; find a pair of strategies $(\gamma, \psi) \in \Gamma \times \Psi$ that minimizes $\mathcal{J}(\gamma, \psi)$ in Eq. (10).

C. Notions of optimality

Throughout the paper we will refer to two notions of optimality commonly used in team-decision theory.

1) Global optimal solutions: A pair of scheduling and estimation strategies $(\gamma^*, \psi^*) \in \Gamma \times \Psi$ is *globally optimal* if

$$\mathcal{J}(\gamma^*, \psi^*) \leq \mathcal{J}(\gamma, \psi), \quad (\gamma, \psi) \in \Gamma \times \Psi. \tag{11}$$

2) Person-by-person optimal solutions: A pair of scheduling and estimation strategies $(\gamma^*, \psi^*) \in \Gamma \times \Psi$ is *person-by-person optimal* if

$$\begin{aligned} \mathcal{J}(\gamma^*, \psi^*) &\leq \mathcal{J}(\gamma^*, \psi), \quad \psi \in \Psi; \\ \mathcal{J}(\gamma^*, \psi^*) &\leq \mathcal{J}(\gamma, \psi^*), \quad \gamma \in \Gamma. \end{aligned} \tag{12}$$

Person-by-person optimality is a necessary but not sufficient condition for global optimality, and therefore it is a weaker notion. However, it is a useful concept that can be used to characterize properties of globally optimal solutions.

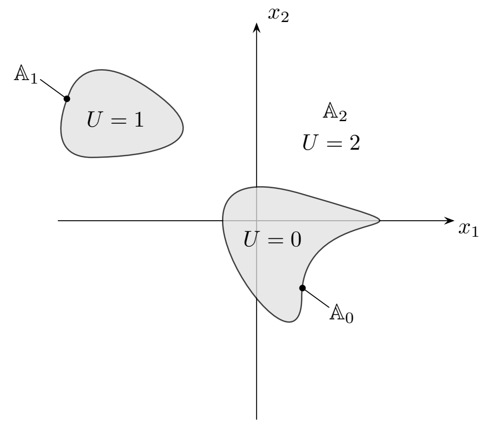

Fig. 2. Partition of the observation space used to illustrate the issue of signaling in problems of networked estimation.

D. A brief discussion on signaling

In problems of decentralized control and estimation with nonclassical information structures, optimal solutions usually use a form of implicit communication known as *signaling*. Signaling is the effect of conveying information through actions [13], and it is the reason why this class of problems are difficult to solve, cf. [14].

Consider an instance of Problem 1 with two zero-mean independent scalar subsystems and information structure $\mathcal{I}_1$. Assume that the network manager makes its decision according to the partition of the observation space shown in Fig. 2, where $\mathbb{A}_0 \cup \mathbb{A}_1 \cup \mathbb{A}_2 = \mathbb{R}^2$ and $(x_1, x_2) \in \mathbb{A}_i$ implies that $U = i$, $i \in \{0, 1, 2\}$. Suppose that the network manager observes $(x_1, x_2) \in \mathbb{A}_1$, which implies that $U = 1$ and consequently, $Y_1 = (1, x_1)$ and $Y_2 = \varnothing$. The optimal estimate used by $\mathcal{E}_2$ in this case is

$$\begin{aligned} \hat{X}_2 &= \mathbf{E}[X_2 \mid Y_2 = \varnothing]; \\ &= \mathbf{E}[X_2 \mid (X_1, X_2) \in \mathbb{A}_2^c], \end{aligned} \tag{13}$$

which may correspond to a different numerical value than if we simply used the naïve blind estimate $\mathbf{E}[X_2] = 0$.

If the problem has information structure $\mathcal{I}_2$, the observations at the estimators are $Y_1 = Y_2 = (1, x_1)$. This situation is more complicated than what we had previously: the optimal estimate used by $\mathcal{E}_2$ in this case is

$$\begin{aligned} \hat{X}_2 &= \mathbf{E}[X_2 \mid Y_2 = (1, x_1)]; \\ &= \mathbf{E}[X_2 \mid (X_1, X_2) \in \mathbb{A}_1, X_1 = x_1], \end{aligned} \tag{14}$$

which is a function of the scheduling strategy and x_1. Therefore, the optimal estimation strategy is always a function of the scheduling strategy, even if the subsystems are independent. This coupling between scheduling and estimation strategies is what makes these problems nontrivial even in one-shot scenarios. In the following sections, we elaborate on these issues and take a few steps toward obtaining optimal solutions to Problem 1.

III. Main results

In Section II-D we argued that the naïve estimate may not be optimal when "no-transmission" symbols are observed, for *arbitrary scheduling strategies*. However, under a few assumptions on the probabilistic model of the system, they turn out to be optimal when paired with an *appropriate scheduling strategy*. The following scheduling and estimation strategies play an important role in the two optimality results we present next.

Definition 1 (Max-norm scheduling strategy): Let $c \geq 0$ and $x \in \mathbb{R}^n$. The *max-norm* scheduling strategy is defined as

$$\gamma^{\max}(x) \stackrel{\text{def}}{=} \begin{cases} 0, & \text{if } \|x_i\| \leq \sqrt{c},\ i \in \{1, \cdots, N\} \\ \arg\max\limits_{i} \|x_i\|, & \text{otherwise.} \end{cases} \tag{15}$$

Definition 2 (Mean-value estimation strategy): The mean value estimation strategy profile ψ^{mean} is a vector of functions where each component is given by

$$\psi_i^{\text{mean}}(y_i) \stackrel{\text{def}}{=} \begin{cases} x_i, & \text{if } y_i = (i, x_i) \\ \mathbf{E}[X_i], & \text{otherwise} \end{cases} \quad i \in \{1, \cdots, N\}. \tag{16}$$

In this section we present two results for a class of systems with independent observations. In the first result, we completely characterize a solution of Problem 1 with the unicast information structure under a symmetry and modality assumption on the probability density functions.

Theorem 1: Consider Problem 1 with the unicast information structure $\mathcal{I}_1$. If $X_1, \cdots, X_N$ are mutually independent, zero-mean continuous random vectors distributed according to circularly symmetric unimodal densities, then the pair of strategies $(\gamma^{\max}, \psi^{\text{mean}})$ is a global optimal solution to Problem 1.

Under an assumption on the symmetry of the probability density functions, the next result provides a person-by-person optimal solution to Problem 1 for the broadcast network model.

Theorem 2: Consider Problem 1 with the broadcast information structure $\mathcal{I}_2$. If $X_1, \cdots, X_N$ are mutually independent, zero-mean continuous random vectors distributed according to circularly symmetric densities, then the pair of strategies $(\gamma^{\max}, \psi^{\text{mean}})$ is a person-by-person optimal solution to Problem 1.

Remark 2: The scheduling strategy $\gamma^{\max}$ may be implemented in a *two-stage* decentralized architecture: the comparisons with the threshold $\sqrt{c}$ can be performed locally at the subsystems; and the comparison among states is performed by the network manager. The idea is that only the subsystems whose states have a norm larger than the threshold should request a transmission by the network manager. Therefore, the parameter c controls the communication rate between subsystems and estimators.

Remark 3: Theorem 1 is also valid for discrete random vectors X_i under appropriate symmetry assumptions on the probability mass functions p_{X_i}, $i \in \{1, \cdots, N\}$. We also note that the broadcast information structure $\mathcal{I}_2$ provides more information to the estimators than $\mathcal{I}_1$, which means that the optimal performance of a globally optimal solution of Problem 1 under $\mathcal{I}_2$ should be at least as good as the optimal performance under $\mathcal{I}_1$.

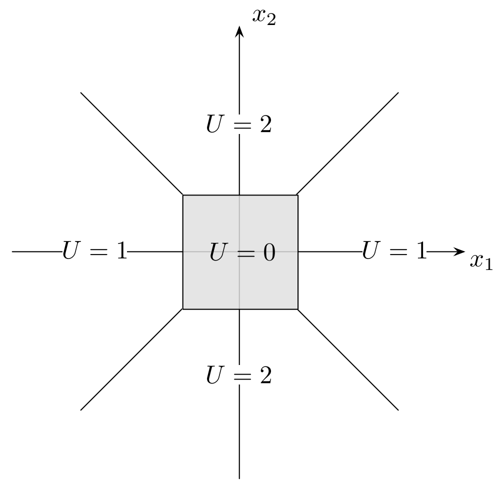

Fig. 3. Partition of $\mathbb{R}^2$ induced by $\gamma^{\max}$ for the case of two scalar subsystems. The sides of the shaded square have length $2\sqrt{c}$.

IV. Proofs of Theorems 1 and 2

A. Unicast information structure and proof of Theorem 1

Lemma 1: Problem 1 can be written as a finite dimensional optimization problem over $\mathbb{R}^n$.

Proof: Using total expectation, we decompose the cost functional as

$$\mathcal{J}(\gamma, \psi) = \mathbf{E}\left[\sum_{i=1}^{N} \|X_i - \hat{X}_i\|^2 \;\Big|\; U = 0\right] \mathbf{P}(U = 0) + \sum_{j=1}^{N} \mathbf{E}\left[\sum_{i=1}^{N} \|X_i - \hat{X}_i\|^2 + c \;\Big|\; U = j\right] \mathbf{P}(U = j). \tag{17}$$

Let $i \in \{1, \cdots, N\}$. Under $\mathcal{I}_1$, the estimator $\mathcal{E}_i$ either receives $Y_i = (i, x_i)$ or $Y_i = \varnothing$. For any given $\gamma \in \Gamma$, the optimal estimation strategy profile ψ^*_γ is given by

$$\psi^*_{\gamma,i}(y_i) = \begin{cases} x_i, & \text{if } y_i = (i, x_i) \\ \mathbf{E}[X_i \mid \gamma(X) \neq i], & \text{if } y_i = \varnothing \end{cases} \quad i \in \{1, \cdots, N\}. \tag{18}$$

Constrain the search for estimation strategies to the following class parameterized by $\hat{x} \stackrel{\text{def}}{=} [\hat{x}_1, \cdots, \hat{x}_N] \in \mathbb{R}^n$. This class consists of functions of the form

$$\psi_{\hat{x}_i, i}(y_i) = \begin{cases} x_i, & \text{if } y_i = (i, x_i) \\ \hat{x}_i, & \text{otherwise} \end{cases} \quad i \in \{1, \cdots, N\}, \tag{19}$$

where $\hat{x}_i \in \mathbb{R}^{n_i}$. We denote this estimation strategy profile by $\psi_{\hat{x}}$. The optimal estimation strategy profile belongs to this class. Therefore, constraining to this class is without loss of optimality.

The jointly optimal scheduling and estimation strategies can be found by solving the following optimization problem

$$\underset{\hat{x} \in \mathbb{R}^n}{\text{minimize}} \left\{ \underset{\gamma \in \Gamma}{\text{minimize}} \;\; \mathcal{J}(\gamma, \psi_{\hat{x}}) \right\}, \tag{20}$$

where

$$\mathcal{J}(\gamma,\psi_{\hat{x}}) = \int_{\mathbb{R}^n} \Bigg[\Bigg(\sum_{i=1}^{N} \|x_i - \hat{x}_i\|^2 \Bigg) \mathbf{1}(\gamma(x)=0) + \sum_{j=1}^{N} \Bigg(\sum_{i\neq j} \|x_i - \hat{x}_i\|^2 + c \Bigg) \mathbf{1}(\gamma(x)=j) \Bigg] f_X(x)dx, \quad (21)$$

where we have used the following notation

$$\mathbf{1}(A) \overset{\text{def}}{=} \begin{cases} 1, & \text{if } A \text{ is true;} \\ 0, & \text{otherwise.} \end{cases} \quad (22)$$

Denote $\tilde{\mathcal{J}} : \mathbb{R}^n \to \mathbb{R}$ such that

$$\tilde{\mathcal{J}}(\hat{x}) \overset{\text{def}}{=} \min_{\gamma\in\Gamma} \quad \mathcal{J}(\gamma,\psi_{\hat{x}}). \quad (23)$$

Let $\mathbb{A}_j^{\gamma} \overset{\text{def}}{=} \{x \in \mathbb{R}^n \mid \gamma(x) = j\}$, $j \in \{0,1,\cdots,N\}$. The sets $\mathbb{A}_j^{\gamma}$ are disjoint, therefore, at a given $x \in \mathbb{R}^n$ only one of the indicator functions in Eq. (21) is equal to one and the remaining are equal to zero. Assigning the points $x \in \mathbb{R}^n$ to the sets $\mathbb{A}_j^{\gamma}$ such as to minimize Eq. (21), we construct a scheduling strategy parametrized by $\hat{x} \in \mathbb{R}^n$ that is optimal for a given $\psi_{\hat{x}}$. Denote this scheduling strategy by $\gamma_{\hat{x}}^*$, where $x \in \mathbb{R}^n$ is assigned to $\mathbb{A}_0^{\gamma_{\hat{x}}^*}$ if and only if

$$\|x_i - \hat{x}_i\| \leq \sqrt{c}, \quad i \in \{1,\cdots,N\}. \quad (24)$$

Similarly, $x \in \mathbb{R}^n$ is assigned to $\mathbb{A}_j^{\gamma_{\hat{x}}^*}$, $j \in \{1,\cdots,N\}$, if and only if

$$\begin{cases} \|x_j - \hat{x}_j\| > \sqrt{c}; \\ \|x_j - \hat{x}_j\| \geq \|x_k - \hat{x}_k\|, \quad k \neq j. \end{cases} \quad (25)$$

The resulting cost is a function of $\hat{x} \in \mathbb{R}^n$ and is equal to

$$\tilde{\mathcal{J}}(\hat{x}) = \mathbf{E}\Bigg[\min\Bigg\{ \sum_{i=1}^{N} \|X_i - \hat{x}_i\|^2, \sum_{i\neq 1} \|X_i - \hat{x}_i\|^2 + c, \sum_{i\neq 2} \|X_i - \hat{x}_i\|^2 + c, \cdots, \sum_{i\neq N} \|X_i - \hat{x}_i\|^2 + c \Bigg\} \Bigg]. \quad (26)$$

Therefore, Problem 1 reduces to

$$\underset{\hat{x}\in\mathbb{R}^n}{\text{minimize}} \quad \tilde{\mathcal{J}}(\hat{x}). \quad (27)$$

■

Let $\mathbf{0}_k$ denote the zero vector of dimension k.

Lemma 2: If $\hat{x}^* = \mathbf{0}_n$, then the optimal scheduling strategy is $\gamma^{\max}$.

Proof: The proof is an immediate consequence of Eqs. (24) and (25), replacing $\hat{x}_i = \mathbf{0}_{n_i}$, $i \in \{1,\cdots,N\}$. ■

We are now ready to present the proof of Theorem 1.

Proof: From Lemma 1, Problem 1 is equivalent to a finite dimensional optimization problem with variable $\hat{x} \in \mathbb{R}^n$. We shall show that under the assumptions of Theorem 1, the point $\hat{x}^* = \mathbf{0}_n$ is a global minimizer of $\tilde{\mathcal{J}}(\hat{x})$.

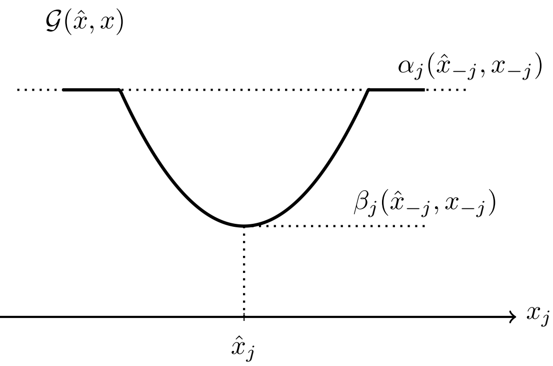

Fig. 4. Conceptual plot of $\mathcal{G}(\hat{x},x)$ as a function of x_j while keeping its remaining arguments fixed.

Define the following function

$$\mathcal{G}(\hat{x},x) \overset{\text{def}}{=} \min\Bigg\{ \sum_{i=1}^{N} \|x_i - \hat{x}_i\|^2, \sum_{i\neq 1} \|x_i - \hat{x}_i\|^2 + c, \cdots, \sum_{i\neq N} \|x_i - \hat{x}_i\|^2 + c \Bigg\}. \quad (28)$$

Let $j \in \{1,\cdots,N\}$ and note that $\mathcal{G}(\hat{x},x)$ can be expressed as

$$\mathcal{G}(\hat{x},x) = \min\Big\{ \alpha_j(\hat{x}_{-j},x_{-j}), \|x_j - \hat{x}_j\|^2 + \beta_j(\hat{x}_{-j},x_{-j}) \Big\}, \quad (29)$$

where

$$\alpha_j(\hat{x}_{-j},x_{-j}) \overset{\text{def}}{=} \sum_{i\neq j} \|x_i - \hat{x}_i\|^2 + c \quad (30)$$

and

$$\beta_j(\hat{x}_{-j},x_{-j}) \overset{\text{def}}{=} \min\Bigg\{ \sum_{i\neq j} \|x_i - \hat{x}_i\|^2, \sum_{i\neq\{1,j\}} \|x_i - \hat{x}_i\|^2 + c, \cdots, \sum_{i\neq\{j-1,j\}} \|x_i - \hat{x}_i\|^2 + c, \sum_{i\neq\{j+1,j\}} \|x_i - \hat{x}_i\|^2 + c, \cdots, \sum_{i\neq\{N,j\}} \|x_i - \hat{x}_i\|^2 + c \Bigg\}. \quad (31)$$

Therefore, $\mathcal{G}(\hat{x},x)$ satisfies the following property:

$$\lim_{\|x_j\|\to+\infty} \mathcal{G}(\hat{x},x) = \alpha_j(\hat{x}_{-j},x_{-j}), \quad (32)$$

which can be visualized in Fig. 4.

Using the mutual independence assumption, we rewrite the cost as

$$\tilde{\mathcal{J}}(\hat{x}) = \int_{\mathbb{R}^{n_1}} \cdots \int_{\mathbb{R}^{n_N}} \mathcal{G}(\hat{x},x) f_{X_1}(x_1) \cdots f_{X_N}(x_N) dx_1 \cdots dx_N. \quad (33)$$

Define the following function

$$\tilde{\mathcal{J}}_j(\hat{x}_j,\hat{x}_{-j},x_{-j}) \overset{\text{def}}{=} \int_{\mathbb{R}^{n_j}} \mathcal{G}(\hat{x},x) f_{X_j}(x_j) dx_j. \quad (34)$$

Our argument makes use of the Hardy-Littlewood inequality (see Appendix I) to show that

$$\tilde{\mathcal{J}}_j(\hat{x}_j, \hat{x}_{-j}, x_{-j}) \geq \tilde{\mathcal{J}}_j(\mathbf{0}_{n_j}, \hat{x}_{-j}, x_{-j}). \tag{35}$$

In order to show that, let $\mathcal{H}_j$ be defined as

$$\mathcal{H}_j(\hat{x}_j, x_j, \hat{x}_{-j}, x_{-j}) \stackrel{\text{def}}{=} \alpha_j(\hat{x}_{-j}, x_{-j}) - \mathcal{G}(\hat{x}, x). \tag{36}$$

This function vanishes at infinity, i.e.,

$$\lim_{\|x_j\| \to +\infty} \mathcal{H}_j(\hat{x}_j, x_j, \hat{x}_{-j}, x_{-j}) = 0. \tag{37}$$

Moreover, the *symmetric decreasing rearrangement* of $\mathcal{H}_j$ with respect to x_j while keeping its remaining arguments fixed is

$$\mathcal{H}_j^{\downarrow}(\hat{x}_j, x_j, \hat{x}_{-j}, x_{-j}) = \mathcal{H}_j(\mathbf{0}_{n_j}, x_j, \hat{x}_{-j}, x_{-j}). \tag{38}$$

Recall that

$$\tilde{\mathcal{J}}_j(\hat{x}_j, \hat{x}_{-j}, x_{-j}) = \alpha_j(\hat{x}_{-j}, x_{-j}) - \int_{\mathbb{R}^{n_j}} \mathcal{H}_j(\hat{x}_j, x_j, \hat{x}_{-j}, x_{-j}) f_{X_j}(x_j) dx_j. \tag{39}$$

Since f_{X_j} is a circularly symmetric unimodal density with mean equal to $\mathbf{0}_{n_j}$, its symmetric decreasing rearrangement is given by

$$f_{X_j}^{\downarrow} = f_{X_j}. \tag{40}$$

The Hardy-Littlewood inequality implies that

$$\begin{aligned} \int_{\mathbb{R}^{n_j}} & \mathcal{H}_j(\hat{x}_j, x_j, \hat{x}_{-j}, x_{-j}) f_{X_j}(x_j) dx_j \\ & \leq \int_{\mathbb{R}^{n_j}} \mathcal{H}_j^{\downarrow}(\hat{x}_j, x_j, \hat{x}_{-j}, x_{-j}) f_{X_j}^{\downarrow}(x_j) dx_j; \quad (41) \\ & = \int_{\mathbb{R}^{n_j}} \mathcal{H}_j(\mathbf{0}_{n_j}, x_j, \hat{x}_{-j}, x_{-j}) f_{X_j}(x_j) dx_j. \quad (42) \end{aligned}$$

Therefore,

$$\tilde{\mathcal{J}}_j(\hat{x}_j, \hat{x}_{-j}, x_{-j}) \geq \tilde{\mathcal{J}}_j(\mathbf{0}_{n_j}, \hat{x}_{-j}, x_{-j}) \Rightarrow \hat{x}_j^* = \mathbf{0}_{n_j}. \tag{43}$$

Since this is true for any $\hat{x}_{-j} \in \mathbb{R}^{n-n_j}$, it is also true for $\hat{x}_{-j} = \mathbf{0}_{n-n_j}$. Repeating this argument for $j = 1, \cdots, N$ it follows that $\hat{x}^* = \mathbf{0}_n$. From Lemma 2, the strategy pair $(\gamma^{\max}, \psi^{\text{mean}})$ is globally optimal for Problem 1. ■

B. Broadcast information structure and proof of Theorem 2

Proof: Let $i, j \in \{1, \cdots, N\}$ such that $i \neq j$. Under $\mathcal{I}_2$, the estimator $\mathcal{E}_i$ either receives $Y_i = (i, x_i)$, $Y_i = (j, x_j)$ or $Y_i = \varnothing$. For any given $\gamma \in \Gamma$, the optimal estimation strategy profile ψ_γ^* is given by

$$\psi_{\gamma,i}^*(y_i) = \begin{cases} x_i, & \text{if } y_i = (i, x_i); \\ \mathbf{E}[X_i \mid \gamma(X) = j, X_j = x_j], & \text{if } y_i = (j, x_j); \\ \mathbf{E}[X_i \mid \gamma(X) = 0], & \text{if } y_i = \varnothing. \end{cases} \tag{44}$$

Without loss in optimality, constrain the search for optimal estimation strategies to the class of functions $\Psi_{\hat{x},\hat{g}}$ in which the functions are of the form

$$\psi_i(y_i) = \begin{cases} x_i, & \text{if } y = (i, x_i); \\ \hat{g}_{ij}(x_j), & \text{if } y = (j, x_j); \\ \hat{x}_i, & \text{if } y = \varnothing, \end{cases} \tag{45}$$

where $\hat{x}_i \in \mathbb{R}^{n_i}$ and $\hat{g}_{ij} : \mathbb{R}^{n_j} \to \mathbb{R}^{n_i}$. Note that the optimal representation points and functions depend on γ and are given by

$$\hat{x}_i^* = \mathbf{E}[X_i \mid \gamma(X) = 0]; \tag{46}$$

$$\hat{g}_{ij}^*(x_j) = \mathbf{E}[X_i \mid \gamma(X) = j, X_j = x_j]. \tag{47}$$

For estimation strategies within $\Psi_{\hat{x},\hat{g}}$, the cost becomes

$$\mathcal{J}(\gamma, \psi) = \int_{\mathbb{R}^n} \Big(\sum_{i=1}^{N} \|x_i - \hat{x}_i\|^2 \Big) \mathbf{1}(\gamma(x) = 0) f_X(x) dx + \sum_{j=1}^{N} \int_{\mathbb{R}^n} \Big(\sum_{i \neq j} \|x_i - \hat{g}_{ij}(x_j)\|^2 + c \Big) \mathbf{1}(\gamma(x) = j) f_X(x) dx. \tag{48}$$

For fixed vectors $\hat{x}_i$ and maps $\hat{g}_{ij}$, an optimal scheduling strategy can be constructed and is given by:

$$\gamma_{\hat{x},\hat{g}}^*(x) = 0 \Leftrightarrow \sum_{i=1}^{N} \|x_i - \hat{x}_i\|^2 \leq \sum_{i \neq j} \|x_i - \hat{g}_{ij}(x_j)\|^2 + c, \ \ j \in \{1, \cdots, N\}; \tag{49}$$

and

$$\gamma_{\hat{x},\hat{g}}^*(x) = j \Leftrightarrow \begin{cases} \sum_{i \neq j} \|x_i - \hat{g}_{ij}(x_j)\|^2 + c < \sum_{i=1}^{N} \|x_i - \hat{x}_i\|^2 \\ \sum_{i \neq j} \|x_i - \hat{g}_{ij}(x_j)\|^2 \leq \sum_{i \neq k} \|x_i - \hat{g}_{ik}(x_k)\|^2, \end{cases} \quad k \in \{1, \cdots, N\}. \tag{50}$$

Using this scheduling strategy, we can rewrite the cost as a function of the estimation strategy profile $\psi \in \Psi_{\hat{x},\hat{g}}$:

$$\tilde{\mathcal{J}}(\hat{x}, \hat{g}) \stackrel{\text{def}}{=} \mathbf{E}\Big[\min \Big\{ \sum_{i=1}^{N} \|X_i - \hat{x}_i\|^2, \sum_{i \neq 1} \|X_i - \hat{g}_{i1}(X_j)\|^2 + c, \cdots, \sum_{i \neq N} \|X_i - \hat{g}_{iN}(X_N)\|^2 + c \Big\} \Big]. \tag{51}$$

Minimizing $\tilde{\mathcal{J}}(\hat{x}, \hat{g})$ is a nonconvex *functional* optimization problem.[1]

If $\hat{x}_i = \mathbf{0}_{n_i}$ and $\hat{g}_{ij}(\cdot) \equiv \mathbf{0}_{n_i}$, $i, j \in \{1, \cdots, N\}$, then Eqs. (49) and (50) imply that $\gamma_{\hat{x},\hat{g}}^* = \gamma^{\max}$. Conversely, if $\gamma = \gamma^{\max}$, then

$$\hat{x}_i^* = \frac{\int_{\mathbb{R}^n} x_i \mathbf{1}(\gamma(x) = 0) f_X(x) dx}{\int_{\mathbb{R}^n} \mathbf{1}(\gamma(x) = 0) f_X(x) dx}. \tag{52}$$

The set

$$\big\{ x \in \mathbb{R}^n \mid \gamma(x) = 0 \big\} = \big\{ x \in \mathbb{R}^n \mid \|x_1\| \leq \sqrt{c}, \cdots, \|x_N\| \leq \sqrt{c} \big\}. \tag{53}$$

[1]Finding global minimizers of $\tilde{\mathcal{J}}(\hat{x}, \hat{g})$ is an open problem.

Using the independence assumption, the integral in the numerator of Eq. (52) is

$$\int_{\|x_1\|\leq\sqrt{c}} \cdots \int_{\|x_N\|\leq\sqrt{c}} x_i f_{X_1}(x_1)\cdots f_{X_N}(x_N) dx_1 \cdots dx_N. \tag{54}$$

Since f_{X_i} is a circularly symmetric probability density function,

$$\int_{\|x_i\|\leq\sqrt{c}} x_i f_{X_i}(x_i) dx_i = \mathbf{0}_{n_i} \Rightarrow \hat{x}_i^* = \mathbf{0}_{n_i}. \tag{55}$$

Finally, for $\tau \in \mathbb{R}^{n_j}$ we compute

$$\hat{g}_{ij}^*(\tau) = \frac{\int_{\mathbb{R}^n} x_i \mathbf{1}(\gamma(x) = j, x_j = \tau) f_X(x) dx}{\int_{\mathbb{R}^n} \mathbf{1}(\gamma(x) = j, x_j = \tau) f_X(x) dx}. \tag{56}$$

The set over which the indicator function is equal to one is

$$\big\{x \in \mathbb{R}^n \mid \|x_j\| \geq \sqrt{c}, \|x_j\| \geq \|x_1\|, \cdots, \|x_j\| \geq \|x_N\|, x_j = \tau\big\}. \tag{57}$$

Using the independence assumption, the numerator in Eq. (56) is determined by

$$\int_{\|x_i\|\leq\|\tau\|} x_i f_{X_i}(x_i) dx_i \equiv \mathbf{0}_{n_i} \Rightarrow \hat{g}_{ij}^*(\tau) \equiv \mathbf{0}_{n_i}. \tag{58}$$

Therefore, the pair of strategies $\gamma^{\max}$ and ψ^{mean} is person-by-person optimal. ■

V. Conclusions

In this paper, we studied the problem of optimal scheduling in a distributed remote estimation system communicating over a shared communication medium. The access to the communication resources is granted by a network manager, which implements a medium access control strategy used to avoid packet collisions. Two network models were considered: the unicast and broadcast models. Under independence and symmetry assumptions on the probabilistic model that specifies the problem, we obtained the following optimality results. In the unicast case, we established the global optimality of the max-norm scheduling and mean-value estimation strategies. In the broadcast case, we showed the person-by-person optimality of the same pair. For the latter, we conjecture that the person-by-person solution is also globally optimal. These results are evidence that architectures that use dynamic scheduling in conjunction with event-based strategies used in previous works, may in fact be optimal for certain classes of systems. Future research on this topic include: extending the framework to consider correlated subsystems; sequential problem formulations; and proving the global optimality of the max-norm scheduling and mean-value estimation for the broadcast network.

Appendix I
Auxiliary results

The following two definitions and theorem can be found in [15] and in [16].

Definition 3 (Symmetric rearrangement): Let $\mathbb{A}$ be a measurable set of finite volume in $\mathbb{R}^n$. Its symmetric rearrangement $\mathbb{A}^*$ is defined as the open ball centered at $\mathbf{0}_n$ whose volume agrees with $\mathbb{A}$.

Definition 4 (Symmetric decreasing rearrangement): Let $f : \mathbb{R}^n \to \mathbb{R}$ be a nonnegative measurable function that vanishes at infinity. The symmetric decreasing rearrangement $f^{\downarrow}$ of f is

$$f^{\downarrow}(x) \overset{\mathrm{def}}{=} \int_0^{\infty} \mathbf{1}\Big(x \in \{\xi \in \mathbb{R}^n \mid f(\xi) > t\}^*\Big) dt. \tag{59}$$

Theorem 3 (Hardy-Littlewood Inequality): If f and g are two nonnegative measurable functions defined on $\mathbb{R}^n$ which vanish at infinity, then the following holds:

$$\int_{\mathbb{R}^n} f(x)g(x)dx \leq \int_{\mathbb{R}^n} f^{\downarrow}(x)g^{\downarrow}(x)dx, \tag{60}$$

where $f^{\downarrow}$ and $g^{\downarrow}$ are the symmetric decreasing rearrangements of f and g, respectively.

References

[1] K.-D. Kim and P. R. Kumar, "Cyber–physical systems: A perspective at the centennial," *Proceedings of the IEEE*, vol. 100, pp. 1287–1308, May 2012.

[2] S. Yuksel and T. Basar, *Stochastic Networked Control Systems*. Springer, 2013.

[3] M. M. Vasconcelos and N. C. Martins, "Optimal estimation over the collision channel," *IEEE Transactions on Automatic Control*, vol. 62, no. 1, pp. 321–336, January 2017.

[4] C. Ramesh, "Contention-based multiple access architectures for networked control systems," Ph.D. dissertation, KTH Royal Institute of Technology, 2011.

[5] G. C. Walsh, H. Ye, and L. G. Bushnell, "Stability analysis of networked control systems," *IEEE Transaction on Control Systems Technology*, vol. 10, no. 3, pp. 438–446, May 2002.

[6] A. Cervin and T. Henningsson, "Scheduling of event-triggered controllers on a shared network," in *Proceedings of the 47th IEEE Conference on Decision and Control*, 2008, pp. 3601–3606.

[7] E. Henriksson, D. E. Quevedo, E. G. W. Peters, H. Sandberg, and K. H. Johansson, "Multiple loop self-triggered model predictive control for network scheduling and control," *IEEE Transaction on Control Systems Technology*, vol. 23, no. 6, pp. 2167–2181, 2015.

[8] A. Molin and S. Hirche, "Price-based adaptive scheduling in multiloop control systems with resource constraints," *IEEE Transactions on Automatic Control*, vol. 59, no. 12, pp. 3282–3295, December 2014.

[9] M. Rabi, L. Stabellini, A. Proutiere, and M. Johansson, "Networked estimation under contention-based medium access," *International Journal of Robust and Nonlinear Control*, vol. 20, no. 2, pp. 140–155, January 2009.

[10] M. Xia, V. Gupta, and P. J. Antsaklis, "Networked state estimation over a shared communication medium," *IEEE Transactions on Automatic Control*, 2017.

[11] G. M. Lipsa and N. C. Martins, "Remote state estimation with communication costs for first-order LTI systems," *IEEE Transactions on Automatic Control*, vol. 56, no. 9, pp. 2013–2025, September 2011.

[12] A. Nayyar, T. Basar, D. Teneketzis, and V. Veeravalli, "Optimal strategies for communication and remote estimation with an energy harvesting sensor," *IEEE Transactions on Automatic Control*, vol. 58, no. 9, pp. 2246–2260, September 2013.

[13] Y.-C. Ho, M. Kastner, and E. Wong, "Teams, signaling, and information theory," *IEEE Transactions on Automatic Control*, vol. 23, no. 2, pp. 305–312, April 1978.

[14] H. S. Witsenhausen, "A counter-example in stochastic optimum control," *SIAM Journal on Control*, vol. 6, no. 1, pp. 131–147, 1968.

[15] A. Burchard, "A short course on rearrangement inequalities," in *Lecture notes of the IMDEA Winter School*, 2009.

[16] G. H. Hardy, J. E. Littlewood, and G. Pólya, *Inequalities*. Cambridge University Press, 1952.