2018 Annual American Control Conference (ACC)
June 27–29, 2018. Wisconsin Center, Milwaukee, USA

Sensing-Constrained LQG Control

Vasileios Tzoumas,[1,2] Luca Carlone,[2] George J. Pappas,[1] Ali Jadbabaie[2]

***Abstract*—Linear-Quadratic-Gaussian (LQG) control is concerned with the design of an optimal controller and estimator for linear Gaussian systems with imperfect state information. Standard LQG assumes the set of sensor measurements, to be fed to the estimator, to be given. However, in many problems, arising in networked systems and robotics, one may not be able to use all the available sensors, due to power or payload constraints, or may be interested in using the smallest subset of sensors that guarantees the attainment of a desired control goal. In this paper, we introduce the *sensing-constrained LQG control* problem, in which one has to jointly design sensing, estimation, and control, under given constraints on the resources spent for sensing. We focus on the realistic case in which the sensing strategy has to be selected among a finite set of possible sensing modalities. While the computation of the optimal sensing strategy is intractable, we present the first scalable algorithm that computes a near-optimal sensing strategy with provable sub-optimality guarantees. To this end, we show that a separation principle holds, which allows the design of sensing, estimation, and control policies in isolation. We conclude the paper by discussing two applications of sensing-constrained LQG control, namely, *sensing-constrained formation control* and *resource-constrained robot navigation*.**

I. Introduction

Traditional approaches to control of systems with partially observable state assume the choice of sensors used to observe the system is given. The choice of sensors usually results from a preliminary design phase in which an expert designer selects a suitable sensor suite that accommodates estimation requirements (e.g., observability, desired estimation error) and system constraints (e.g., size, cost). Modern control applications, from large networked systems to miniaturized robotics systems, pose serious limitations to the applicability of this traditional paradigm. In large-scale networked systems (e.g., smart grids or robot swarms), in which new nodes are continuously added and removed from the network, a manual re-design of the sensors becomes cumbersome and expensive, and it is simply not scalable. In miniaturized robot systems, while the set of onboard sensors is fixed, it may be desirable to selectively activate only a subset of the sensors during different phases of operation, in order to minimize power consumption. In both application scenarios, one usually has access to a (possibly large) list of potential sensors, but, due to resource constraints (e.g., cost, power), can only utilize a subset of them. Moreover, the need for online and large-scale sensor selection demands for automated approaches that efficiently select a subset of sensors to maximize system performance.

Motivated by these applications, in this paper we consider the problem of jointly designing control, estimation, and sensor selection for a system with partially observable state.

Related work. One body of related work is *control over band-limited communication channels*, which investigates the trade-offs between communication constraints (e.g., data rate, quantization, delays) and control performance (e.g., stability) in networked control systems. Early work provides results on the impact of quantization [1], finite data rates [2], [3], and separation principles for LQG design with communication constraints [4]; more recent work focuses on privacy constraints [5]. We refer the reader to the surveys [6]–[8]. A second set of related work is *sensor selection and scheduling*, in which one has to select a (possibly time-varying) set of sensors in order to monitor a phenomenon of interest. Related literature includes approaches based on randomized sensor selection [9], dual volume sampling [10], [11], convex relaxations [12], [13], and submodularity [14]–[16]. The third set of related works is *information-constrained (or information-regularized)* LQG control [17], [18]. Shafieepoorfard and Raginsky [17] study rationally inattentive control laws for LQG control and discuss their effectiveness in stabilizing the system. Tanaka and Mitter [18] consider the co-design of sensing, control, and estimation, propose to augment the standard LQG cost with an information-theoretic regularizer, and derive an elegant solution based on semidefinite programming. The main difference between our proposal and [18] is that we consider the case in which the choice of sensors, rather than being arbitrary, is restricted to a finite set of available sensors.

Contributions. We extend the Linear-Quadratic-Gaussian (LQG) control to the case in which, besides designing an optimal controller and estimator, one has to select a set of sensors to be used to observe the system state. In particular, we formulate the *sensing-constrained* (finite-horizon) LQG problem as the joint design of an optimal control and estimation policy, as well as the selection of a subset of k out of N available sensors, that minimize the LQG objective, which quantifies tracking performance and control effort. We first leverage a separation principle to show that the design of sensing, control, and estimation, can be performed independently. While the computation of the optimal sensing strategy is combinatorial in nature, a key contribution of this paper is to provide the first scalable algorithm that computes a near-optimal sensing strategy with provable sub-optimality guarantees. We motivate the importance of the sensing-constrained LQG problem, and demonstrate the effectiveness of the proposed algorithm in nu-

[1]The authors are with the Department of Electrical and Systems Engineering, University of Pennsylvania, Philadelphia, PA 19104 USA (email: {pappagsg, vtzoumas}@seas.upenn.edu).

[2]The authors are with the Institute for Data, Systems and Society, and the Laboratory for Information and Decision Systems, Massachusetts Institute of Technology, Cambridge, MA 02139 USA (email: {jadbabai, lcarlone, vtzoumas}@mit.edu).

This work was supported in part by TerraSwarm, one of six centers of STARnet, a Semiconductor Research Corporation program sponsored by MARCO and DARPA, and in part by AFOSR Complex Networks Program.

merical experiments, by considering two application scenarios, namely, *sensing-constrained formation control* and *resource-constrained robot navigation*, which, due to page limitations, we include in the full version of this paper, located at the authors' websites. All proofs can be found also in the full version of this paper, located at the authors' websites.

Notation. Lowercase letters denote vectors and scalars, and uppercase letters denote matrices. We use calligraphic fonts to denote sets. The identity matrix of size n is denoted with $\mathbf{I}_n$ (dimension is omitted when clear from the context). For a matrix M and a vector v of appropriate dimension, we define $\|v\|_M^2 \triangleq v^\mathsf{T} M v$. For matrices $M_1, M_2, \ldots, M_k$, we define $\operatorname{diag}(M_1, M_2, \ldots, M_k)$ as the block diagonal matrix with diagonal blocks the $M_1, M_2, \ldots, M_k$.

II. Sensing-Constrained LQG Control

In this section we formalize the sensing-constrained LQG control problem considered in this paper. We start by introducing the notions of *system*, *sensors*, and *control policies*.

a) System: We consider a standard discrete-time (possibly time-varying) linear system with additive Gaussian noise:

$$x_{t+1} = A_t x_t + B_t u_t + w_t, \quad t = 1, 2, \ldots, T, \tag{1}$$

where $x_t \in \mathbb{R}^{n_t}$ represents the state of the system at time t, $u_t \in \mathbb{R}^{m_t}$ represents the control action, w_t represents the process noise, and T is a finite time horizon. In addition, we consider the system's initial condition x_1 to be a Gaussian random variable with covariance $\Sigma_{1|0}$, and w_t to be a Gaussian random variable with mean zero and covariance W_t, such that w_t is independent of x_1 and $w_{t'}$ for all $t' = 1, 2, \ldots, T$, $t' \neq t$.

b) Sensors: We consider the case where we have a (potentially large) set of available sensors, which take noisy linear observations of the system's state. In particular, let $\mathcal{V}$ be a set of indices such that each index $i \in \mathcal{V}$ uniquely identifies a sensor that can be used to observe the state of the system. We consider sensors of the form

$$y_{i,t} = C_{i,t} x_t + v_{i,t}, \quad i \in \mathcal{V}, \tag{2}$$

where $y_{i,t} \in \mathbb{R}^{p_{i,t}}$ represents the measurement of sensor i at time t, and $v_{i,t}$ represents the measurement noise of sensor i. We assume $v_{i,t}$ to be a Gaussian random variable with mean zero and positive definite covariance $V_{i,t}$, such that $v_{i,t}$ is independent of x_1, and of $w_{t'}$ for any $t' \neq t$, and independent of $v_{i',t'}$ for all $t' \neq t$, and any $i' \in \mathcal{V}$, $i' \neq i$.

In this paper we are interested in the case in which we cannot use all the available sensors, and as a result, we need to select a convenient subset of sensors in $\mathcal{V}$ to maximize our control performance (formalized in Problem 1 below).

Definition 1 (Active sensor set and measurement model). *Given a set of available sensors $\mathcal{V}$, we say that $\mathcal{S} \subset \mathcal{V}$ is an* active sensor set *if we can observe the measurements from each sensor $i \in \mathcal{S}$ for all $t = 1, 2, \ldots, T$. Given an active sensor set $\mathcal{S} = \{i_1, i_2 \ldots, i_{|\mathcal{S}|}\}$, we define the following quantities*

$$\begin{aligned} y_t(\mathcal{S}) &\triangleq [y_{i_1,t}^\mathsf{T}, y_{i_2,t}^\mathsf{T}, \ldots, y_{i_{|\mathcal{S}|},t}^\mathsf{T}]^\mathsf{T}, \\ C_t(\mathcal{S}) &\triangleq [C_{i_1,t}^\mathsf{T}, C_{i_2,t}^\mathsf{T}, \ldots, C_{i_{|\mathcal{S}|},t}^\mathsf{T}]^\mathsf{T}, \\ V_t(\mathcal{S}) &\triangleq \mathit{diag}[V_{i_1,t}, V_{i_2,t}, \ldots, V_{i_{|\mathcal{S}|},t}] \end{aligned} \tag{3}$$

which lead to the definition of the measurement model*:*

$$y_t(\mathcal{S}) = C_t(\mathcal{S}) x_t + v_t(\mathcal{S}) \tag{4}$$

where $v_t(\mathcal{S})$ is a zero-mean Gaussian noise with covariance $V_t(\mathcal{S})$. Despite the availability of a possibly large set of sensors $\mathcal{V}$, our observer will only have access to the measurements produced by the active sensors.

The following paragraph formalizes how the choice of the active sensors affects the control policies.

c) Control policies: We consider control policies u_t for all $t = 1, 2, \ldots, T$ that are only informed by the measurements collected by the active sensors:

$$u_t = u_t(\mathcal{S}) = u_t(y_1(\mathcal{S}), y_2(\mathcal{S}), \ldots, y_t(\mathcal{S})), \quad t = 1, 2, \ldots, T.$$

Such policies are called *admissible*.

In this paper, we want to find a small set of active sensors $\mathcal{S}$, and admissible controllers $u_1(\mathcal{S}), u_2(\mathcal{S}), \ldots, u_T(\mathcal{S})$, to solve the following sensing-constrained LQG control problem.

Problem 1 (Sensing-constrained LQG control). *Find a sensor set $\mathcal{S} \subset \mathcal{V}$ of cardinality at most k to be active across all times $t = 1, 2, \ldots, T$, and control policies $u_{1:T}(\mathcal{S}) \triangleq \{u_1(\mathcal{S}), u_2(\mathcal{S}), \ldots, u_T(\mathcal{S})\}$, that minimize the LQG cost function:*

$$\min_{\substack{\mathcal{S} \subseteq \mathcal{V},\, |\mathcal{S}| \leq k,\\ u_{1:T}(\mathcal{S})}} \sum_{t=1}^{T} \mathbb{E}\left[\|x_{t+1}(\mathcal{S})\|_{Q_t}^2 + \|u_t(\mathcal{S})\|_{R_t}^2\right], \tag{5}$$

where the state-cost matrices $Q_1, Q_2, \ldots, Q_T$ are positive semi-definite, the control-cost matrices $R_1, R_2, \ldots, R_T$ are positive definite, and the expectation is taken with respect to the initial condition x_1, the process noises $w_1, w_2, \ldots, w_T$, and the measurement noises $v_1(\mathcal{S}), v_2(\mathcal{S}), \ldots, v_T(\mathcal{S})$.

Problem 1 generalizes the imperfect state-information LQG control problem from the case where all sensors in $\mathcal{V}$ are active, and only optimal control policies are to be found [19, Chapter 5], to the case where only a few sensors in $\mathcal{V}$ can be active, and both optimal sensors and control policies are to be found jointly. While we already noticed that admissible control policies depend on the active sensor set $\mathcal{S}$, it is worth noticing that this in turn implies that the state evolution also depends on $\mathcal{S}$; for this reason we write $x_{t+1}(\mathcal{S})$ in eq. (5). The intertwining between control and sensing calls for a joint design strategy. In the following section we focus on the design of a jointly optimal control and sensing solution to Problem 1.

III. Joint Sensing and Control Design

In this section we first present a separation principle that decouples sensing, estimation, and control, and allows designing them in cascade (Section III-A). We then present a scalable algorithm for sensing and control design (Section III-B).

Algorithm 1 Joint Sensing and Control design for Problem 1.

Input: Time horizon T, available sensor set $\mathcal{V}$, covariance matrix $\Sigma_{1|0}$ of initial condition x_1; for all $t = 1, 2, \ldots, T$, system matrix A_t, input matrix B_t, LQG cost matrices Q_t and R_t, process noise covariance matrix W_t; and for all sensors $i \in \mathcal{V}$, measurement matrix $C_{i,t}$, and measurement noise covariance matrix $V_{i,t}$.

Output: Active sensors $\widehat{\mathcal{S}}$, and control matrices $K_1, \ldots, K_T$.

```
Ŝ is returned by Algorithm 2 that finds a (possibly approximate) solution to the optimization problem in eq. (6);
K_1, ..., K_T are computed using the recursion in eq. (8).
```

A. Separability of Optimal Sensing and Control Design

We characterize the jointly optimal control and sensing solutions to Problem 1, and prove that they can be found in two separate steps, where first the sensing design is computed, and second the corresponding optimal control design is found.

Theorem 1 (Separability of optimal sensing and control design). *Let the sensor set $\mathcal{S}^\star$ and the controllers $u_1^\star, u_2^\star, \ldots, u_T^\star$ be a solution to the sensing-constrained LQG Problem 1. Then, $\mathcal{S}^\star$ and $u_1^\star, u_2^\star, \ldots, u_T^\star$ can be computed in cascade as follows:*

$$\mathcal{S}^\star \in \underset{\mathcal{S}\subseteq\mathcal{V}, |\mathcal{S}|\leq k}{\arg\min} \sum_{t=1}^{T} tr[\Theta_t \Sigma_{t|t}(\mathcal{S})], \tag{6}$$

$$u_t^\star = K_t \hat{x}_{t,\mathcal{S}^\star}, \quad t = 1, \ldots, T \tag{7}$$

where $\hat{x}_t(\mathcal{S})$ is the Kalman estimator of the state x_t, i.e., $\hat{x}_t(\mathcal{S}) \triangleq \mathbb{E}(x_t | y_1(\mathcal{S}), y_2(\mathcal{S}), \ldots, y_t(\mathcal{S}))$, and $\Sigma_{t|t}(\mathcal{S})$ is $\hat{x}_t(\mathcal{S})$'s error covariance, i.e., $\Sigma_{t|t}(\mathcal{S}) \triangleq \mathbb{E}[(\hat{x}_t(\mathcal{S}) - x_t)(\hat{x}_t(\mathcal{S}) - x_t)^\top]$ [19, Appendix E]. In addition, the matrices Θ_t and K_t are independent of the selected sensor set $\mathcal{S}$, and they are computed as follows: the matrices Θ_t and K_t are the solution of the backward Riccati recursion

$$\begin{aligned} S_t &= Q_t + N_{t+1}, \\ N_t &= A_t^\top (S_t^{-1} + B_t R_t^{-1} B_t^\top)^{-1} A_t, \\ M_t &= B_t^\top S_t B_t + R_t, \\ K_t &= -M_t^{-1} B_t^\top S_t A_t, \\ \Theta_t &= K_t^\top M_t K_t, \end{aligned} \tag{8}$$

with boundary condition $N_{T+1} = 0$.

Remark 1 (Certainty equivalence principle). *The control gain matrices $K_1, K_2, \ldots, K_T$ are the same as the ones that make the controllers $(K_1 x_1,\ K_1 x_2, \ldots, K_T x_T)$ optimal for the perfect state-information version of Problem 1, where the state x_t is known to the controllers [19, Chapter 4].*

Theorem 1 decouples the design of the sensing from the controller design. Moreover, it suggests that once an optimal sensor set $\mathcal{S}^\star$ is found, then the optimal controllers are equal to $K_t \hat{x}_t(\mathcal{S})$, which correspond to the standard LQG control policy. This should not come as a surprise, since for a given sensing strategy, Problem 1 reduces to standard LQG control.

We conclude this section with a remark providing a more intuitive interpretation of the sensor design step in eq. (6).

Algorithm 2 Sensing design for Problem 1.

Input: Time horizon T, available sensor set $\mathcal{V}$, covariance matrix $\Sigma_{1|0}$ of system's initial condition x_1, and for any time $t = 1, 2, \ldots, T$, any sensor $i \in \mathcal{V}$, process noise covariance matrix W_t, measurement matrix $C_{i,t}$, and measurement noise covariance matrix $V_{i,t}$.

Output: Sensor set $\widehat{\mathcal{S}}$.

Compute $\Theta_1, \Theta_2, \ldots, \Theta_T$ using recursion in eq. (8);
$\widehat{\mathcal{S}} \leftarrow \emptyset$; $\quad i \leftarrow 0$;
while $i < k$ **do**
  for all $a \in \mathcal{V} \setminus \widehat{\mathcal{S}}$ **do**
    $\widehat{\mathcal{S}}_a \leftarrow \widehat{\mathcal{S}} \cup \{a\}$; $\quad \Sigma_{1|0}(\widehat{\mathcal{S}}_a) \leftarrow \Sigma_{1|0}$;
    for all $t = 1, \ldots, T$ **do**
      $\Sigma_{t|t}(\widehat{\mathcal{S}}_a) \leftarrow$
      $[\Sigma_{t|t-1}(\widehat{\mathcal{S}}_a)^{-1} + C_t(\widehat{\mathcal{S}}_a)^\top V_t(\widehat{\mathcal{S}}_a)^{-1} C_t(\widehat{\mathcal{S}}_a)]^{-1}$;
      $\Sigma_{t+1|t}(\widehat{\mathcal{S}}_a) \leftarrow A_t \Sigma_{t|t}(\widehat{\mathcal{S}}_a) A_t^\top + W_t$;
    end for
    $\text{cost}_a \leftarrow \sum_{t=1}^{T} \text{tr}[\Theta_t \Sigma_{t|t}(\widehat{\mathcal{S}}_a)]$;
  end for
  $a_i \leftarrow \arg\min_{a\in\mathcal{V}\setminus\mathcal{S}} \text{cost}_a$;
  $\widehat{\mathcal{S}} \leftarrow \widehat{\mathcal{S}} \cup \{a_i\}$; $\quad i \leftarrow i + 1$;
end while

Remark 2 (Control-aware sensor design). *In order to provide more insight on the cost function in* (6)*, we rewrite it as:*

$$\begin{aligned} \sum_{t=1}^{T} tr[\Theta_t \Sigma_{t|t}(\mathcal{S})] &= \sum_{t=1}^{T} \mathbb{E}\left(tr\{[x_t - \hat{x}_t(\mathcal{S})]^\top \Theta_t [x_t - \hat{x}_t(\mathcal{S})]\} \right) \\ &= \sum_{t=1}^{T} \mathbb{E}\left(\|K_t x_t - K_t \hat{x}_t(\mathcal{S})\|_{M_t}^2 \right), \end{aligned} \tag{9}$$

where in the first line we used the fact that $\Sigma_{t|t}(\mathcal{S}) = \mathbb{E}\left[(x_t - \hat{x}_t(\mathcal{S}))(x_t - \hat{x}_t(\mathcal{S}))^\top\right]$, and in the second line we substituted the definition of $\Theta_t = K_t^\top M_t K_t$ from eq. (8)*.*

From eq. (9)*, it is clear that each term $tr[\Theta_t \Sigma_{t|t}(\mathcal{S})]$ captures the expected control mismatch between the imperfect state-information controller $u_t(\mathcal{S}) = K_t \hat{x}_t(\mathcal{S})$ (which is only aware of the measurements from the active sensors) and the perfect state-information controller $K_t x_t$. This is an important distinction from the existing sensor selection literature. In particular, while standard sensor selection attempts to minimize the estimation covariance, for instance by minimizing*

$$\sum_{t=1}^{T} tr[\Sigma_{t|t}(\mathcal{S})] \triangleq \sum_{t=1}^{T} \mathbb{E}\left(\|x_t - \hat{x}_t(\mathcal{S})\|_2^2 \right), \tag{10}$$

the proposed LQG cost formulation attempts to minimize the estimation error of only the informative states to the perfect state-information controller: for example, the contribution of all $x_t - \hat{x}_t(\mathcal{S})$ in the null space of K_t to the total control mismatch in eq. (9) *is zero. Hence, in contrast to minimizing the cost function in eq.* (10)*, minimizing the cost function in eq.* (9) *results to a control-aware sensing design.*

B. Scalable Near-optimal Sensing and Control Design

This section proposes a practical design algorithm for Problem 1. The pseudo-code of the algorithm is presented in Algorithm 1. Algorithm 1 follows the result of Theorem 1, and jointly designs sensing and control by first computing an active sensor set (line 1 in Algorithm 1) and then computing the control policy (line 2 in Algorithm 1). We discuss each step of the design process in the rest of this section.

1) Near-optimal Sensing design: The optimal sensor design can be computed by solving the optimization problem in eq. (6). The problem is combinatorial in nature, since it requires to select a subset of elements of cardinality k out of all the available sensors that induces the smallest cost.

In this section we propose a greedy algorithm, whose pseudo-code is given in Algorithm 2, that computes a (possibly approximate) solution to the problem in eq. (6). Our interest towards this greedy algorithm is motivated by the fact that it is scalable (in Section IV we show that its complexity is linear in the number of available sensors) and is provably close to the optimal solution of the problem in eq. (6) (we provide suboptimality bounds in Section IV).

Algorithm 2 computes the matrices Θ_t ($t = 1, 2, \ldots, T$) which appear in the cost function in eq. (6) (line 1). Note that these matrices are independent on the choice of sensors. The set of active sensors $\widehat{\mathcal{S}}$ is initialized to the empty set (line 2). The "while loop" in line 3 will be executed k times and at each time a sensor is greedily added to the set of active sensors $\widehat{\mathcal{S}}$. In particular, the "for loop" in lines 4-12 computes the estimation covariance resulting by adding a sensor to the current active sensor set and the corresponding cost (line 11). Finally, the sensor inducing the smallest cost is selected (line 13) and added to the current set of active sensors (line 14).

2) Control policy design: The optimal control design is computed as in eq. (7), where the control policy matrices $K_1, K_2, \ldots, K_T$ are obtained from the recursion in eq. (8).

In the following section we characterize the approximation and running-time performance of Algorithm 1.

IV. Performance Guarantees for Joint Sensing and Control Design

We prove that Algorithm 1 is the first scalable algorithm for the joint sensing and control design Problem 1, and that it achieves a value for the LQG cost function in eq. (5) that is finitely close to the optimal. We start by introducing the notion of supermodularity ratio (Section IV-A), which will enable to bound the sub-optimality gap of Algorithm 1 (Section IV-B).

A. Supermodularity ratio of monotone functions

We define the supermodularity ratio of monotone functions. We start with the notions of monotonicity and supermodularity.

Definition 2 (Monotonicity). *Consider any finite ground set $\mathcal{V}$. The set function $f : 2^{\mathcal{V}} \mapsto \mathbb{R}$ is non-increasing if and only if for any $\mathcal{A} \subseteq \mathcal{A}' \subseteq \mathcal{V}$, $f(\mathcal{A}) \geq f(\mathcal{A}')$.*

Definition 3 (Supermodularity [20, Proposition 2.1]). *Consider any finite ground set $\mathcal{V}$. The set function $f : 2^{\mathcal{V}} \mapsto \mathbb{R}$ is supermodular if and only if for any $\mathcal{A} \subseteq \mathcal{A}' \subseteq \mathcal{V}$ and $x \in \mathcal{V}$, $f(\mathcal{A}) - f(\mathcal{A} \cup \{x\}) \geq f(\mathcal{A}') - f(\mathcal{A}' \cup \{x\})$.*

In words, a set function f is supermodular if and only if it satisfies the following intuitive diminishing returns property: for any $x \in \mathcal{V}$, the marginal drop $f(\mathcal{A}) - f(\mathcal{A} \cup \{x\})$ diminishes as $\mathcal{A}$ grows; equivalently, for any $\mathcal{A} \subseteq \mathcal{V}$ and $x \in \mathcal{V}$, the marginal drop $f(\mathcal{A}) - f(\mathcal{A} \cup \{x\})$ is non-increasing.

Definition 4 (Supermodularity ratio). *Consider any finite ground set $\mathcal{V}$, and a non-increasing set function $f : 2^{\mathcal{V}} \mapsto \mathbb{R}$. We define the supermodularity ratio of f as*

$$\gamma_f = \min_{\mathcal{A} \subseteq \mathcal{V}, x, x' \in \mathcal{V} \setminus \mathcal{A}} \frac{f(\mathcal{A}) - f(\mathcal{A} \cup \{x\})}{f(\mathcal{A} \cup \{x'\}) - f[(\mathcal{A} \cup \{x'\}) \cup \{x\}]}.$$

In words, the supermodularity ratio of a monotone set function f measures how far f is from being supermodular. In particular, per the Definition 4 of supermodularity ratio, the supermodularity ratio γ_f takes values in $[0, 1]$, and

- $\gamma_f = 1$ if and only if f is supermodular, since if $\gamma_f = 1$, then Definition 4 implies $f(\mathcal{A}) - f(\mathcal{A} \cup \{x\}) \geq f(\mathcal{A} \cup \{x'\}) - f[(\mathcal{A} \cup \{x'\}) \cup \{x\}]$, i.e., the drop $f(\mathcal{A}) - f(\mathcal{A} \cup \{x\})$ is non-increasing as new elements are added in $\mathcal{A}$.
- $\gamma_f < 1$ if and only if f is *approximately supermodular*, in the sense that if $\gamma_f < 1$, then Definition 4 implies $f(\mathcal{A}) - f(\mathcal{A} \cup \{x\}) \geq \gamma_f \{f(\mathcal{A} \cup \{x'\}) - f[(\mathcal{A} \cup \{x'\}) \cup \{x\}]\}$, i.e., the drop $f(\mathcal{A}) - f(\mathcal{A} \cup \{x\})$ is approximately non-increasing as new elements are added in $\mathcal{A}$; specifically, the supermodularity ratio γ_f captures how much ones needs to discount the drop $f(\mathcal{A} \cup \{x'\}) - f[(\mathcal{A} \cup \{x'\}) \cup \{x\}]$, such that $f(\mathcal{A}) - f(\mathcal{A} \cup \{x\})$ remains greater then, or equal to, $f(\mathcal{A} \cup \{x'\}) - f[(\mathcal{A} \cup \{x'\}) \cup \{x\}]$.

We next use the notion of supermodularity ratio Definition 4 to quantify the sub-optimality gap of Algorithm 1.

B. Performance Analysis for Algorithm 1

We quantify Algorithm 1's running time, as well as, Algorithm 1's approximation performance, using the notion of supermodularity ratio introduced in Section IV-A. We conclude the section by showing that for appropriate LQG cost matrices $Q_1, Q_2, \ldots, Q_T$ and $R_1, R_2, \ldots, R_T$, Algorithm 1 achieves near-optimal approximate performance.

Theorem 2 (Performance of Algorithm 1). *For any active sensor set $\mathcal{S} \subseteq \mathcal{V}$, and admissible control policies $u_{1:T}(\mathcal{S}) \triangleq \{u_1(\mathcal{S}), u_2(\mathcal{S}), \ldots, u_T(\mathcal{S})\}$, let $h[\mathcal{S}, u_{1:T}(\mathcal{S})]$ be Problem 1's cost function, i.e.,*

$$h[\mathcal{S}, u_{1:T}(\mathcal{S})] \triangleq \textstyle\sum_{t=1}^{T} \mathbb{E}(\|x_{t+1}(\mathcal{S})\|^2_{Q_t} + \|u_t(\mathcal{S})\|^2_{R_t});$$

Further define the following set-valued function and scalar:

$$g(\mathcal{S}) \triangleq \min_{u_{1:T}(\mathcal{S})} h[\mathcal{S}, u_{1:T}(\mathcal{S})], \tag{11}$$

$$g^\star \triangleq \min_{\substack{\mathcal{S} \subseteq \mathcal{V}, |\mathcal{S}| \leq k, \\ u_{1:T}(\mathcal{S})}} h[\mathcal{S}, u_{1:T}(\mathcal{S})].$$

The following results hold true:

1) *(Approximation quality) Algorithm 1 returns an active sensor set $\widehat{\mathcal{S}} \subset \mathcal{V}$ of cardinality k, and gain matrices $K_1, K_2, \ldots, K_T$, such that the cost $h[\widehat{\mathcal{S}}, u_{1:T}(\widehat{\mathcal{S}})]$ attained by the sensor set $\widehat{\mathcal{S}}$ and the corresponding control policies $u_{1:T}(\widehat{\mathcal{S}}) \triangleq \{K_1\hat{x}_1(\widehat{\mathcal{S}}), \ldots, K_T\hat{x}_T(\widehat{\mathcal{S}})\}$ satisfies*
$$\frac{h(\widehat{\mathcal{S}}, u_{1:T}(\widehat{\mathcal{S}})) - g^\star}{g(\emptyset) - g^\star} \leq \exp(-\gamma_g) \tag{12}$$
where γ_g is the supermodularity ratio of $g(\mathcal{S})$ in eq. (11).
2) *(Running time) Algorithm 1 runs in $O(k|\mathcal{V}|Tn^{2.4})$ time, where $n \triangleq \max_{t=1,2,\ldots,T}(n_t)$ is the maximum system size in eq.* (1).

Theorem 2 ensures that Algorithm 1 is the first scalable algorithm for the sensing-constrained LQG control Problem 1. In particular, Algorithm 1's running time $O(k|\mathcal{V}|Tn^{2.4})$ is linear both in the number of available sensors $|\mathcal{V}|$, and the sensor set cardinality constraint k, as well as, linear in the Kalman filter's running time across the time horizon $\{1, 2 \ldots, T\}$. Specifically, the contribution $n^{2.4}T$ in Algorithm 1's running time comes from the computational complexity of using the Kalman filter to compute the state estimation error covariances $\Sigma_{t|t}$ for each $t = 1, 2, \ldots, T$ [19, Appendix E].

Theorem 2 also guarantees that for non-zero ratio γ_g Algorithm 1 achieves a value for Problem 1 that is finitely close to the optimal. In particular, the bound in ineq. (12) improves as γ_g increases, since it is decreasing in γ_g, and is characterized by the following extreme behaviors: for $\gamma_g = 1$, the bound in ineq. (12) is $e^{-1} \simeq .37$, which is the minimum for any $\gamma_g \in [0, 1]$, and hence, the best bound on Algorithm 1's approximation performance among all $\gamma_g \in [0, 1]$ (ideally, the bound in ineq. (12) would be 0 for $\gamma_g = 1$, in which case Algorithm 1 would be exact, since it would be implied $h(\widehat{\mathcal{S}}, u_{1:T}(\widehat{\mathcal{S}})) = g^\star$; however, even for supermodular functions, the best bound one can achieve in the worst-case is e^{-1} [21]); for $\gamma_g = 0$, ineq. (12) is uninformative since it simplifies to $h(\widehat{\mathcal{S}}, u_{1:T}(\widehat{\mathcal{S}})) \leq g(\emptyset) = h(\emptyset, u_{1:T}(\emptyset))$, which is trivially satisfied.[1]

In the remaining of the section, we first prove that if the strict inequality $\sum_{t=1}^T \Theta_t \succ 0$ holds, where each Θ_t is defined as in eq. (8), then the ratio γ_g in ineq. (12) is non-zero, and as result Algorithm 1 achieves a near-optimal approximation performance (Theorem 3). Then, we prove that the strict inequality $\sum_{t=1}^T \Theta_t \succ 0$ holds true in all LQG control problem instances where a zero controller would result in a suboptimal behavior of the system and, as a result, LQG control design (through solving Problem 1) is necessary to achieve their desired system performance (Theorem 4).

Theorem 3 (Lower bound for supermodularity ratio γ_g). *Let Θ_t for all $t = 1, 2, \ldots, T$ be defined as in eq.* (8), *$g(\mathcal{S})$ be defined as in eq.* (11), *and for any sensor $i \in \mathcal{V}$, $\bar{C}_{i,t}$ be the normalized measurement matrix $V_{i,t}^{-1/2}C_{i,t}$.*

If $\sum_{t=1}^T \Theta_t \succ 0$, the supermodularity ratio γ_g is non-zero. In addition, if we consider for simplicity that the Frobenius norm of each $\bar{C}_{i,t}$ is 1, i.e., tr $\left(\bar{C}_{i,t}\bar{C}_{i,t}^\mathsf{T}\right) = 1$, and that tr$[\Sigma_{t|t}(\emptyset)] \leq \lambda^2_{\max}[\Sigma_{t|t}(\emptyset)]$, γ_g's lower bound is
$$\gamma_g \geq \frac{\lambda_{\min}(\sum_{t=1}^T \Theta_t)}{\lambda_{\max}(\sum_{t=1}^T \Theta_t)} \frac{\min_{t\in\{1,2,\ldots,T\}} \lambda^2_{\min}[\Sigma_{t|t}(\mathcal{V})]}{\max_{t\in\{1,2,\ldots,T\}} \lambda^2_{\max}[\Sigma_{t|t}(\emptyset)]} \frac{1 + \min_{i\in\mathcal{V},t\in\{1,2\ldots,T\}} \lambda_{\min}[\bar{C}_i\Sigma_{t|t}(\mathcal{V})\bar{C}_i^\mathsf{T}]}{2 + \max_{i\in\mathcal{V},t\in\{1,2\ldots,T\}} \lambda_{\max}[\bar{C}_i\Sigma_{t|t}(\emptyset)\bar{C}_i^\mathsf{T}]}. \tag{13}$$

The supermodularity ratio bound in ineq. (13) suggests two cases under which γ_g can increase, and correspondingly, the performance bound of Algorithm 1 in eq. (12) can improve:

a) Case 1 where γ_g's bound in ineq. (13) *increases:* When the fraction $\lambda_{\min}(\sum_{t=1}^T \Theta_t)/\lambda_{\max}(\sum_{t=1}^T \Theta_t)$ increases to 1, then the right-hand-side in ineq. (13) increases. Equivalently, the right-hand-side in ineq. (13) increases when on average all the directions $x_t^{(i)} - \hat{x}_t^{(i)}$ of the estimation errors $x_t - \hat{x}_t = (x_t^{(1)} - \hat{x}_t^{(1)}, x_t^{(2)} - \hat{x}_t^{(2)}, \ldots, x_t^{(n_t)} - \hat{x}_t^{(n_t)})$ become equally important in selecting the active sensor set. To see this, consider for example that $\lambda_{\max}(\Theta_t) = \lambda_{\min}(\Theta_t) = \lambda$; then, the cost function in eq. (6) that Algorithm 1 minimizes to select the active sensor set becomes
$$\sum_{t=1}^T \text{tr}[\Theta_t\Sigma_{t|t}(\mathcal{S})] = \lambda \sum_{t=1}^T \mathbb{E}\left[\text{tr}(\|x_t - \hat{x}_t(\mathcal{S})\|_2^2)\right]$$
$$= \lambda \sum_{t=1}^T \sum_{i=1}^{n_t} \mathbb{E}\left[\text{tr}(\|x_t^{(i)} - \hat{x}_t^{(i)}(\mathcal{S})|_2^2)\right].$$
Overall, it is easier for Algorithm 1 to approximate a solution to Problem 1 as the cost function in eq. (6) becomes the cost function in the standard sensor selection problems where one minimizes the total estimation covariance as in eq. (10).

b) Case 2 where γ_g's bound in ineq. (13) *increases:* When either the numerators of the last two fractions in the right-hand-side of ineq. (13) increase or the denominators of the last two fractions in the right-hand-side of ineq. (13) decrease, then the right-hand-side in ineq. (13) increases. In particular, the numerators of the last two fractions in right-hand-side of ineq. (13) capture the estimation quality when all available sensors in $\mathcal{V}$ are used, via the terms of the form $\lambda_{\min}[\Sigma_{t|t}(\mathcal{V})]$ and $\lambda_{\min}[\bar{C}_{i,t}\Sigma_{t|t}(\mathcal{V})\bar{C}_{i,t}^\mathsf{T}]$. Interestingly, this suggests that the right-hand-side of ineq. (13) increases when the available sensors in $\mathcal{V}$ are inefficient in achieving low estimation error, that is, when the terms of the form $\lambda_{\min}[\Sigma_{t|t}(\mathcal{V})]$ and $\lambda_{\min}[\bar{C}_{i,t}\Sigma_{t|t}(\mathcal{V})\bar{C}_{i,t}^\mathsf{T}]$ increase. Similarly, the denominators of the last two fractions in right-hand-side of ineq. (13) capture the estimation quality when no sensors are used, via the terms of the form $\lambda_{\max}[\Sigma_{t|t}(\emptyset)]$ and $\lambda_{\max}[\bar{C}_{i,t}\Sigma_{t|t}(\emptyset)\bar{C}_{i,t}^\mathsf{T}]$. This suggests that the right-hand-side of ineq. (13) increases when the measurement noise increases.

We next give a control-level equivalent condition to Theorem 3's condition $\sum_{t=1}^T \Theta_t \succ 0$ for non-zero ratio γ_g.

Theorem 4 (Control-level condition for near-optimal sensor selection). *Consider the LQG problem where for any time $t =$*

[1]The inequality $h(\widehat{\mathcal{S}}, u_{1:T}(\widehat{\mathcal{S}})) \leq h(\emptyset, u_{1:T}(\emptyset))$ simply states that a control policy that is informed by the active sensor set $\mathcal{S}$ has better performance than a policy that does not use any sensor; for a more formal proof we refer the reader to Appendix B.

$1, 2, \ldots, T$*, the state* x_t *is known to each controller* u_t *and the process noise* w_t *is zero, i.e., the optimization problem*

$$\min_{u_{1:T}} \sum_{t=1}^{T} \left[\|x_{t+1}\|^2_{Q_t} + \|u_t(x_t)\|^2_{R_t}\right]\Big|_{\Sigma_{t|t}=W_t=0}. \quad (14)$$

Let A_t *to be invertible for all* $t = 1, 2, \ldots, T$*; the strict inequality* $\sum_{t=1}^{T} \Theta_t \succ 0$ *holds if and only if for all non-zero initial conditions* x_1*,*

$$0 \notin \arg\min_{u_{1:T}} \sum_{t=1}^{T} \left[\|x_{t+1}\|^2_{Q_t} + \|u_t(x_t)\|^2_{R_t}\right]\Big|_{\Sigma_{t|t}=W_t=0}.$$

Theorem 4 suggests that Theorem 3's sufficient condition $\sum_{t=1}^{T} \Theta_t \succ 0$ for non-zero ratio γ_g holds if and only if for any non-zero initial condition x_1 the all-zeroes control policy $u_{1:T} = (0, 0, \ldots, 0)$ is suboptimal for the noiseless perfect state-information LQG problem in eq. (14).

Overall, Algorithm 1 is the first scalable algorithm for Problem 1, and (for the LQG control problem instances of interest where a zero controller would result in a suboptimal behavior of the system and, as a result, LQG control design is necessary to achieve their desired system performance) it achieves close to optimal approximate performance.

V. Concluding Remarks

In this paper, we introduced the *sensing-constrained LQG control* Problem 1, which is central in modern control applications that range from large-scale networked systems to miniaturized robotics networks. While the computation of the optimal sensing strategy is intractable, We provided the first scalable algorithm for Problem 1, Algorithm 1, and under mild conditions on the system and LQG matrices, proved that Algorithm 1 computes a near-optimal sensing strategy with provable sub-optimality guarantees. To this end, we showed that a separation principle holds, which allows the design of sensing, estimation, and control policies in isolation. We motivated the importance of the sensing-constrained LQG Problem 1, and demonstrated the effectiveness of Algorithm 1, by considering two application scenarios: *sensing-constrained formation control*, and *resource-constrained robot navigation*.

References

[1] N. Elia and S. Mitter, "Stabilization of linear systems with limited information," *IEEE Trans. on Automatic Control*, vol. 46, no. 9, pp. 1384–1400, 2001.

[2] G. Nair and R. Evans, "Stabilizability of stochastic linear systems with finite feedback data rates," *SIAM Journal on Control and Optimization*, vol. 43, no. 2, pp. 413–436, 2004.

[3] S. Tatikonda and S. Mitter, "Control under communication constraints," *IEEE Trans. on Automatic Control*, vol. 49, no. 7, pp. 1056–1068, 2004.

[4] V. Borkar and S. Mitter, "LQG control with communication constraints," *Comm., Comp., Control, and Signal Processing*, pp. 365–373, 1997.

[5] J. L. Ny and G. Pappas, "Differentially private filtering," *IEEE Trans. on Automatic Control*, vol. 59, no. 2, pp. 341–354, 2014.

[6] G. Nair, F. Fagnani, S. Zampieri, and R. Evans, "Feedback control under data rate constraints: An overview," *Proceedings of the IEEE*, vol. 95, no. 1, pp. 108–137, 2007.

[7] J. Hespanha, P. Naghshtabrizi, and Y.Xu, "A survey of results in networked control systems," *Pr. of the IEEE*, vol. 95, no. 1, p. 138, 2007.

[8] J. Baillieul and P. Antsaklis, "Control and communication challenges in networked real-time systems," *Proceedings of the IEEE*, vol. 95, no. 1, pp. 9–28, 2007.

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

[10] H. Avron and C. Boutsidis, "Faster subset selection for matrices and applications," *SIAM Journal on Matrix Analysis and Applications*, vol. 34, no. 4, pp. 1464–1499, 2013.

[11] C. Li, S. Jegelka, and S. Sra, "Polynomial Time Algorithms for Dual Volume Sampling," *ArXiv e-prints: 1703.02674*, 2017.

[12] S. Joshi and S. Boyd, "Sensor selection via convex optimization," *IEEE Transactions on Signal Processing*, vol. 57, no. 2, pp. 451–462, 2009.

[13] J. L. Ny, E. Feron, and M. A. Dahleh, "Scheduling continuous-time kalman filters," *IEEE Trans. on Aut. Control*, vol. 56, no. 6, pp. 1381–1394, 2011.

[14] M. Shamaiah, S. Banerjee, and H. Vikalo, "Greedy sensor selection: Leveraging submodularity," in *Proceedings of the 49th IEEE Conference on Decision and control*, 2010, pp. 2572–2577.

[15] S. T. Jawaid and S. L. Smith, "Submodularity and greedy algorithms in sensor scheduling for linear dynamical systems," *Automatica*, vol. 61, pp. 282–288, 2015.

[16] V. Tzoumas, A. Jadbabaie, and G. J. Pappas, "Sensor placement for optimal Kalman filtering," in *Amer. Contr. Conf.*, 2016, pp. 191–196.

[17] E. Shafieepoorfard and M. Raginsky, "Rational inattention in scalar LQG control," in *Proceedings of the 52th IEEE Conference in Decision and Control*, 2013, pp. 5733–5739.

[18] T. Tanaka and H. Sandberg, "SDP-based joint sensor and controller design for information-regularized optimal LQG control," in *54th IEEE Conference on Decision and Control*, 2015, pp. 4486–4491.

[19] D. P. Bertsekas, *Dynamic programming and optimal control, Vol. I.* Athena Scientific, 2005.

[20] G. Nemhauser, L. Wolsey, and M. Fisher, "An analysis of approximations for maximizing submodular set functions – I," *Mathematical Programming*, vol. 14, no. 1, pp. 265–294, 1978.

[21] U. Feige, "A threshold of $ln(n)$ for approximating set cover," *Journal of the ACM*, vol. 45, no. 4, pp. 634–652, 1998.

[22] L. Carlone and S. Karaman, "Attention and anticipation in fast visual-inertial navigation," in *Proceeding of the IEEE International Conference on Robotics and Automation*, 2017, pp. 3886–3893.