

Deterministic and Randomized Actuator Scheduling With Guaranteed Performance Bounds

M. Siami, A. Olshevsky, and A. Jadbabaie

Abstract—In this paper, we investigate the problem of actuator selection for linear dynamical systems. We develop a framework to design a sparse actuator schedule for a given large-scale linear system with guaranteed performance bounds using deterministic polynomial-time and randomized approximately linear-time algorithms. First, we introduce systemic controllability metrics for linear dynamical systems that are monotone and homogeneous with respect to the controllability Gramian. We show that several popular and widely used optimization criteria in the literature belong to this class of controllability metrics. Our main result is to provide a polynomial-time actuator schedule that on average selects only a constant number of actuators at each time step, independent of the dimension, to furnish a guaranteed approximation of the controllability metrics in comparison to when all actuators are in use. Our results naturally apply to the dual problem of sensor selection, in which we provide a guaranteed approximation to the observability Gramian. We illustrate the effectiveness of our theoretical findings via several numerical simulations using benchmark examples.

I. INTRODUCTION

Over the past few years, controllability and observability properties of complex dynamical networks have been subjects of intense study in the controls community [1]–[12]. This interest stems from the need to steer or observe the state of large-scale, networked systems such as the power grids [13], social networks, biological and genetic regulatory networks [14]–[16], and traffic networks [17]. While the classical notion of controllability and observability, introduced by Kalman in [18] is quite well understood, the dependence of various measures of controllability or observability on number and location of sensors and actuators in linear systems have been subject of study for nearly 5 decades [19]. Often times, there is a need to steer or estimate the state of a large-scale, networked control system with as few actuators/sensors as possible, due to issues related to cost and energy depletion. The desire to perform control/estimation using a sparse set of actuators/sensors spans various application domains, ranging from infrastructure networks (e.g., water and power networks) to multi-robot systems and the study of the human connectome. For example, energy conservation through efficient utilization of sensors and actuators can help extend the duration of battery life in networks of mobile sensors and multi-agent robotic networks; estimating the whole state of the power grid using fewer measurement units will help reduce the cost of monitoring the network for systemic failures, etc.

It is therefore desirable to have a limited number of sensors and actuators without compromising the control or estimation performance too much. Unfortunately, as the recent works in [1], [20] have shown, the problem of finding a sparse set of input variables such that the resulting system is controllable is NP-hard. Even the presumably easier problem of approximating the minimum number better than a constant multiplicative factor of $\log n$ is also NP-hard.

This research was supported in part by a Vannevar Bush Fellowship from the Office of Secretary of Defense.

M. Siami is with the Department of Electrical and Computer Engineering, Northeastern University, Boston, MA 02115 USA (e-mail: m.siami@northeastern.edu). A. Olshevsky is with the Department of Electrical and Computer Engineering, Boston University, Boston, MA 02215 USA (e-mail: alexols@bu.edu). A. Jadbabaie is with the Institute for Data, Systems, and Society, Massachusetts Institute of Technology, Cambridge, MA 02139 USA (e-mail: jadbabai@mit.edu).

Other results in the literature have studied network controllability by exploring approximation algorithms for the closely related subset selection problem [1], [11], [12]. More recently, some of the authors showed that even the problem of finding a sparse set of actuators to guarantee reachability of a particular state is hard and even hard to approximate [21].

Previous studies have been mainly focused on solving the optimal sensor/actuator placement problem using the greedy heuristic, as approximations of the corresponding sparse-subset selection problem. While these results attempt to find approximation algorithms for finding the best sparse subset, our focus in this paper is to gain new fundamental insights into approximating various controllability metrics compared to the case when all possible actuators are chosen. Specifically, we are interested in actuator/sensor schedules that select a small number of actuators/sensors so as to save the energy while ensuring a suitable level of controllability (observability) performance for the entire network. Due to energy efficiency, we may want to reduce the number of active actuator/sensors at each time. At the same time, we would like to have a performance that closely resembles that of the original system, when all available sensor/actuators are active.

We investigate sparse sensor and actuator selection as particular instances where discrete geometric structures can be utilized to study network controllability and observability problems (cf. [20], [22], [23]). A key observation is the close connection between this problem and some classical problems in statistics such as outlier detection, active learning, and optimal experimental design. In recent years, there has been a renewed interest in optimal experiment design which has a long history going back at least 65 years [24], [25].

One of our main contributions is to show that the time-varying actuator selection problem, which goes back to a paper by Athans in 1972 [19], can be solved via random sampling. We propose an alternative to submodularity-based methods and instead use recent advances in theoretical computer science to develop scalable algorithms for sparsifying control inputs. Current approaches based on polynomial time relaxations of the subset selection problem require an extra multiplicative factor of $\log n$ sensors/actuators times the minimal number in order to just maintain controllability/observability. Using these recent advances [25]–[31], we show that by carefully designing a scheduling strategy, one can choose on average a constant number of sensors and actuators at each time, to approximate the controllability/observability metrics of the system when all sensors and actuators are in use.

Potential application domains. One potential application can be considered as wide-area oscillation damping control using High Voltage DC (HVDC) lines (e.g., [32]–[34]). HVDC systems are increasingly being installed in power grids all around the globe. This trend is expected to continue with recent advancements in power electronics technology, energy harvesting, and usage of renewable energy [32]. In this setup, it seems quite compelling to examine approaches that can support sparse HVDC lines (*i.e.*, actuators) scheduling to improve the controllability of the power grid in order to account for issues related to cost, energy depletion, and the limitations in directly accessing actuators, especially in large networks (cf. Example 2 in Section VII). Moreover, note that the dual problem of

actuator scheduling for control is sensor scheduling for estimation. In the present case, our sparse sensor schedule setup is equal to reducing the number of measurements for data reduction, and the observability Gramian-based measures show how well one can estimate the state of the system [35].

Another potential application is disease spread estimation in networks where testing resources are scarce. There are several models for the spread of infections (see [36] and references therein). Formally, let us consider a network with n nodes. Each node represents a city and has a non-negative scalar state $x_i(k)$ associated with it, which indicates the prevalence of an infection in that node. Since $x_i(t)$ is the *proportion* of the population at node i infected at time t , we assume that the $x_i(t)$ are close to zero (for example, this is valid for the recent COVID-19 pandemic; even though there are a substantial number of infected people still the proportion of the population that is infected is small as of early-April, 2020¹). It therefore makes sense to linearize the epidemic models around the zero state.

After linearization, the states evolve according to an autonomous linear differential equation $x(k+1) = Ax(k)$; in all epidemic models, the off-diagonal entries a_{ij} of the state matrix A indicates the unitized transmission rate of the infection from city j to city i , while the diagonal entries can be positive or negative, reflecting the possibility of either local spread or recovery.

We assume $y(t) = C(t)x + w(t)$, where $C(t)$ is a ‘‘subset’’ of the identity matrix (because one can measure the prevalence of the infection in a node by randomly testing from the population at that node) and $w(t)$ is noise. The sensor scheduling problem in this context amounts estimating the state with as small a variance as possible, while the measurements have to be done over a certain time-horizon and are bounded in number due to scarce resources.²

Some of our results appeared earlier in the conference version of this paper [37], [38]; however, their proofs are presented here for the first time. The manuscript also contains several new results, remarks, numerical examples, and proofs.

II. PRELIMINARIES AND DEFINITIONS

A. Mathematical Notations

Throughout the paper, discrete time index is denoted by k . The sets of real (integer), non-negative real (integer), and positive real (integer) numbers are represented by \mathbb{R} (\mathbb{Z}), \mathbb{R}_+ (\mathbb{Z}_+) and \mathbb{R}_{++} (\mathbb{Z}_{++}), respectively. The set of natural numbers $\{i \in \mathbb{Z}_{++} : i \leq n\}$ is denoted by $[n]$. The cardinality of a set σ is denoted by $\text{card}(\sigma)$. Capital letters, such as A or B , stand for real-valued matrices. For a square matrix X , $\det(X)$ and $\text{Trace}(X)$ refer to the determinant and the summation of on-diagonal elements of X , respectively. \mathbb{S}_+^n is the positive definite cone of n -by- n matrices. The n -by- n identity matrix is denoted by I . Notation $A \preceq B$ is equivalent to matrix $B - A$ being positive semi-definite. The transpose of matrix A is denoted by A^\top . The rank, kernel and image of matrix A are referred to by $\text{rank}(A)$, $\ker(A)$ and $\text{Im}(A)$, respectively. The Moore-Penrose pseudo-inverse of matrix A is denoted by A^\dagger . The ceiling function of $x \in \mathbb{R}$ is denoted by $\lceil x \rceil$ where it returns the least integer greater than or equal to x .

B. Linear Systems and Controllability

We start with the canonical linear discrete-time, time-invariant dynamics

$$x(k+1) = Ax(k) + Bu(k),$$

where $A \in \mathbb{R}^{n \times n}$, $B \in \mathbb{R}^{n \times m}$ and $k \in \mathbb{Z}_+$. The state matrix A describes the underlying structure of the system and the interaction strength between the agents, and matrix B represents how the control input enters the system. Equivalently, the dynamics can be written as

$$x(k+1) = Ax(k) + \sum_{i \in [m]} b_i u_i(k), \quad (1)$$

where b_i 's are columns of matrix $B \in \mathbb{R}^{n \times m}$. Then, the controllability matrix at time t is given by

$$\mathcal{C}(t) = [B \ AB \ A^2B \ \dots \ A^{t-1}B]. \quad (2)$$

In this paper, we assume that $t > 0$ is the time horizon to control (also known as the time-to-control). It is well-known that from a numerical standpoint it is better to characterize controllability in terms of the Gramian matrix at time t defined as follows:

$$\mathcal{W}(t) = \sum_{i=0}^{t-1} A^i B B^\top (A^i)^\top = \mathcal{C}(t) \mathcal{C}^\top(t). \quad (3)$$

When looking at time-varying input/actuator schedules, we will consider the following linear system with time-varying input matrix $\mathcal{B}(\cdot)$

$$x(k+1) = Ax(k) + \mathcal{B}(k)u(k). \quad (4)$$

For the above system, the controllability and Gramian matrices at time step t are defined as

$$\mathcal{C}_*(t) = [\mathcal{B}(t-1) \ A\mathcal{B}(t-2) \ A^2\mathcal{B}(t-3) \ \dots \ A^{t-1}\mathcal{B}(0)],$$

and

$$\begin{aligned} \mathcal{W}_*(t) &= \sum_{i=0}^{t-1} A^i \mathcal{B}(t-i-1) \mathcal{B}^\top(t-i-1) (A^i)^\top \\ &= \mathcal{C}_*(t) \mathcal{C}_*^\top(t), \end{aligned} \quad (5)$$

respectively.

Assumption 1: Throughout the paper, we assume that the system (1) is controllable (i.e., the controllability matrix has full row rank and the Gramian is positive definite). However, all results presented in this paper can be modified/extended to uncontrollable systems.

C. Matrix Reconstruction and Sparsification

The key idea through out the paper is to approximate the time- t controllability Gramian as a sparse sum of rank-1 matrices, while controlling the approximation error. To this end, we present a key lemma from the sparsification literature and state the necessary modification. We then use this result later in our deterministic algorithm to find a sparse actuator schedule.

Lemma 1 (Dual Set Spectral Sparsification [39]): Let $V = \{v_1, \dots, v_t\}$ and $U = \{u_1, \dots, u_t\}$ be two equal cardinality decompositions of identity matrices (i.e., $\sum_{i=1}^t v_i v_i^\top = I_n$ and $\sum_{i=1}^t u_i u_i^\top = I_\ell$) where $v_i \in \mathbb{R}^n$ ($n < t$) and $u_i \in \mathbb{R}^\ell$ ($\ell \leq t$). Given an integer κ with $n < \kappa \leq t$, Algorithm 1 computes a set of weights $c_i \geq 0$ where $i \in [t]$, such that

$$\lambda_{\min} \left(\sum_{i=1}^t c_i v_i v_i^\top \right) \geq \left(1 - \sqrt{\frac{n}{\kappa}} \right)^2,$$

$$\lambda_{\max} \left(\sum_{i=1}^t c_i u_i u_i^\top \right) \leq \left(1 + \sqrt{\frac{\ell}{\kappa}} \right)^2,$$

and

$$\text{card} \{i : c_i > 0, i \in [t]\} \leq \kappa.$$

Due to space limitations, we refer the interested readers to [39] for more details on Algorithm 1. However, roughly speaking, Algorithm

¹<https://www.cdc.gov/coronavirus/2019-ncov/cases-updates/cases-in-us.html>

²From now on we will focus the paper on the actuator selection problem. The dual notion of sensor selection follows similar ideas.

1 is based on choosing vectors in a greedy fashion that satisfy a set of desired properties at each step, leading to bounds on eigenvalues. In Algorithm 1, lower and upper barriers or potential functions are defined as follows:

$$\underline{\phi}(\underline{\mu}, \underline{\mathcal{A}}) = \sum_{i=1}^n \frac{1}{\lambda_i(\underline{\mathcal{A}}) - \underline{\mu}}, \quad (6)$$

and

$$\bar{\phi}(\bar{\mu}, \bar{\mathcal{A}}) = \sum_{i=1}^{\ell} \frac{1}{\bar{\mu} - \lambda_i(\bar{\mathcal{A}})}, \quad (7)$$

respectively. These potential functions quantify how far the eigenvalues of $\underline{\mathcal{A}}$ and $\bar{\mathcal{A}}$ are from the barriers $\underline{\mu}$ and $\bar{\mu}$. These potential functions become unbounded as any eigenvalue nears the barriers.³ We control the maximum eigenvalue of $\bar{\mathcal{A}}$ using an upper barrier $\bar{\mu}$ and the minimum eigenvalue of $\underline{\mathcal{A}}$ using a lower barrier $\underline{\mu}$. Two parameters \mathfrak{L} and \mathfrak{U} are defined as follows:

$$\mathfrak{L}(v, \underline{\delta}, \underline{\mathcal{A}}, \underline{\mu}) = \frac{v^\top (\underline{\mathcal{A}} - (\underline{\mu} + \underline{\delta})I_n)^{-2} v}{\underline{\phi}(\underline{\mu} + \underline{\delta}, \underline{\mathcal{A}}) - \underline{\phi}(\underline{\mu}, \underline{\mathcal{A}})} - v^\top (\underline{\mathcal{A}} - (\underline{\mu} + \underline{\delta})I_n)^{-1} v,$$

and

$$\mathfrak{U}(u, \bar{\delta}, \bar{\mathcal{A}}, \bar{\mu}) = \frac{u^\top ((\bar{\mu} + \bar{\delta})I_\ell - \bar{\mathcal{A}})^{-2} u}{\bar{\phi}(\bar{\mu}, \bar{\mathcal{A}}) - \bar{\phi}(\bar{\mu} + \bar{\delta}, \bar{\mathcal{A}})} + u^\top ((\bar{\mu} + \bar{\delta})I_\ell - \bar{\mathcal{A}})^{-1} u.$$

The Sherman-Morrison-Woodbury formula inspires the structure of the above quantities for more details on the barrier method (cf. [41, Section 1.2]). These potential functions (6) and (7) are chosen to guide the selection of vectors and scalings at each timestep τ and to ensure steady progress of the algorithm. Small values of these potentials indicate that the eigenvalues of $\bar{\mathcal{A}}$ and $\underline{\mathcal{A}}$ do not concentrate near $\bar{\mu}$ and $\underline{\mu}$, respectively. In Algorithm 1, at each iteration, we increase the upper barrier $\bar{\mu}$ by a fixed constant $\bar{\delta}$ and the lower barrier $\underline{\mu}$ by another fixed constant $\underline{\delta}$. It can be shown that as long as the potentials remain bounded, there must exist (at every step τ) a choice of an index j and weight c_j so that the addition of associated rank-1 matrices to $\bar{\mathcal{A}}$ and $\underline{\mathcal{A}}$, and the increments of barriers do not increase either potential and keep all the eigenvalues of the updated matrix between the barriers (see Algorithm 1). Repeating these steps ensures steady growth of all the eigenvalues and yields the desired result.

This algorithm is a generalization of an algorithm from [28] which is deterministic and at most needs $\mathcal{O}(\kappa t(n^2 + \ell^2))$. Furthermore, the algorithm needs $\mathcal{O}(\kappa t n^2)$ operations if U contains the standard basis of \mathbb{R}^t ; we refer the reader to [39] for more details.

Remark 1: We modify the 5th line of Algorithm 1; at each step, we choose an index j that maximizes

$$\mathfrak{L}(v_j, \underline{\delta}, \underline{\mathcal{A}}(\tau), \underline{\mu}(\tau)) - \mathfrak{U}(u_j, \bar{\delta}, \bar{\mathcal{A}}(\tau), \bar{\mu}(\tau)), \quad (8)$$

instead of only finding an index j such that

$$\mathfrak{U}(u_j, \bar{\delta}, \bar{\mathcal{A}}(\tau), \bar{\mu}(\tau)) \leq \mathfrak{L}(v_j, \underline{\delta}, \underline{\mathcal{A}}(\tau), \underline{\mu}(\tau)). \quad (9)$$

We should note that if an index j maximizes (8), then it will satisfy (9). Therefore, Lemma 1 still holds for the modified algorithm, and hence the theoretical bounds are valid. Based on our simulations, we observe that this modification can help to improve Algorithm 1 by producing smaller ratio $\lambda_{\max}(\sum_{i=1}^t c_i v_i v_i^\top) / \lambda_{\min}(\sum_{i=1}^t c_i v_i v_i^\top)$ (in Section V, we will see that this quantity is closely related to approximation factor ϵ). We denote the application of the algorithm

³These potentials are equal to constant multiples of the Stieltjes transform of $\underline{\mathcal{A}}$ and $\bar{\mathcal{A}}$ evaluated at $\underline{\mu}$ and $\bar{\mu}$, respectively [40].

Algorithm 1: A Deterministic Dual Set Spectral Sparsification
DualSet(V, U, κ).

Input : $V = [v_1, \dots, v_\ell]$, with $VV^\top = I_n$
 $U = [u_1, \dots, u_t]$, with $UU^\top = I_\ell$
 $\kappa \in \mathbb{Z}_+$, with $n < \kappa \leq t$
Output: $c = [c_1, c_2, \dots, c_t] \in \mathbb{R}_+^{1 \times t}$ with $\|c\|_0 \leq \kappa$

1 Set $c(0) = 0_{t \times 1}$, $\underline{\mathcal{A}}(0) = 0_{n \times n}$, $\bar{\mathcal{A}}(0) = 0_{\ell \times \ell}$, $\underline{\delta} = 1$,

$$\bar{\delta} = \frac{1 + \sqrt{\frac{\ell}{\kappa}}}{1 - \sqrt{\frac{\ell}{\kappa}}}$$

2 **for** $\tau = 0 : \kappa - 1$ **do**

3 $\underline{\mu}(\tau) = \tau - \sqrt{\kappa n}$

4 $\bar{\mu}(\tau) = \bar{\delta}(\tau + \sqrt{\kappa \ell})$

5 Find an index j such that

$$\mathfrak{U}(u_j, \bar{\delta}, \bar{\mathcal{A}}(\tau), \bar{\mu}(\tau)) \leq \mathfrak{L}(v_j, \underline{\delta}, \underline{\mathcal{A}}(\tau), \underline{\mu}(\tau))$$

6 Set $\Delta = 2(\mathfrak{U}(u_j, \bar{\delta}, \bar{\mathcal{A}}(\tau), \bar{\mu}(\tau)) + \mathfrak{L}(v_j, \underline{\delta}, \underline{\mathcal{A}}(\tau), \underline{\mu}(\tau)))^{-1}$

7 Update the j -th component of $c(\tau)$:

$$c(\tau + 1) = c(\tau) + \Delta e_j,$$

8 $\underline{\mathcal{A}}(\tau + 1) = \underline{\mathcal{A}}(\tau) + \Delta v_j v_j^\top$

9 $\bar{\mathcal{A}}(\tau + 1) = \bar{\mathcal{A}}(\tau) + \Delta u_j u_j^\top$

10 **end**

11 **return** $c = \kappa^{-1} (1 - \sqrt{\frac{\ell}{\kappa}}) c(\kappa)$

to V and U by

$$[c_1, c_2, \dots, c_t] = \text{DualSet}^*(V, U, \kappa).$$

We now recall the concentration lemma of Rudelson-Vershynin [29] as follows. We are going to use this result in the proof of Theorem 2.

Lemma 2: [29, Thm. 3.1] Let $y \in \mathbb{R}^p$ be a random vector such that $\|y\| \leq b$ almost surely and $\|\mathbf{E} y y^\top\|_2 \leq 1$. Let y_1, \dots, y_n be i.i.d. copies of y . Then

$$\mathbf{E} \left\| \frac{1}{n} \sum_{i=1}^n y_i y_i^\top - \mathbf{E} y y^\top \right\|_2 \leq \min \left(1, c b \sqrt{\frac{\log n}{n}} \right), \quad (10)$$

where $c > 0$ is some universal constant.

In the next section, we show how various controllability measures can be approximated by selecting a sparse set of actuators via the above algorithm.

III. SYSTEMIC CONTROLLABILITY METRICS

Similar to the *systemic* notions introduced in [42]–[44], we define various controllability metrics. These measures are real-valued operators defined on the set of all linear dynamical systems governed by (4) and quantify various measures of the required control energy. All of the metrics depend on the controllability Gramian matrix of the system which is a positive definite matrix. Therefore, one can define a systemic controllability performance measure as an operator on the set of Gramian matrices of all controllable systems with n states which we represent by \mathbb{S}_+^n .⁴

Definition 1 (Systemic Criteria): A controllability metric $\rho : \mathbb{S}_+^n \rightarrow \mathbb{R}$ is systemic if and only if

⁴For any $X \in \mathbb{S}_+^n$ and given $t \in \mathbb{Z}_{++}$, there exists at least one controllable system with $\mathcal{W}(t) = X$ (e.g., $x(k+1) = X^{\frac{1}{2}} u(k)$), and for any controllable system, it is well known that the Gramian matrix is positive definite (see (5)). Therefore, the set of Gramian matrices of all controllable systems with n states is equal to \mathbb{S}_+^n .

Optimality-criteria	Systemic Controllability Measure	Matrix Operator Form
A-optimality	Average control energy	Trace ($\mathcal{W}^{-1}(t)$)
D-optimality	The volume of the ellipsoid	$(\det \mathcal{W}(t))^{-1/n}$
T-optimality	Inverse of the trace	$1/\text{Trace}(\mathcal{W}(t))$
E-optimality	Inverse of the minimum eigenvalue	$1/\lambda_{\min}(\mathcal{W}(t))$

TABLE I: Some important examples of systemic controllability metrics.

1. *Homogeneity:* For all $\kappa > 1$,

$$\rho(\kappa A) = \kappa^{-1} \rho(A);$$

2. *Monotonicity:* If $B \preceq A$, then

$$\rho(A) \leq \rho(B).$$

For many popular choices of ρ , one can see that they satisfy the properties presented in Definition 1. Some of them are listed in Table I. We note that similar criteria have been developed [24], [25], [45] in the experiment design literature (cf. Table I). In what follows, we will make this statement formal.

Proposition 1: For given dynamics (4) with Gramian matrix $\mathcal{W}(t)$, the metrics presented in Table 1 are systemic controllability measures.

Proof: One can easily see that all these measures satisfy the homogeneity, and monotonicity properties in Definition 1 (cf. [44], [46]). ■

In the next section, we show how various measures can be approximated by selecting a sparse set of actuators.

IV. SPARSE ACTUATOR SELECTION PROBLEMS

For a given linear system (1) with a general underlying structure, the actuator scheduling problem seeks to construct a schedule of the control inputs that keeps the number of active actuators much less than the original system such that the controllability matrices of the original and the new systems are similar in an appropriately defined sense. Specifically, given a canonical linear, time-invariant system (1) with m actuators and controllability Gramian matrix $\mathcal{W}(t)$ at time t , our goal is to find a sparse actuator schedule such that the resulting system with controllability Gramian $\mathcal{W}_s(t)$ is well-approximated, i.e.,

$$\left| \frac{\rho(\mathcal{W}(t)) - \rho(\mathcal{W}_s(t))}{\rho(\mathcal{W}(t))} \right| \leq \epsilon, \quad (11)$$

where ρ is any systemic controllability metric that quantifies the difficulty of the control problem for example as a function of the required control energy, and $\epsilon \geq 0$ is the approximation factor. The systemic controllability metrics are defined based on the controllability Gramian, therefore “close” Gramian matrices result in approximately the same values. Our goal here is to answer the following questions:

- What is the minimum number of actuators to be chosen to achieve a good approximation of the system with the full set of actuators utilized?
- What is the relation between the number of selected actuators and performance/controllability loss?
- Does a sparse approximation schedule exist with at most a constant number of active actuators at each time?
- What is the time complexity of choosing the subset of actuators with guaranteed performance bounds?

In the rest of this paper, we show how some fairly recent advances in theoretical computer science and the probabilistic method can be utilized to answer these questions. The probabilistic method is one of the most important tools of modern combinatorics which was introduced by Erdős. The idea is that a deterministic solution

is shown to exist by constructing a random candidate satisfying all the requirements of the problem with positive probability. Recently, Marcus, Spielman, and Srivastava introduced a new variant of the probabilistic method which ends up solving the so-called Kadison-Singer (KS) conjecture [30]. We use the solution approach to the KS conjecture together with a combination of tools from Sections V to find a sparse approximation of the actuator selection problem with algorithms that have favorable time-complexity.

V. A WEIGHTED SPARSE ACTUATOR SCHEDULE

As a starting point, we allow for scaling of the input signals at chosen inputs while keeping the input scaling bounded. The input scaling allows for an extra degree of freedom that could allow for choosing a sparser set of inputs. Given (1), we define a weighted actuator schedule by $\sigma = \{\sigma_k\}_{k=0}^{t-1}$ and scalings $s_i(k) \geq 0$ where $i \in [m]$, $k+1 \in [t]$, and $\sigma_k = \{i | s_i(k) > 0\} \subseteq [m]$. The resulting system with this schedule is

$$x(k+1) = Ax(k) + \sum_{i \in \sigma_k} s_i(k) b_i u_i(k), \quad k \in \mathbb{Z}_+ \quad (12)$$

where $s_i(k) \geq 0$ shows the strength of the i -th control input at time k . The controllability Gramian (5) at time t for this system can be rewritten as

$$\mathcal{W}_s(t) = \sum_{k=0}^{t-1} \sum_{j \in \sigma_k} s_j^2(k) \left(A^{t-k-1} b_j \right) \left(A^{t-k-1} b_j \right)^\top. \quad (13)$$

Our goal is to reduce the number of active actuators *on average* d , where

$$d := \frac{\sum_{k=0}^{t-1} \text{card}\{\sigma_k\}}{t}, \quad (14)$$

such that the controllability Gramian of the fully actuated and the new sparsely actuated system are “close.” Of course, this approximation will require horizon lengths that are potentially longer than the dimension of the state. The definition below formalizes this approximation.

Definition 2: Given a time horizon $t \geq n$, system (12) with a weighted actuator schedule is (ϵ, d) -approximation of system (1) if and only if

$$(1 - \epsilon) \mathcal{W}(t) \preceq \mathcal{W}_s(t) \preceq (1 + \epsilon) \mathcal{W}(t), \quad (15)$$

where $\mathcal{W}(t)$ and $\mathcal{W}_s(t)$ are the controllability Gramian matrices of (1) and (12), respectively, and parameter d is defined by (14) as the average number of active actuators, and $\epsilon \in (0, 1)$ is the approximation factor.

Remark 2: While it might appear that allowing for the choice of $s_i(k)$ might lead to amplification of input signals, we note that the scaling cannot be too large because the approximation is two-sided. Specifically, by taking the trace from both sides of (15), we can see that the weighted summation of $s_i^2(k)$ ’s is bounded. Moreover, based on Definition 2, the ranks of matrices $\mathcal{W}(t)$ and $\mathcal{W}_s(t)$ are the same. Thus, the resulting (ϵ, d) -approximation remains controllable (recall that we assume that the original system is controllable).

Remark 3: The results presented in this paper also work for the case of linear time-varying systems, and it is straightforward to extend

Algorithm 2: A deterministic greedy-based algorithm to construct a sparse weighted actuator schedule (Theorem 1).

Input : $A \in \mathbb{R}^{n \times n}$, $B \in \mathbb{R}^{n \times m}$, t and d

Output: $s_i(k) \geq 0$ for $(i, k+1) \in [m] \times [t]$

- 1 $C(t) := [B \ AB \ A^2B \ \dots \ A^{t-1}B]$
- 2 Set $V = (C(t)C^\top(t))^{-\frac{1}{2}} C(t)$
- 3 Set $U = V$
- 4 Run $[c_1, \dots, c_{mt}] = \text{DualSet}^*(V, U, dt)$
- 5 **return** $s_i(k) := \sqrt{c_{i+mk}/(1 + \frac{n}{dt})}$ for $(i, k+1) \in [m] \times [t]$

Algorithm 3: A deterministic greedy-based algorithm to construct a sparse weighted actuator schedule (Corollary 1).

Input : $A \in \mathbb{R}^{n \times n}$, $B \in \mathbb{R}^{n \times m}$, t and d

Output: $s_i(k) \geq 0$ for $(i, k+1) \in [m] \times [t]$

- 1 $C(t) := [B \ AB \ A^2B \ \dots \ A^{t-1}B]$
- 2 Set $V = (C(t)C^\top(t))^{-\frac{1}{2}} C(t)$
- 3 Set

$$U = \begin{bmatrix} \underbrace{e_1, \dots, e_{mt}}_{=I_{mt}} \end{bmatrix}$$

// where $e_i \in \mathbb{R}^{mt}$ for $i \in [mt]$ are the standard basis vectors for \mathbb{R}^{mt}
- 4 Run $[c_1, \dots, c_{mt}] = \text{DualSet}^*(V, U, dt)$
- 5 **return** $s_i(k) := \sqrt{c_{i+mk}}$ for $(i, k+1) \in [m] \times [t]$

them for affine nonlinear discrete-time systems as well.

A. Deterministic Approach: Sparsifying Sums of Rank-one Matrices

The next theorem constructs a solution for the sparse weighted actuator schedule problem in polynomial time.

Theorem 1: Given the time horizon $t \geq n$, model (1), and $d > 1$, Algorithm 2 deterministically constructs an actuator schedule such that the resulting system (12) is a (ϵ, d) -approximation of (1) with $\epsilon = \frac{2}{\sqrt{\frac{dt}{n} + \sqrt{\frac{n}{dt}}}}$ in at most $\mathcal{O}(dm(tn)^2)$ operations.

Proof: The controllability Gramian of (1) at time t is given by

$$\begin{aligned} \mathcal{W}(t) &= \sum_{i=0}^{t-1} \sum_{j=1}^m \underbrace{(A^i b_j)(A^i b_j)^\top}_{v_{ij}} \\ &= \sum_{i=0}^{t-1} \sum_{j=1}^m v_{ij} v_{ij}^\top. \end{aligned} \quad (16)$$

By multiplying $\mathcal{W}^{-\frac{1}{2}}(t)$ on both sides of (16), it follows that

$$\begin{aligned} I &= \sum_{i=0}^{t-1} \sum_{j=1}^m \underbrace{(\mathcal{W}^{-\frac{1}{2}}(t) A^i b_j)(\mathcal{W}^{-\frac{1}{2}}(t) A^i b_j)^\top}_{\bar{v}_{ij}} \\ &= \sum_{i=0}^{t-1} \sum_{j=1}^m \bar{v}_{ij} \bar{v}_{ij}^\top. \end{aligned} \quad (17)$$

Next, we define $U := \{\bar{v}_{ij} | i+1 \in [t], j \in [m]\}$ and $V := U$. According to (3), (16), and (17), elements of U are the columns of matrix $(C(t)C^\top(t))^{-\frac{1}{2}} C(t)$. We now apply Lemma 1, which shows

that there exist scalars $\bar{c}_{ij} \geq 0$ with

$$\text{card} \{(i, j) : i+1 \in [t], j \in [m], \bar{c}_{ij} > 0\} \leq \frac{dt}{n} \times n, \quad (18)$$

such that

$$\left(1 - \sqrt{\frac{n}{dt}}\right)^2 I \preceq \sum_{i=0}^{t-1} \sum_{j=1}^m \bar{c}_{ij} \bar{v}_{ij} \bar{v}_{ij}^\top,$$

and

$$\sum_{i=0}^{t-1} \sum_{j=1}^m \bar{c}_{ij} \bar{v}_{ij} \bar{v}_{ij}^\top \preceq \left(1 + \sqrt{\frac{n}{dt}}\right)^2 I,$$

or equivalently,

$$\left(1 - \sqrt{\frac{n}{dt}}\right)^2 \mathcal{W}(t) \preceq \sum_{i=0}^{t-1} \sum_{j=1}^m \bar{c}_{ij} v_{ij} v_{ij}^\top \preceq \left(1 + \sqrt{\frac{n}{dt}}\right)^2 \mathcal{W}(t). \quad (19)$$

We can of course write the controllability Gramian of (12) at time t as

$$\begin{aligned} \mathcal{W}_s(t) &= \sum_{i=0}^{t-1} \sum_{j=1}^m s_j^2(t-i-1) \underbrace{(A^i b_j)(A^i b_j)^\top}_{v_{ij}} \\ &= \sum_{i=0}^{t-1} \sum_{j=1}^m s_j^2(t-i-1) v_{ij} v_{ij}^\top. \end{aligned}$$

Define $\epsilon := \frac{2}{\sqrt{\frac{dt}{n} + \sqrt{\frac{n}{dt}}}}$, and

$$s_j(t-i-1) := \sqrt{\bar{c}_{ij}/(1 + \frac{n}{dt})}. \quad (20)$$

Then, by substituting $(1 + \frac{n}{dt}) s_j^2(t-i-1)$ for \bar{c}_{ij} in (19), we get

$$(1 - \epsilon)\mathcal{W}(t) \preceq \mathcal{W}_s(t) \preceq (1 + \epsilon)\mathcal{W}(t). \quad (21)$$

Finally, using (21), (18), and Definition 2, we obtain the desired result. Moreover, this algorithm runs in dt iterations; In each iteration, the functions \mathcal{U} and \mathcal{L} are evaluated at most mt times. All mt evaluations for both functions need at most $\mathcal{O}(n^3 + mtn^2)$ time, because for all of them the matrix inversions and eigenvalue decompositions can be calculated once. Finally, the updating step needs an additional $\mathcal{O}(n^2)$ time. Overall, the complexity of the algorithm is of the order $\mathcal{O}(dm(tn)^2)$. ■

Remark 4: For a given $d \geq 1$, while choosing dt columns of the controllability matrix that form a full row rank matrix (i.e., the system is controllable) is an easy task but finding dt columns of the controllability matrix that approximate the full Gramian matrix is what we are interested in here. To do so, we should note that approximating the full Gramian matrix while keeping the number of active actuators less than a constant d at each time is not possible in general. For example, in the case that $A = \mathbf{0}_{n \times n}$ and $B = I_n$, at least all actuators at time $k = t-1$ are needed to form a full row rank matrix (or to approximate the full Gramian matrix). However, as we mentioned earlier, the number of active actuators on average can be kept constant in order to approximate the full Gramian matrix. Furthermore, condition $dt \geq n$ is needed for any algorithm that has a hope of success. Indeed, taking $B = I_n$ and $A = I_n$, it is straightforward to see that if $dt < n$, then we cannot hope to approximate the controllability Gramian because the controllability matrix of any schedule with d active actuators on average is not full rank.

Tradeoffs: Theorem 1 illustrates a tradeoff between the average number of active actuators d and the time horizon t (also known as the time-to-control). This implies that the reduction in the average number of active actuators comes at the expense of increasing time horizon t in order to get the same approximation factor ϵ . Moreover,

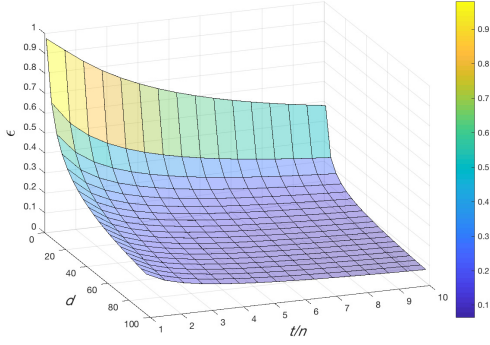


Fig. 1: This plot presents the approximation factor ϵ given by Theorem 1 versus the average number of active actuators $d \in (1, 100]$ and the normalized time horizon t/n .

the approximation becomes more accurate as t and d are increased. Of course, increasing d will require more active actuators and larger t requires a larger control time window.

Fig. 1 depicts the approximation ratio ϵ given by Theorem 1 versus the average number of active actuators d and the normalized time horizon t/n . We note that the approximation factor improves as t become larger than n . Moreover, because of $\frac{2}{x+\frac{1}{x}} \leq 1$ for $x > 0$, the approximation factor $\epsilon = \frac{2}{\sqrt{\frac{dt}{n}} + \sqrt{\frac{n}{dt}}}$ is always less than or equal to one. Hence, the upper bound ratio in (15) is at most two.

Sparse Actuator Schedules with Energy Constraints: In this subsection, based on the energy/budget constraints on the scalings $s_i(k)$'s where $i \in [m]$ and $k+1 \in [t]$; three cases are considered as follows

(i) the scaling ratios are bounded, i.e.,

$$\max_{i \in [m], k+1 \in [t]} s_i^2(k) \leq \gamma,$$

(ii) the sum of scaling ratios for each input is bounded, i.e.,

$$\max_{i \in [m]} \sum_{k+1 \in [t]} s_i^2(k) \leq \gamma,$$

(iii) the sum of scaling ratios at each time is bounded, i.e.,

$$\max_{k+1 \in [t]} \sum_{i \in [m]} s_i^2(k) \leq \gamma.$$

In the next corollaries, we present deterministic sparse actuator schedules with the above energy/budget constraints. These corollaries trade one of the inequalities in Theorem 1 with a single fixed bound on the size of scalings.

Corollary 1: Given the time horizon $t \geq n$, model (1), and $d > 1$, Algorithm 3 deterministically constructs an actuator schedule for (12) in at most $\mathcal{O}(dm(tn)^2)$ operations such that, on average, at most d active actuators are selected, and the following bound

$$\rho(\mathcal{W}_s(t)) \leq \left(1 - \sqrt{\frac{n}{dt}}\right)^{-2} \rho(\mathcal{W}(t))$$

holds for all systemic controllability measures. Moreover, the maximum scaling ratio over all time and inputs is bounded by

$$\max_{i \in [m], k+1 \in [t]} s_i^2(k) \leq \gamma,$$

where $\gamma = \left(1 + \sqrt{\frac{m}{d}}\right)^2$.

Proof: The proof is a simple variation on the proof of Theorem 1, and is not repeated here. ■

As expected, the above result shows that the scaling becomes

Algorithm 4: A deterministic greedy-based algorithm to construct a sparse weighted actuator schedule (Corollary 2).

Input : $A \in \mathbb{R}^{n \times n}$, $B \in \mathbb{R}^{n \times m}$, t and d

Output: $s_i(k) \geq 0$ for $(i, k+1) \in [m] \times [t]$

1 $C(t) := [B \ AB \ A^2B \ \dots \ A^{t-1}B]$

2 Set $V = (C(t)C^\top(t))^{-\frac{1}{2}} C(t)$

3 Set

$$U = \frac{1}{\sqrt{t}} \begin{bmatrix} \underbrace{e_1, \dots, e_m}_{=I_m}, \dots, \underbrace{e_1, \dots, e_m}_{=I_m} \end{bmatrix}$$

// where $e_i \in \mathbb{R}^m$ for $i \in [m]$ are the standard basis vectors for \mathbb{R}^m and $UU^\top = I_m$

4 Run $[c_1, \dots, c_{mt}] = \text{DualSet}^*(V, U, dt)$

5 **return** $s_i(k) := \sqrt{c_{i+mk}}$ for $(i, k+1) \in [m] \times [t]$

Algorithm 5: A deterministic greedy-based algorithm to construct a sparse weighted actuator schedule (Corollary 3).

Input : $A \in \mathbb{R}^{n \times n}$, $B \in \mathbb{R}^{n \times m}$, t and d

Output: $s_i(k) \geq 0$ for $(i, k+1) \in [m] \times [t]$

1 $C(t) := [B \ AB \ A^2B \ \dots \ A^{t-1}B]$

2 Set $V = (C(t)C^\top(t))^{-\frac{1}{2}} C(t)$

3 Set

$$U = \frac{1}{\sqrt{m}} \begin{bmatrix} \underbrace{e_1, \dots, e_1}_{m \text{ times}}, \dots, \underbrace{e_t, \dots, e_t}_{m \text{ times}} \end{bmatrix}$$

// where $e_i \in \mathbb{R}^t$ for $i \in [t]$ are the standard basis vectors for \mathbb{R}^t and $UU^\top = I_t$

4 Run $[c_1, \dots, c_{mt}] = \text{DualSet}^*(V, U, dt)$

5 **return** $s_i(k) := \sqrt{c_{i+mk}}$ for $(i, k+1) \in [m] \times [t]$

smaller as the ratio m/d decreases.

Corollary 2: Given the time horizon $t \geq n$, model (1), and $d > 1$, Algorithm 4 deterministically constructs an actuator schedule for (12) in $\mathcal{O}(dm(tn)^2)$ operations such that it has, on average, at most d active actuators, and the following

$$\rho(\mathcal{W}_s(t)) \leq \left(1 - \sqrt{\frac{n}{dt}}\right)^{-2} \rho(\mathcal{W}(t))$$

holds for all systemic controllability measures. Moreover, the sum of scaling ratios for all inputs is bounded by

$$\max_{i \in [m]} \sum_{k=0}^{t-1} s_i^2(k) \leq \gamma,$$

where $\gamma = t \left(1 + \sqrt{\frac{m}{d}}\right)^2$.

Proof: The proof is a simple variation on the proof of Theorem 1, and is not repeated here. ■

Corollary 3: Given the time horizon $t \geq n$, model (1), and $d > 1$, Algorithm 5 deterministically constructs an actuator schedule for (12) in $\mathcal{O}(dm(tn)^2)$ operations such that it has, on average, at most d active actuators, and the following

$$\rho(\mathcal{W}_s(t)) \leq \left(1 - \sqrt{\frac{n}{dt}}\right)^{-2} \rho(\mathcal{W}(t))$$

holds for all systemic controllability measures. Moreover, the sum of

scaling ratios at each time is bounded by

$$\max_{k+1 \in [t]} \sum_{i=1}^m s_i^2(k) \leq \gamma,$$

where $\gamma = m \left(1 + \sqrt{\frac{1}{d}}\right)^2$.

Proof: The proof is a simple variation on the proof of Theorem 1, and is not repeated here. ■

We use a different idea in Subsection V-B, to develop scalable algorithms that sparsify control inputs by employing a sub-sampling method for a time-varying actuator schedule. This however come at the cost of an extra log factor in terms of the average number of selected actuators.

B. Randomized Approach: Sampling Based on the Leverage Score

In this subsection, we focus on a computationally tractable method for the weighted sparse actuator scheduling problem that achieve near optimal solution.

Definition 3: The leverage score of the i -th column of matrix $P \in \mathbb{R}^{n \times m}$ is defined as

$$\ell_i = p_i^\top (PP^\top)^\dagger p_i,$$

where p_i is the i -th column of matrix P .

This quantity encodes the importance of the i -th column compared to the other columns. A larger leverage score shows that the corresponding column has more influence on the spectrum of P . Based on the leverage score definition, we get $\ell_i \in [0, 1]$ for all $i \in [m]$. Because ℓ_i 's are the diagonal elements of the projection matrix $P^\top (PP^\top)^{-1} P$ and the diagonal elements of the projection matrix are between zero and one. Leverage score $\ell_i = 1$ means that the i -th column has a component orthogonal to the rest of the columns. Therefore, eliminating that column will decrease the rank of matrix P . On the other hand, $\ell_i = 0$ means that the i -th column is parallel to the rest of the columns. When the corresponding matrix is the graph Laplacian, this quantity reduces to the effective resistance of each link in a graph [27].

We group the columns of $\mathcal{C}(t)$ in the following form

$$\mathcal{C}(t) = \left[\underbrace{[b_1 \quad Ab_1 \quad \dots \quad A^{t-1}b_1]}_{\mathcal{C}_1(t)} \quad \dots \quad \underbrace{[b_m \quad Ab_m \quad \dots \quad A^{t-1}b_m]}_{\mathcal{C}_m(t)} \right],$$

where b_j is the j -th column of matrix B . Matrix $\mathcal{C}_j(t)$ presents the controllability matrix of input j at time t . The leverage score for each column of $\mathcal{C}(t)$ is defined as

$$\ell(A^i b_j) = (A^i b_j)^\top \left(\mathcal{C}(t) \mathcal{C}^\top(t) \right)^\dagger A^i b_j, \quad (22)$$

where $(i+1) \in [t]$ and $j \in [m]$. For these scores, we have

$$\begin{aligned} \sum_{i=0}^{t-1} \sum_{j=1}^m \ell(A^i b_j) &= \text{Trace} \left(\mathcal{C}^\top(t) (\mathcal{C}(t) \mathcal{C}^\top(t))^\dagger \mathcal{C}(t) \right) \\ &= \text{Trace} \left((\mathcal{C}(t) \mathcal{C}^\top(t))^\dagger \mathcal{C}(t) \mathcal{C}^\top(t) \right) \\ &= \text{Trace}(I_n) = n. \end{aligned} \quad (23)$$

In (23), we use the fact that $\text{Trace}(AB) = \text{Trace}(BA)$ (i.e., the matrices in a trace of a product can be switched without changing the result as long as A and B^\top have the same dimensions), and $\text{rank}(\mathcal{C}(t)) = n$ (i.e., the system is controllable).

We now randomly sample the actuators with probabilities proportional to their leverage scores to sparsify control inputs. This sampling occurs across time and over all possible actuators at each time (see Algorithm 6). At every time, each actuator is kept active or inactive

Algorithm 6: A simple randomized algorithm to compute a sparse weighted actuator schedule $\{\sigma_i\}_{i=0}^{t-1}$ (Theorem 2).

Input : $A \in \mathbb{R}^{n \times n}$, $B \in \mathbb{R}^{n \times m}$, t and d

Output: $\{\sigma_i\}_{i=0}^{t-1}$ and $s_i(k-1)$ for $(i, k) \in [m] \times [t]$

```

1  $\mathcal{C}(t) := [B \quad AB \quad A^2B \quad \dots \quad A^{t-1}B]$ 
2 set  $\{\sigma_i\}_{i=0}^{t-1}$  to be the empty sets (i.e.  $\sigma_i := \{\}$ )
3 set  $s_i(k-1) = 0$  for  $(i, k) \in [m] \times [t]$ 
4 set  $\pi(i, k) = \frac{\text{Trace}((\mathcal{C}(t)\mathcal{C}^\top(t))^\dagger A^{t-k} b_i (A^{t-k} b_i)^\top)}{n}$  for all
    $(i, k) \in [m] \times [t]$ 
5 for  $j = 1$  to  $M := \lceil dt \rceil$  do
6    $(i, k) \leftarrow$  sample  $(i, k)$  from  $[m] \times [t]$  with probability
   distribution  $\pi$ 
7    $\sigma_{k-1} = \sigma_{k-1} \cup \{i\}$ 
8    $s_i^2(k-1) = s_i^2(k-1) + \frac{1}{M\pi(i, k)}$ 
9 end
10 return  $\{\sigma_i\}_{i=0}^{t-1}$  and  $s_i(k-1)$  for  $(i, k) \in [m] \times [t]$ 

```

according to probability $\ell(A^i b_j)/n$ where $(i+1) \in [t]$ and $j \in [m]$. Using [27, Thm. 1], we can construct a sampling strategy that utilizes the leverage score to probabilistically choose actuators. The catch is that there is an extra $\log n$ factor in the average number of selected actuators, and potentially different actuators are chosen at different times.

Theorem 2: Assume that dynamics (1), time horizon $t \geq n$, and approximation factor $\epsilon \in [1/\sqrt{n}, 1)$ are given. Choose a real number $d = \frac{9c^2 n \log n}{\epsilon^2}$, where c is the constant in Lemma 2. Then, Algorithm 6 produces scheduling (12) which is (ϵ, d) -approximation of (1) with probability of at least 0.5 for sufficiently large n .⁵

Proof: The structure of the proof follows from the proof of [27, Thm. 4]. Let us start with the following projection matrix

$$\Pi = \mathcal{C}(t)^\top \mathcal{W}^{-1}(t) \mathcal{C}(t), \quad (24)$$

where $\mathcal{C}(t)$ is n -by- tm controllability matrix (2) and matrix $\mathcal{W}(t) = \mathcal{C}(t) \mathcal{C}^\top(t)$ is given by (3). The tm -by- tm projection matrix Π has eigenvalue at 0 with multiplicity $t \times m - n$ and eigenvalue at 1 with multiplicity n . Therefore, we get

$$\text{Trace}(\Pi) = \text{rank}(\Pi) = n. \quad (25)$$

The set X is obtained based on columns of Π as follows

$$X = \{y_j \in \mathbb{R}^n : y_j = (\pi_j)^{-1/2} \Pi(\cdot, j), \text{ and } j \in [tm]\},$$

where matrix Π is given by (24), vector $\Pi(\cdot, j)$ is the j -th column of Π , and π_j is probability of selecting vector y_j (i.e., $\pi(y_j) = \pi_j$). The probability distribution π over X is defined by

$$\pi(y_j) := \pi_j = \frac{\Pi(j, j)}{\text{rank}(\Pi)} = \frac{\Pi(j, j)}{n}, \quad (26)$$

where $\Pi(j, j)$ is the j -th diagonal element of matrix Π and $j \in [tm]$. Based on (24), each columns of Π corresponds to the i -th input at time $k-1$ where $(i, k) \in [m] \times [t]$. The mapping comes from the

⁵We should note that one can repeat Algorithm 6, for example $c = 4$ times to get the desired results with more than $1 - 1/c^4 = 0.9375$ probability. Moreover, the probability is improved by increasing the number of iteration c . Assume c is a constant and does not depend on n, t, d or m . Then, repeating the algorithm c times does not change the time complexity of the algorithm, and we still have an approximately linear time algorithm. Therefore, by repeating the algorithm and choosing the best result, we can obtain the same error bound with higher probability.

controllability matrix structure (2), and we define it as $\mathbf{m}(\cdot) : [mt] \rightarrow [m] \times [t]$, where

$$\mathbf{m}(j) = (i, k) = (j - t \lfloor \frac{j}{t} \rfloor, t - \lfloor \frac{j}{m} \rfloor). \quad (27)$$

This means the j -th column of Π corresponds to (i, k) , where $k = t - \lfloor \frac{j}{m} \rfloor$ and $i = j - t \lfloor \frac{j}{t} \rfloor$. Thus, in Algorithm 6, for notational simplicity we denote $\pi(y_j) := \pi(\mathbf{m}(j)) = \pi(i, k)$. For each element of X , we have

$$\begin{aligned} \|y_j\| &= (\pi_j)^{-1/2} \|\Pi(\cdot, j)\| = \left(\frac{n}{\Pi(j, j)} \right)^{1/2} \times (\Pi(j, j))^{1/2} \\ &= \sqrt{n}. \end{aligned} \quad (28)$$

where, we use the fact that

$$\Pi(\cdot, j)^\top \Pi(\cdot, j) = \Pi(j, j),$$

because Π is an orthogonal projection matrix (*i.e.*, $\Pi\Pi = \Pi$). Then, using (25) and (28), we have

$$\begin{aligned} \mathbf{E}(yy^\top) &= \sum_{j=1}^{tm} \pi_j y_j y_j^\top = \sum_{j=1}^{tm} \pi_j \frac{1}{\pi_j} \Pi(\cdot, j) \Pi(\cdot, j)^\top \\ &= \Pi\Pi = \Pi, \end{aligned} \quad (29)$$

where y is a random variable vector with a countable set of outcomes X occurring with probabilities π defined by (26). Let $\hat{y}_1, \dots, \hat{y}_M$ be independent samples drawn from π , then, based on Algorithm 6, then we have

$$\begin{aligned} \Pi\Pi\Pi &= \sum_{j=1}^{tm} \Gamma(j, j) \Pi(\cdot, j) \Pi(\cdot, j)^\top \\ &= \sum_{j=1}^{tm} s_i^2(k-1) \Pi(\cdot, j) \Pi(\cdot, j)^\top \\ &= \sum_{j=1}^{tm} \frac{\# \text{ of times } (i, k) \text{ is sampled}}{M\pi(i, k)} \Pi(\cdot, j) \Pi(\cdot, j)^\top \\ &= \frac{1}{M} \sum_{j=1}^{tm} \frac{\# \text{ of times } (i, k) \text{ is sampled}}{\pi(i, k)} \Pi(\cdot, j) \Pi(\cdot, j)^\top \\ &= \frac{1}{M} \sum_{j=1}^M \hat{y}_j \hat{y}_j^\top, \end{aligned} \quad (30)$$

where Γ is a nonnegative diagonal matrix and the random entry $\Gamma(j, j)$ specifies the ‘‘amount’’ of the i -th input at time $k-1$ (where $\mathbf{m}(j) = (i, k)$) included in the sparse actuator scheduling by Algorithm 6. For instance, $\Gamma(j, j) = 1/M\pi(\mathbf{m}(j))$ if the i -th input at time $k-1$ is sampled once, $2/M\pi(\mathbf{m}(j))$ if it is sampled twice, and zero if it is not sampled at all. The scaling of the i -th input at time $k-1$ in the scheduling is given by $s_i^2(k-1) = \Gamma(j, j)$ where $\mathbf{m}(j) = (i, k)$. We next use a concentration lemma to prove this theorem. Using Lemma 2, (28), (29), and (30), we get

$$\begin{aligned} \mathbf{E} \left\| \frac{1}{M} \sum_{i=1}^M \hat{y}_i \hat{y}_i^\top - \mathbf{E} yy^\top \right\|_2 &= \mathbf{E} \|\Pi\Pi\Pi - \Pi\|_2 \quad (31) \\ &\leq \min \left(1, c\sqrt{\frac{n \log M}{M}} \right), \end{aligned} \quad (32)$$

where c is an absolute constant. Assuming $M = 9c^2 n \log n / \epsilon^2$ gives⁶

$$\mathbf{E} \|\Pi\Pi\Pi - \Pi\|_2 \leq c\sqrt{\frac{n \log M}{M}}$$

$$\leq \epsilon \sqrt{\frac{\log(9c^2 n \log n / \epsilon^2)}{9 \log n}} \leq \epsilon/2, \quad (33)$$

for n sufficiently large, and ϵ is assumed to be in $[1/\sqrt{n}, 1)$. By Markov’s inequality and (33), we have

$$\Pr \{ \|\Pi\Pi\Pi - \Pi\| > \epsilon \} \leq 0.5,$$

which means we have

$$\|\Pi - \Pi\Pi\Pi\|_2 \leq \epsilon, \quad (34)$$

with probability of at least 0.5. Note that Γ is a non-negative diagonal matrix with weights $s_i^2(k)$ on its diagonal such that $\mathcal{W}_s(t) = \mathcal{C}(t) \Gamma \mathcal{C}^\top(t)$. Based on [27, Lemma 4], the inequality (34) is equivalent to

$$\sup_{\substack{x \in \mathbb{R}^{tm} \\ x \neq 0}} \frac{|x^\top (\Pi - \Pi\Pi\Pi)x|}{x^\top x} \leq \epsilon. \quad (35)$$

Since we have $\text{Im}\{\mathcal{C}^\top(t)\} \subset \mathbb{R}^{mt}$, it follows that

$$\sup_{\substack{x \in \text{Im}\{\mathcal{C}^\top(t)\} \\ x \neq 0}} \frac{|x^\top (\Pi - \Pi\Pi\Pi)x|}{x^\top x} \leq \sup_{\substack{x \in \mathbb{R}^{mt} \\ x \neq 0}} \frac{|x^\top (\Pi - \Pi\Pi\Pi)x|}{x^\top x} \leq \epsilon.$$

Let us define $x = \mathcal{C}^\top(t)x'$. Then, we rewrite (35) as follows

$$\sup_{\substack{x' \in \mathbb{R}^n \\ x' \notin \ker\{\mathcal{C}^\top(t)\}}} \frac{|x'^\top (\mathcal{W}(t) - \mathcal{W}_s(t))x'|}{x'^\top \mathcal{W}(t)x'} \leq \epsilon. \quad (36)$$

As a result, it follows that

$$\sup_{\substack{x' \in \mathbb{R}^n \\ x' \neq 0}} \frac{|x'^\top (\mathcal{W}(t) - \mathcal{W}_s(t))x'|}{x'^\top \mathcal{W}(t)x'} \leq \epsilon, \quad (37)$$

which implies that

$$(1 - \epsilon)\mathcal{W}(t) \preceq \mathcal{W}_s(t) = \mathcal{C}(t)\Gamma\mathcal{C}^\top(t) \preceq (1 + \epsilon)\mathcal{W}(t). \quad (38)$$

Finally, using (38) and Definition 2, it is straightforward to show that for every systemic controllability measure $\rho : \mathbb{S}_+^n \rightarrow \mathbb{R}_+$, we have

$$\left| \frac{\rho(\mathcal{W}(t)) - \rho(\mathcal{W}_s(t))}{\rho(\mathcal{W}(t))} \right| \leq \epsilon.$$

Therefore, we conclude the desired result. \blacksquare

This result shows that with a simple randomized sampling strategy, one can choose on average less than $\mathcal{O}(\log n / \epsilon^2)$ number of actuators at each time, to approximate any of the controllability metrics when $t = n$. Moreover, this result shows that it is possible to have a time-varying actuator schedule with a constant number of active actuators on average over a time horizon a little longer than n (*i.e.*, $t = \mathcal{O}(n \log n)$) via random sampling. Algorithm 6 computes the sparse actuator schedule using a nearly-linear time $\tilde{\mathcal{O}}(mt)$ algorithm⁷ with guaranteed performance bounds, where mt is the total number of actuations (time-to-control \times number of inputs). This favorable almost-linear-time complexity is achieved by random sampling of actuators in both time and domain based on their leverage scores [27]. According to Theorem 1, the average number of active actuators can be reduced to $\mathcal{O}(1/\epsilon^2)$, at the expense of either solving SDPs [26] or greedily handling certain eigenvalue bounds (see Algorithm 2). Algorithm 6 is conceptually simpler than Algorithm 2 and the SDP-based algorithm presented in [26], which provide $d = \mathcal{O}(1/\epsilon^2)$ in $\mathcal{O}(m(tn)^2/\epsilon^2)$ and $\tilde{\mathcal{O}}(mt/\epsilon^{\mathcal{O}(1)})$ time, respectively.

The concept of a leverage score for each column can be generalized

⁶It can be shown that $M = 4n \log n \epsilon^2$ would be enough to get ϵ approximation with high probability [47].

⁷ $f(n) \in \tilde{\mathcal{O}}(g(n))$ means that there exists $c > 0$ such that $f(n) \in \mathcal{O}(g(n) \log^c g(n))$.

to a group of columns as follows

$$\ell_{c_i} = \text{Trace} \left(C_i^\top(t) \left(\mathcal{C}(t) \mathcal{C}^\top(t) \right)^\dagger C_i(t) \right). \quad (39)$$

Using group leverage scores, one can also use a greedy heuristic algorithm to obtain an approximation solution for the static scheduling problem. We note that the problem of approximation of the controllability Gramian with a sparse, static actuator set is considerably more challenging as it doesn't lend itself to a sampling-based strategy: any choice made at one time has to be consistent with the next.

When using a time-varying schedule, the contribution of each actuator to the Gramian at each time is a rank-one matrix. Therefore, we can use the machinery developed for the Kadison-Singer conjecture to find a sparse subset of actuators over time to approximate the (potentially very large) sum of rank-one matrices. In the static case, however, the choices of actuators at different times are all the same. As a result, the Gramian can be written as a sum of positive semi-definite matrices corresponding to the selected actuators at each time. Finding a sparse approximation in this case would require a generalization of the Kadison-Singer conjecture from sums of rank-one to sums of higher ranked positive semidefinite matrices. Such a result has remained elusive as of yet.

VI. AN UNWEIGHTED SPARSE ACTUATOR SCHEDULE

In the previous section, we allowed for re-scaling of the input to come up with a sparse approximation of the Gramian. Here, we assume that the actuator/signal strength cannot be arbitrarily set for individual active actuators and only can be 0 or 1. Given a time horizon $t \geq n$, our problem is to compute an actuator schedule $\sigma = \{\sigma_k\}_{k=0}^{t-1}$ where $\sigma_k \subset [m]$ for the system (1), i.e.,

$$x(k+1) = Ax(k) + \sum_{i \in \sigma_k} b_i u_i(k), \quad k \in \mathbb{Z}_+. \quad (40)$$

As before, the controllability Gramian at time t for schedule (40) is given by

$$\mathcal{W}_\sigma(t) := \sum_{i=0}^{t-1} \sum_{j \in \sigma_i} (A^{t-i-1} b_j)(A^{t-i-1} b_j)^\top. \quad (41)$$

Optimal actuator selection can now be formulated as a combinatorial optimization problem. We consider both static and dynamic actuator schedules, corresponding to time-invariant and time-varying input matrices.

1) *The Static Scheduling Problem:* In this case, all sets $\sigma_i \subset [m]$ for $i+1 \in [t]$ are identical, which means we keep the same schedule at every point in time for the whole time horizon t :

$$\min_{\sigma \in \mathcal{S}(m, d_{\max})} \rho \left(\sum_{i=0}^{n-1} \sum_{j \in \sigma} (A^i b_j)(A^i b_j)^\top \right), \quad (42)$$

where

$$\mathcal{S}(m, d_{\max}) := \{\sigma : \sigma \subset [m], \text{card}(\sigma) \leq d_{\max}\}, \quad (43)$$

where d_{\max} is a given upper bound on the number of active actuators at each time, and m is the total number of actuators.

2) *The Time-varying Scheduling Problem:* In this case, the optimal dynamic strategy is given as:

$$\min_{\{\sigma_i\}_{i=0}^{t-1} \in \mathcal{S}(m, d_{\max}, t)} \rho \left(\sum_{i=0}^{t-1} \sum_{j \in \sigma_i} (A^{t-i-1} b_j)(A^{t-i-1} b_j)^\top \right), \quad (44)$$

where

$$\mathcal{S}(m, d_{\max}, t) := \left\{ \{\sigma_i\}_{i=0}^{t-1} : \sigma_i \subset [m], \sum_{i=0}^{t-1} \text{card}(\sigma_i) \leq t d_{\max} \right\}, \quad (45)$$

and d_{\max} is a given upper bound on the average number of active actuators at each time, i.e., $d_{\max} \geq \sum_{i=0}^{t-1} \text{card}(\sigma_i)/t$, where t is a time horizon, and m is the total number of actuators.

The exact combinatorial optimization problems (42) and (44) are intractable and NP-hard optimization problems; however, it is straightforward to solve a continuous relaxation of these optimization problems where the cost function ρ is convex. To find a near-optimal solution of optimization problems (42) and (44), one can use a variety of standard methods for optimal experimental design (greedy methods, sampling methods, the classical pipage rounding method combined with SDP). Specifically, in the case of submodular systemic controllability measures (e.g., D- and T-optimality), the classical rounding method (e.g., pipage and randomized rounding) combined with SDP relaxation results in computationally fast algorithms with a constant approximation ratio [45]. These approaches are not applicable to non-submodular systemic measures, such as A- and E-optimality [25], [48].

In the following result, we use a result based on regret minimization of the least eigenvalues of positive semi-definite matrices (cf. [25]) to obtain a constant approximation ratio for all systemic controllability metrics.

Theorem 3: Assume that time horizon $t \geq n$, dynamics (1), systemic controllability metric $\rho : \mathbb{S}_+^n \rightarrow \mathbb{R}$, and $d_{\max} > 2$ are given. Then there exists a polynomial-time algorithm which computes a schedule $\hat{\sigma} = \{\hat{\sigma}_i\}_{i=0}^{t-1}$ that satisfies

$$\rho(\mathcal{W}_{\hat{\sigma}}(t)) \leq \gamma \left(\frac{d_{\max} t}{n} \right) \cdot \min_{\{\sigma_i\}_{i=0}^{t-1} \in \mathcal{S}(m, d_{\max}, t)} \rho(\mathcal{W}_\sigma(t)),$$

where $\gamma(d_{\max} t/n)$ is a positive constant depending only on $d_{\max} t/n$.

Proof: The proof is a simple variation on the proof of [25, thm. 1.1], and is not repeated here. ■

The positive constant $\gamma(\cdot)$ in Theorem 3 is defined as follows

$$\gamma(\zeta) = \min_{y > \frac{3\zeta}{\zeta-2}} \frac{v(2 + \frac{v}{\zeta})}{(1 - \frac{2}{\zeta})v - 3}, \quad \text{where } \zeta > 2, \quad (46)$$

see the proof of [25, thm. 1.1]. For example, for $d_{\max} t/n \in \{4, 10, 50\}$, using (46), we get

$$\gamma(4) = 2 \left(5 + \sqrt{21} \right) \approx 19.1652,$$

$$\gamma(10) = 5 \left(\frac{11 + \sqrt{57}}{16} \right) \approx 5.79682,$$

and

$$\gamma(50) = 25 \left(\frac{17 + \sqrt{33}}{192} \right) \approx 2.96153.$$

Next, we use the results from Section V to obtain an unweighted sparse actuator schedule with guaranteed performance bound.

Corollary 4: Assume that time horizon $t \geq n$, dynamics (1), and $d_{\max} > 1$ are given. Then polynomial-time Algorithm 7 deterministically constructs an actuator schedule for (12) with $s_i(k) \in \{0, 1\}$ such that it has, on average, at most d_{\max} active actuators, and the following

$$\rho(\mathcal{W}_\sigma(t)) \leq \left(\frac{1 + \sqrt{\frac{m}{d_{\max}}}}{1 - \sqrt{\frac{n}{d_{\max} t}}} \right)^2 \rho(\mathcal{W}(t)),$$

holds for all systemic controllability measures.

Proof: The proof is a simple variation on the proof of Theorem 1, and is not repeated here. ■

In view of this result, one can choose any constant number greater than one as the number of active actuators on average to construct a sparse unweighted actuator schedule in order to approximate

Algorithm 7: A deterministic greedy-based algorithm to construct a sparse unweighted actuator schedule (Corollary 4).

Input : $A \in \mathbb{R}^{n \times n}$, $B \in \mathbb{R}^{n \times m}$, t and d_{\max}
Output: $s_i(k)$ for $(i, k + 1) \in [m] \times [t]$

```

1  $C(t) := [B \ AB \ A^2B \ \dots \ A^{t-1}B]$ 
2 Set  $V = (C(t)C^\top(t))^{-\frac{1}{2}} C(t)$ 
3 Set
   
$$U = \begin{bmatrix} \mathbf{e}_1, \dots, \mathbf{e}_{mt} \\ \hline = I_{mt} \end{bmatrix}$$

   // where  $\mathbf{e}_i \in \mathbb{R}^{mt}$  for  $i \in [mt]$  are the standard
   // basis vectors for  $\mathbb{R}^{mt}$ 
4 Run  $[c_1, \dots, c_{mt}] = \text{DualSet}^*(V, U, d_{\max}t)$ 
5 return  $s_i(k) := \left\lceil \frac{\sqrt{c_{i+mk}}}{1 + \sqrt{\frac{m}{d_{\max}}}} \right\rceil$  for
    $(i, k + 1) \in [m] \times [t]$ 

```

Algorithm 8: A greedy heuristic for given $\rho(\cdot)$ which sequentially picks inputs GreedyStatic(A, B, t, d).

Input : $A \in \mathbb{R}^{n \times n}$, $B \in \mathbb{R}^{n \times m}$, t and d
Output: $B_s \in \mathbb{R}^{n \times d}$, $\rho(\mathcal{W}_s)$

```

1  $\mathcal{W}_s := \mathbf{0}_{n \times n}$ 
2 for  $k = 1$  to  $d$  do
3    $j \leftarrow$  find a column of  $B$  that returns the maximum value for
   
$$\rho(\mathcal{W}_s + \alpha I_n) - \rho\left(\mathcal{W}_s + \sum_{i=0}^{t-1} A^i B(:, j) B(:, j)^\top (A^i)^\top + \alpha I_n\right)$$

   //  $\alpha > 0$  is sufficiently small to avoid
   // singularity
4    $B_s \leftarrow [B_s, B(:, j)]$ 
5    $\mathcal{W}_s = \sum_{i=0}^{t-1} A^i B_s B_s^\top (A^i)^\top$ 
6    $B(:, j) \leftarrow []$ 
7 end
8 return  $B_s, \rho(\mathcal{W}_s)$ 

```

controllability measures. This, however, comes at the cost of an extra $\left(1 + \sqrt{\frac{m}{d_{\max}}}\right)^2$ factor in terms of the energy cost compared to the weighted sparse actuator schedule (cf. Corollary 1).

VII. NUMERICAL EXAMPLES

In this section, we consider three numerical examples to demonstrate the results.

We compare our results with a greedy heuristic that sequentially picks control inputs to maximize the systemic metric decrease of the controllability matrix (see Algorithm 8). The selected inputs are active at all times. It is shown that the greedy method works well and matches the inapproximability barrier⁸ in polynomial time [1]. We also compare our results with a greedy algorithm for a time-varying actuator schedule that sequentially picks both control inputs and activation times to maximize the decrease in the systemic metric of the controllability Gramian (see Algorithm 9). Without loss of generality, we assume time horizon $t = n$.

⁸It approximates the minimum number of inputs in the system that need to be affected for controllability within a factor of $c \log n$ for some $c > 0$.

Algorithm 9: A greedy heuristic for given $\rho(\cdot)$ which sequentially picks inputs and activation times GreedyTimeVarying(A, B, t, d).

Input : $A \in \mathbb{R}^{n \times n}$, $B \in \mathbb{R}^{n \times m}$, t and d
Output: $\rho(\mathcal{W}_s)$

```

1  $C := [B \ AB \ A^2B \ \dots \ A^{t-1}B]$ 
2  $C_s := \mathbf{0}_{n \times mt}$ 
3 for  $k = 1$  to  $M := [dt]$  do
4    $j \leftarrow$  find a column of  $C$  that returns the maximum value for
   
$$\rho(\mathcal{W}_s + \alpha I_n) - \rho(\mathcal{W}_s + C(:, j) C(:, j)^\top + \alpha I_n)$$

   //  $\alpha > 0$  is sufficiently small to avoid
   // singularity
5    $C_s \leftarrow [C_s, C(:, j)]$ 
6    $\mathcal{W}_s = C_s C_s^\top$ 
7    $C(:, j) \leftarrow []$ 
8 end
9 return  $\rho(\mathcal{W}_s)$ 

```

Example 1 ([1]): Assume that the state space matrices of system (1) are given by

$$A = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 & 0 & 0 & \frac{-7}{2} \\ 0 & 2 & 0 & 0 & 0 & 0 & 0 & -3 \\ 0 & 0 & 3 & 0 & 0 & 0 & 0 & \frac{-5}{2} \\ \frac{3}{4} & \frac{1}{3} & 0 & 4 & 0 & 0 & 0 & \frac{13}{8} \\ 0 & \frac{3}{4} & \frac{1}{2} & 0 & 5 & 0 & 0 & \frac{8}{3} \\ \frac{5}{4} & 0 & \frac{3}{4} & 0 & 0 & 6 & 0 & \frac{11}{3} \\ \frac{4}{2} & \frac{3}{2} & 1 & 0 & 0 & 0 & 7 & \frac{9}{4} \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 8 \end{bmatrix} \quad (47)$$

and

$$B_{\min} = \text{diag}[1, 1, 0, 0, 0, 0, 0, 1] \quad (48)$$

Direct computation shows that choosing (48) makes the system controllable and no diagonal-matrix sparser than B_{\min} renders A controllable. For this case ($B = B_{\min}$), the performance is:

$$\text{Trace} \left(\sum_{i=0}^{n-1} A^i B_{\min} B_{\min}^\top (A^i)^\top \right)^{-1} = 0.503,$$

and for the fully actuated case (i.e., $B = I_8$), we have

$$\text{Trace} \left(\sum_{i=0}^{n-1} A^i B B^\top (A^i)^\top \right)^{-1} = 0.132.$$

We compare our method with simple-random and periodical switching methods which are depicted in Fig. 2, and obtain systemic controllability performances, which are presented in Table II.

Example 2: Let us consider a dynamic network consisting of $n = 200$ agents/nodes, which are randomly distributed in a 1×1 square-shape area in space and are coupled over a proximity graph. Every agent is connected to all of its spatial neighbors within a closed ball of radius $r = 0.125$. Assume that the state space matrices of this network are given by

$$A = I_n - \frac{1}{n}L, \text{ and } B = I_n, \quad (49)$$

where L is the Laplacian matrix of the underlying graph given by Fig. 3. Now, we consider the actuator scheduling problem discussed in Section V. For undirected consensus networks, a similar problem arises in assignment of a pre-specified number of active agents, as

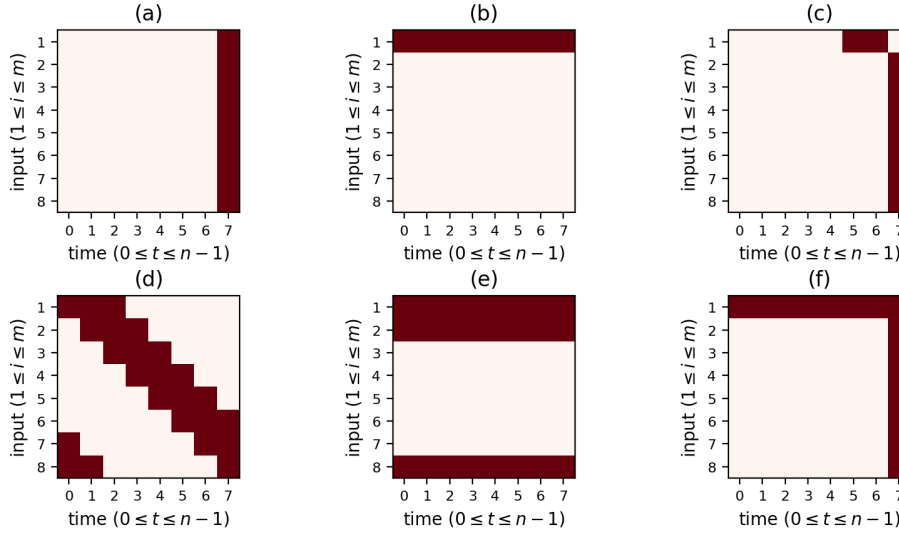


Fig. 2: Six unweighted actuator schedules for Example 1: (a) all actuators are active at time 7 (b) actuator one is active at each time (c) the schedule is obtained Algorithm 7 (d) three actuators are active at all time and each actuator is used three times (e) three fixed actuators $\{1, 2, 8\}$ are active at all time (f) the proposed sparse schedule based on Algorithm 7 with less than two active actuators at each time on average. The color of element (i, k) is red when $s_i(k) = 1$ and white otherwise where $i \in [8]$, $k + 1 \in [8]$ and $s_i(k) \in \{0, 1\}$. For Figs. 2 (c)&(f), which are obtained based on Algorithm 7, we can observe that the actuator schedule has procrastination in actuator activations (i.e., more active actuators at the end of the time horizon); however, in Example 3 we can see “front-loaded” behavior (i.e., more active actuators early in the time horizon) due to different dynamics in this example.

	Figs. 2.(a)&(b)	Fig. 2.(c)	Fig. 2.(d)	Fig. 2.(e)	Fig. 2.(f)	Algorithm 8	Algorithm 9	Fully Actuated
Trace $(\mathcal{W}^{-1}(n))$	uncontrollable	0.628	uncontrollable	0.503	0.161	uncontrollable	0.294	0.132
d	1	1.125	3	3	1.875	3	3	8

TABLE II: The values of controllability performance and average number of active actuators at each time for the unweighted actuator schedule presented in Fig. 2 and based on greedy algorithms 8 and 9. The unweighted schedules presented in Figs. 2 (c)&(f) are obtained based on Algorithm 7. It is not possible to greedily select three inputs (active at all time) to make the system in Example 1 controllable.

	Fig. 4 (Algorithm 6)	Static Leader Schedule	Fully Actuated
Trace $(\mathcal{W}^{-1}(n))$	93.64	676.68	18.16
Average Number of Leaders: d	40	160	200

TABLE III: The values of controllability performance for three different actuator schedules in Example 2: 1) the weighted actuator schedule in Fig. 4 based on Algorithm 6, 2) the static leader schedule with 160 leaders active at all time, 3) the fully actuated case. To have a fair comparison, we normalize the resulting schedule of Algorithm 6 such that the sum of the scalings satisfies $\sum_{k=0}^{n-1} \sum_{i=1}^m s_i^2(k) = dn$ where $d = 40$. The value of the controllability metric for the materialized result of Algorithm 6 is 18.54, which is much closer to the controllability metric of the fully actuated case.

leaders, in order to minimize the controllability metric, e.g., the average controllability energy (cf. [49], [50]). In our setup, each leader i in addition to relative information exchange with its neighbors (based on Laplacian matrix L), it also has access to a control input $u_i(\cdot)$. This system is controllable with only a few inputs/leaders⁹; however, the amount of the average control energy with a static actuator/leader schedule is too large even for a large number of leaders (see Table III). On the other hand, with a time-varying strategy, the resulting performance is close to the fully actuated case even with a small number of leaders. Therefore, instead of choosing the same leaders at every time step, we choose/switch leaders over a given time horizon to further decrease the controllability metric.

Fig. 3 shows the underlying graph, and node colors are proportional to the total number of active steps during time steps 0 to 199 from

⁹The system is not controllable with only one input, because A does not have distinct eigenvalues [50].

least (white) to greatest (red). Fig. 4 depicts a sparse schedule based on Algorithm 6.

Example 3 (Power Network): The problem is to select a set of generators to be involved in the wide-area damping control of power systems. We apply our sparse scheduling approach on the IEEE 39-bus test system (a.k.a. the 10-machine New England Power System; see Fig. 5) [33], [34]. The single line diagram presented in this figure comprises generators (G_i where $i \in [10]$), loads (arrows), transformers (double circles), buses (bold line segments with number $i \in [39]$), and lines between buses (see [33], [34]).

The goal of the wide-area damping control is to damp the fluctuations between generators and synchronize all generators. The voltage at each generator is adjusted by the control inputs (e.g., HVDC lines and storages) to regulate the power output.

We start with a model representing the interconnection between

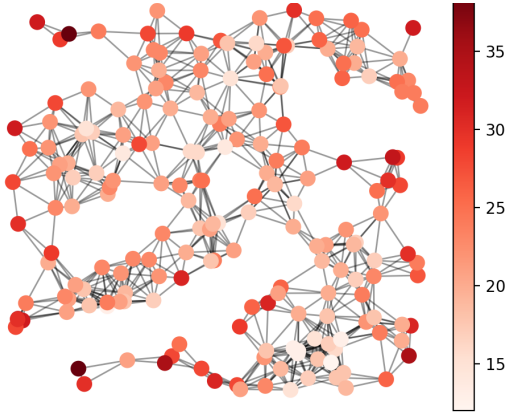


Fig. 3: A dynamical network consists of 200 agents that are randomly distributed in a 1×1 square-shape area in space and are coupled over a proximity graph. Every agent is connected to all of its spatial neighbors within a closed ball of radius $r = 0.125$. Node colors are proportional to the total number of active steps during time steps 0 to 199 from least (white) to greatest (red) based on Algorithm 6 where $d = 40$ (i.e., which means that, on average, only 20% of agents are controlled at each time).

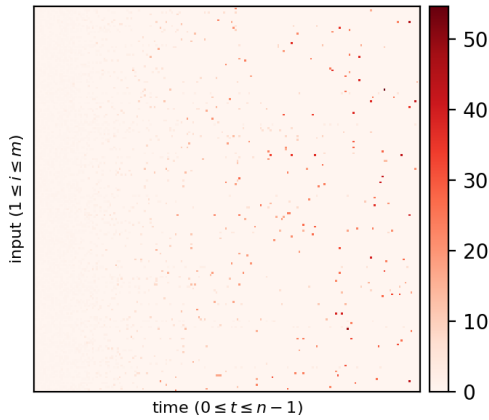


Fig. 4: A sparse schedule based on Algorithm 6 for a given network in Fig. 3 where $d = 40$. This dynamical network has $m = 200$ inputs; however, on average, only 20% are active at each time between 0 to $n - 1$. The color of element (i, k) is proportional to the scaling factor $s_i^2(k)$ where $i \in [200]$ and $k + 1 \in [200]$.

subsystems. Consider the swing dynamics

$$m_i \ddot{\theta}_i + d_i \dot{\theta}_i = - \sum_{j \sim i} k_{ij} (\theta_i - \theta_j) + u_i,$$

where θ_i is the rotor angle state and $w_i := \dot{\theta}_i$ is the frequency state of generator i . We assume this power grid model consists of $n = 10$ generators [33], [34]. The state space model of the swing equation used for frequency control in power networks can be written as follows

$$\begin{aligned} \begin{bmatrix} \dot{\theta}(t) \\ \dot{w}(t) \end{bmatrix} &= \begin{bmatrix} 0 & I \\ -M^{-1}L & -M^{-1}D \end{bmatrix} \begin{bmatrix} \theta(t) \\ w(t) \end{bmatrix} \\ &+ \begin{bmatrix} 0 \\ M^{-1} \end{bmatrix} u(t) \\ y(t) &= \begin{bmatrix} \theta(t) \\ w(t) \end{bmatrix} \end{aligned}$$

where M and D are diagonal matrices with inertia coefficients and

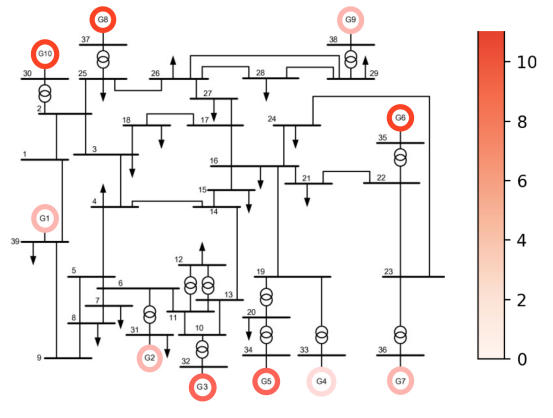


Fig. 5: IEEE 10-generator 39-bus power system network (figure is adapted from [33]). Generator colors are proportional to the total number of active steps during time steps 0 to 19 from least (white) to greatest (red) based on Algorithm 7 where $d = 4$ (i.e., which means that, on average, only four generators are controlled at each time).

damping coefficients of generators and their diagonals, respectively.

We assume that both rotor angle and frequency are available for measurement at each generator. This means each subsystem in the power network has a phase measurement unit (PMU). The PMU is a device that measures the electrical waves on an electricity grid using a common time source for synchronization. The system is discretized to the discrete-time LTI system with state matrices A , B , and C and the sampling time of 0.2 second (the matrices are borrowed from [51]).

Fig. 6 depicts nine sparse schedules based on the proposed deterministic method (Algorithms 3) for different values of d . The sparsity degree of each schedule is captured by d . As d increases the number of non-zero scalings (i.e., activations) increases while the controllability metric decreases (improves). Fig. 7 compares the results of Algorithms 3, 7, 8, and 9. The plot presents the values of the average controllability energy (A-optimality) versus the average number of active actuators. To have a fair comparison, we normalize the resulting schedules of all the methods such that the sum of all the scalings satisfies $\sum_{k=0}^{n-1} \sum_{i=1}^m s_i^2(k) = nd$.

As one expects, Algorithms 3, 7, and 9 outperform Algorithm 8. One observes that Algorithms 3, 7 perform nearly as optimal as the time-varying greedy method 9; however, based on our results, we have theoretical guaranteed performance bounds for Algorithms 3 and 7. Furthermore, the usefulness of Algorithms 3, 7 accentuates itself when the number of active actuators on average is not too small; and potentially can result in a better solution compare to Algorithm 9 (see Fig. 7).

VIII. CONCLUDING REMARKS

In this paper, we have shown how recent advances in matrix reconstruction and graph sparsification literature can be utilized to develop subset selection tools for choosing a relatively small subset of actuators to approximate certain controllability measures. Current approaches based on polynomial time relaxations of the subset selection problem require an extra multiplicative factor of $\log n$ sensors/actuators times the minimal number in order to just maintain controllability/observability. Furthermore, when the control energy is chosen as the cost, submodularity-based approaches fail to guarantee the performance using greedy methods. In contrast, we show that there exists a polynomial-time actuator schedule that on average selects only a constant number of actuators at each time, to approximate

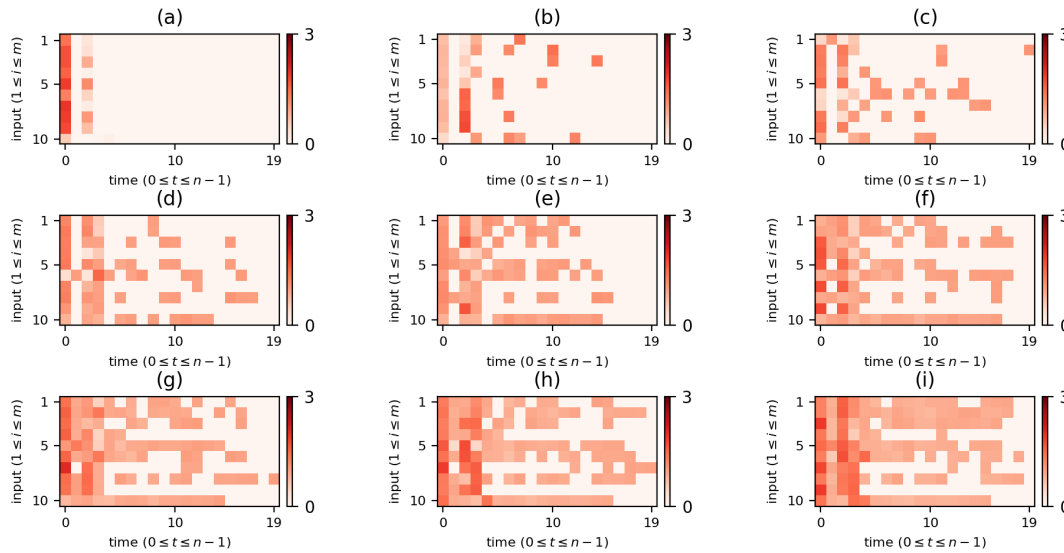


Fig. 6: Subplots (a)-(i) presents nine weighted sparse schedules for Example 3 based on the proposed deterministic method (Algorithm 3) where $d \in \{1.05, 1.75, 2.30, 3.10, 3.95, 4.60, 5.25, 5.75, 6.35\}$ is the average number of active actuators at each time, respectively. The color of element (i, k) is proportional to the scaling factor $s_i^2(k)$ where $i \in [10]$ and $k + 1 \in [20]$.

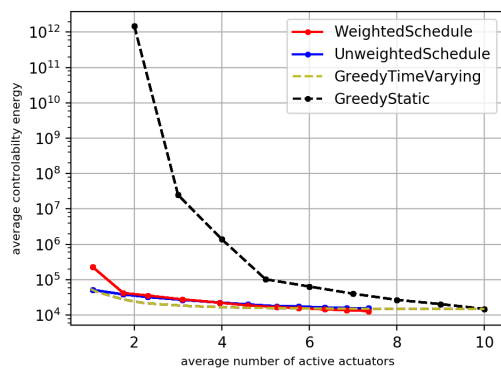


Fig. 7: This plot compares four different methods (Algorithms 3, 7, 8 and 9) for obtaining sparse actuator schedules of the 10-machine New England Power System in Example 3. The plot presents the values of average controllability energy (A-optimality) versus the average number of active actuators at each time (d).

controllability measures. Similar results can be developed for the sensor selection problem. A potential future direction is to see whether this approach can be used to develop an efficient scheme for minimal reachability problems.

REFERENCES

- [1] A. Olshevsky, "Minimal controllability problems," *IEEE Transactions on Control of Network Systems*, vol. 1, no. 3, pp. 249–258, Sept 2014.
- [2] F. Pasqualetti, S. Zampieri, and F. Bullo, "Controllability metrics, limitations and algorithms for complex networks," *IEEE Transactions on Control of Network Systems*, vol. 1, no. 1, pp. 40–52, 2014.
- [3] Y.-Y. Liu and A.-L. Barabási, "Control principles of complex systems," *Reviews of Modern Physics*, vol. 88, no. 3, p. 035006, 2016.
- [4] P. V. Chanekar, N. Chopra, and S. Azarm, "Optimal actuator placement for linear systems with limited number of actuators," in *2017 American Control Conference (ACC)*, May 2017, pp. 334–339.
- [5] P. Müller and H. Weber, "Analysis and optimization of certain qualities of controllability and observability for linear dynamical systems," *Automatica*, vol. 8, no. 3, pp. 237–246, 1972.
- [6] V. Tzoumas, M. A. Rahimian, G. J. Pappas, and A. Jadbabaie, "Minimal actuator placement with bounds on control effort," *IEEE Transactions on Control of Network Systems*, vol. 3, no. 1, pp. 67–78, 2016.
- [7] A. Olshevsky, "Minimum input selection for structural controllability," in *2015 American Control Conference (ACC)*, July 2015, pp. 2218–2223.
- [8] S. Pequito, G. Ramos, S. Kar, A. P. Aguiar, and J. Ramos, "The robust minimal controllability problem," *Automatica*, vol. 82, pp. 261 – 268, 2017.
- [9] E. Nozari, F. Pasqualetti, and J. Cortés, "Time-invariant versus time-varying actuator scheduling in complex networks," in *2017 American Control Conference (ACC)*, May 2017, pp. 4995–5000.
- [10] A. Yazıcıoğlu, W. Abbas, and M. Egerstedt, "Graph distances and controllability of networks," *IEEE Transactions on Automatic Control*, vol. 61, no. 12, pp. 4125–4130, 2016.
- [11] T. H. Summers, F. L. Cortesi, and J. Lygeros, "On submodularity and controllability in complex dynamical networks," *IEEE Transactions on Control of Network Systems*, vol. 3, no. 1, pp. 91–101, 2016.
- [12] S. Pequito, S. Kar, and A. P. Aguiar, "On the complexity of the constrained input selection problem for structural linear systems," *Automatica*, vol. 62, pp. 193–199, 2015.
- [13] A. Chakraborty and M. D. Ilić, *Control and optimization methods for electric smart grids*. Springer, 2011, vol. 3.
- [14] F. A. Chandra, G. Buzi, and J. Doyle, "Glycolytic oscillations and limits on robust efficiency," *Science*, vol. 333, no. 6039, pp. 187–192, 2011.
- [15] L. Marucci, D. A. Barton, I. Cantone, M. A. Ricci, M. P. Cosma, S. Santini, D. di Bernardo, and M. di Bernardo, "How to turn a genetic circuit into a synthetic tunable oscillator, or a bistable switch?" *PLoS One*, vol. 4, no. 12, p. e8083, 2009.
- [16] I. Rajapakse, M. Groudine, and M. Mesbahi, "What can systems theory of networks offer to biology?" *PLoS computational biology*, vol. 8, no. 6, p. e1002543, 2012.
- [17] M. Siami and J. Skaf, "Structural analysis and optimal design of distributed system throttlers," *IEEE Transactions on Automatic Control*, vol. 63, no. 12, pp. 540–547, Feb. 2018.
- [18] R. E. Kalman, "Mathematical description of linear dynamical systems," *Journal of the Society for Industrial and Applied Mathematics, Series A: Control*, vol. 1, no. 2, pp. 152–192, 1963.

- [19] M. Athans, "On the determination of optimal costly measurement strategies for linear stochastic systems," *Automatica*, vol. 8, no. 4, pp. 397–412, 1972.
- [20] V. Tzoumas, M. A. Rahimian, G. J. Pappas, and A. Jadbabaie, "Minimal actuator placement with bounds on control effort," *IEEE Transactions Control of Network Systems*, vol. 3, no. 1, pp. 67–78, 2016.
- [21] A. Jadbabaie, A. Olshevsky, G. J. Pappas, and V. Tzoumas, "Minimal reachability is hard to approximate," *arXiv preprint arXiv:1710.10244*, 2017.
- [22] V. Tzoumas, A. Jadbabaie, and G. J. Pappas, "Minimal reachability problems," in *Proceedings of the 54th IEEE Conference on Decision and Control*, 2015, pp. 4220–4225.
- [23] V. Tzoumas, K. Gatsis, A. Jadbabaie, and G. J. Pappas, "Resilient monotone submodular function maximization," *arXiv preprint arXiv:1703.07280*, 2017.
- [24] O. Kempthorne, *The design and analysis of experiments*. Wiley, 1952.
- [25] Z. Allen-Zhu, Y. Li, A. Singh, and Y. Wang, "Near-optimal design of experiments via regret minimization," in *Proceedings of the 34th International Conference on Machine Learning*, vol. 70. PMLR, 06–11 Aug 2017, pp. 126–135.
- [26] Y. T. Lee and H. Sun, "An sdp-based algorithm for linear-sized spectral sparsification," *arXiv preprint arXiv:1702.08415*, 2017.
- [27] D. A. Spielman and N. Srivastava, "Graph sparsification by effective resistances," in *Proceedings of the Fortieth Annual ACM Symposium on Theory of Computing*, ser. STOC '08. New York, NY, USA: ACM, 2008, pp. 563–568.
- [28] J. Batson, D. A. Spielman, and N. Srivastava, "Twice-ramanujan sparsifiers," *SIAM Journal on Computing*, vol. 41, no. 6, pp. 1704–1721, 2012.
- [29] M. Rudelson and R. Vershynin, "Sampling from large matrices: An approach through geometric functional analysis," *Journal of the ACM*, vol. 54, no. 4, Jul. 2007. [Online]. Available: <http://doi.acm.org/10.1145/1255443.1255449>
- [30] A. Marcus, D. A. Spielman, and N. Srivastava, "Interlacing families ii: Mixed characteristic polynomials and the kadison-singer problem," *arXiv preprint arXiv:1306.3969*, 2013.
- [31] N. Srivastava and L. Trevisan, "An alon-boppana type bound for weighted graphs and lowerbounds for spectral sparsification," *arXiv preprint arXiv:1707.06364*, 2017.
- [32] M. A. Elizondo, R. Fan, H. Kirkham, M. Ghosal, F. Wilches-Bernal, D. Schoenwald, and J. Lian, "Interarea oscillation damping control using high-voltage dc transmission: A survey," *IEEE Transactions on Power Systems*, vol. 33, no. 6, pp. 6915–6923, 2018.
- [33] I. E. Atawi, "An advance distributed control design for wide-area power system stability," Ph.D. dissertation, University of Pittsburgh, 2013.
- [34] Z. Liu, Y. Long, A. Clark, P. Lee, L. Bushnell, D. Kirschen, and R. Poovendran, "Minimal input selection for robust control," *arXiv preprint arXiv:1712.01232*, 2017.
- [35] S. Boyd, "Lecture notes for EE263," *Introduction to Linear Dynamical Systems*, 2008.
- [36] C. Nowzari, V. M. Preciado, and G. J. Pappas, "Analysis and control of epidemics: A survey of spreading processes on complex networks," *IEEE Control Systems Magazine*, vol. 36, no. 1, pp. 26–46, 2016.
- [37] M. Siami and A. Jadbabaie, "Deterministic polynomial-time actuator scheduling with guaranteed performance," in *2018 European Control Conference (ECC)*, 2018, pp. 113–118.
- [38] A. Jadbabaie, A. Olshevsky, and M. Siami, "Limitations and tradeoffs in minimum input selection problems," in *2018 Annual American Control Conference (ACC)*, 2018, pp. 185–190.
- [39] C. Boutsidis, P. Drineas, and M. Magdon-Ismael, "Near-optimal column-based matrix reconstruction," *SIAM Journal on Computing*, vol. 43, no. 2, pp. 687–717, 2014.
- [40] J. W. Silverstein, "The stieltjes transform and its role in eigenvalue behavior of large dimensional random matrices," *Random Matrix Theory and Its Applications. Lect. Notes Ser. Inst. Math. Sci. Natl. Univ. Singap.*, vol. 18, pp. 1–25, 2009.
- [41] M. A. Camacho, "Spectral sparsification," Ph.D. dissertation, 2014.
- [42] M. Siami and N. Motee, "Network abstraction with guaranteed performance bounds," *IEEE Transactions on Automatic Control*, vol. 63, no. 11, 2018.
- [43] —, "Systemic measures for performance and robustness of large-scale interconnected dynamical networks," in *Proceedings of the 53rd IEEE Conference on Decision and Control*, Dec. 2014.
- [44] —, "Growing linear dynamical networks endowed by spectral systemic performance measures," *IEEE Transactions on Automatic Control*, vol. 63, no. 8, 2018.
- [45] S. N. Ravi, V. Ithapu, S. Johnson, and V. Singh, "Experimental design on a budget for sparse linear models and applications," in *International Conference on Machine Learning*, 2016, pp. 583–592.
- [46] F. Pukelsheim, *Optimal design of experiments*. SIAM, 1993, vol. 50.
- [47] D. A. Spielman, "Lecture notes in spectral and algebraic graph theory," *Yale University*, delivered Nov. 14, 2018, <http://cs-www.cs.yale.edu/homes/spielman/sagt/sagt.pdf>.
- [48] A. Olshevsky, "On (non)supermodularity of average control energy," *IEEE Transactions on Control of Network Systems*, 2018, to appear.
- [49] F. Lin, M. Fardad, and M. R. Jovanović, "Algorithms for leader selection in stochastically forced consensus networks," *IEEE Transactions on Automatic Control*, vol. 59, no. 7, pp. 1789–1802, 2014.
- [50] A. Rahmani, M. Ji, M. Mesbahi, and M. Egerstedt, "Controllability of multi-agent systems from a graph-theoretic perspective," *SIAM Journal on Control and Optimization*, vol. 48, no. 1, pp. 162–186, 2009.
- [51] G. Fazelnia, R. Madani, A. Kalbat, and J. Lavaei, "Convex relaxation for optimal distributed control problems," *IEEE Transactions on Automatic Control*, vol. 62, no. 1, pp. 206–221, Jan 2017.



Milad Siami (S'12-M'18) received his dual B.Sc. degrees in electrical engineering and pure mathematics from Sharif University of Technology in 2009, M.Sc. degree in electrical engineering from Sharif University of Technology in 2011. He received his M.Sc. and Ph.D. degrees in mechanical engineering from Lehigh University in 2014 and 2017 respectively. From 2009 to 2010, he was a research student at the Department of Mechanical and Environmental Informatics at the Tokyo Institute of Technology, Tokyo, Japan. He was a postdoctoral associate in the Institute for Data, Systems, and Society at MIT, from 2017 to 2019. He is currently an Assistant Professor with the Department of Electrical & Computer Engineering, Northeastern University, Boston, MA His research interests include distributed control systems, distributed optimization, and applications of fractional calculus in engineering. Dr. Siami received a Gold Medal at National Mathematics Olympiad, Iran (2003) and the Best Student Paper Award at the 5th IFAC Workshop on Distributed Estimation and Control in Networked Systems (2015). Dr. Siami was awarded RCEAS Fellowship (2012), Byllesby Fellowship (2013), Rossin College Doctoral Fellowship (2015), and Graduate Student Merit Award (2016) at Lehigh University.



Alex Olshevsky received the B.S. degree in applied mathematics and in electrical engineering from the Georgia Institute of Technology, Atlanta, GA, USA, both in 2004, and the M.S. and Ph.D. degrees in electrical engineering and computer science from the Massachusetts Institute of Technology, Cambridge, MA, USA, in 2006 and 2010, respectively. He is currently an Associate Professor at the Department of Electrical and Computer Engineering, Boston University, Boston, MA, USA. Prior to this position, he was a faculty member at the University of Illinois at Urbana-Champaign. He was a Postdoctoral Scholar in the Department of Mechanical and Aerospace Engineering, Princeton University, from 2010 to 2012 before joining the University of Illinois at Urbana-Champaign in 2012. His research interests are in control theory, optimization, and machine learning, especially in distributed, networked, and multi-agent settings. Dr. Olshevsky received the NSF CAREER Award, the Air Force Young Investigator Award, the ICS Prize from INFORMS for best paper on the interface of operations research and computer science, and a SIAM Paper Prize for annual paper from the SIAM Journal on Control and Optimization chosen to be reprinted in SIAM Review, and an award in 2019 from the International Medical Informatics Association for best paper on Clinical Research Informatics.



Ali Jadbabaie (S'99-M'08-SM'13-F'15) received the B.S. degree from Sharif University of Technology, Tehran, Iran, the M.S. degree in electrical and computer engineering from the University of New Mexico, Albuquerque, NM, USA, and the Ph.D. degree in control and dynamical systems from California Institute of Technology, Pasadena, CA, USA. He is the JR East Professor of engineering, the Associate Director of the Institute for Data, Systems and Society, and the Director of the Sociotechnical Systems Research Center, MIT, Cambridge, MA,

USA. He holds faculty appointments in the Department of Civil and Environmental Engineering and is a Principal Investigator in the Laboratory for Information and Decision Systems. He was a Postdoctoral Scholar at Yale University before joining the faculty at Penn in July 2002. Prior to joining MIT faculty, he was the Alfred Fidler Moore Professor of network science and held secondary appointments in computer and information science and operations, information, and decisions in the Wharton School. His current research interests include the interplay of dynamic systems and networks with specific emphasis on multiagent coordination and control, distributed optimization, network science, and network economics. He is the Inaugural Editor-in-Chief of the IEEE Transactions on Network Science and Engineering and an Associate Editor of the *Informa Journal Operations Research*. He received the National Science Foundation Career Award, an Office of Naval Research Young Investigator Award, the O. Hugo Schuck Best Paper Award from the American Automatic Control Council, the George S. Axelby Best Paper Award from the IEEE Control Systems Society, and the 2016 Vannevar Bush Fellowship from the office of Secretary of Defense.