INPUT SELECTION FOR PERFORMANCE AND CONTROLLABILITY OF STRUCTURED LINEAR DESCRIPTOR SYSTEMS*

ANDREW CLARK, BASEL ALOMAIR, LINDA BUSHNELL, AND RADHA POOVENDRAN†

Abstract. A common approach to controlling complex networks is to directly control a subset of input nodes, which then controls the remaining nodes via network interactions. Current approaches for selecting input nodes assume that either all system matrix entries are known and fixed, or are independent free parameters, and focus either on performance or controllability. In this paper, we make two contributions towards input selection in networked systems. First, we propose polynomial-time algorithms for input selection in structured linear descriptor systems, which are systems with dependencies between free parameters due to physical laws or design constraints. Second, we develop a framework for input selection based on joint consideration of controllability and performance. We make both contributions by mapping input selection to a submodular optimization problem under two matroid constraints, which enables development of polynomial-time algorithms with provable optimality guarantees. We provide improved optimality guarantees for special cases such as strongly connected networks, consensus networks, double integrators, and networks where all system parameters can take any arbitrary real values.

Key words. Linear descriptor systems, controllability, submodular optimization, matroids, matroid intersection, networked control systems, structured systems

AMS subject classifications. 93B05, 68W25, 90C27, 93C05

1. Introduction.

Complex networks consist of distributed nodes with locally coupled dynamics in domains including intelligent transportation systems [44], social networks [13], energy systems [39], and biological networks [25]. In many of these applications, the complex network must be controlled to reach a desired state, for example, steering a group of unmanned vehicles to a certain formation [35]. A common approach to controlling complex networks is to directly control one or more states of a subset of nodes, denoted as input nodes, while relying on the local coupling to drive the remaining nodes to the desired state [15, 18].

An important design parameter when controlling complex networks is the set of nodes that act as inputs. The problem of selecting a minimum-size set of input nodes to control a complex network has achieved significant attention [24, 33, 40], with recent work focusing on selecting input nodes to satisfy controllability, defined as the ability to drive the network from any initial state to any desired state in finite time using the input nodes. Since the seminal work of [24], a variety of discrete optimization methods have been proposed for selecting input nodes to achieve controllability [30, 33, 40].

In addition to controllability, networked systems are expected to satisfy performance criteria, including convergence rate of the dynamics [32] and robustness to noise and disturbances [1, 34]. Approaches for selecting input nodes based on these criteria, including submodular optimization [5, 8], convex relaxation [23], and combinatorial [12] algorithms, have been proposed in the literature, but are largely disjoint

*A preliminary version of this work appeared at the 51st IEEE Conference on Decision and Control (CDC), 2012.

†A. Clark, L. Bushnell, and R. Poovendran are with the University of Washington, Seattle, Washington. Email: {awclark, lb2,rp3}@uw.edu. B. Alomair is with the National Center for Cybersecurity Technology, King Abdulaziz City for Science and Technology, Riyadh, Saudi Arabia. Email: alomair@kacst.edu.sa. This work was supported by ONR grant N00014-14-1-0029, ONR grant N00014-16-1-2710, NSF grant CNS-1544173, and a grant from the King Abdulaziz City for Science and Technology.

from the algorithms for selecting input nodes to satisfy controllability. At present, a unifying and computationally efficient approach for selecting input nodes based on both controllability and performance is not available in the literature.

Furthermore, current approaches to input selection assume either that the system matrices are fully known, or that only the structure of the system matrices is known. In the latter case, the system matrix consists of zero entries and free parameters, where the free parameters can take any arbitrary value. Many practical systems, however, lie between these two extremes, with either a mixture of known and unknown matrix entries [26], or structural relationships between the unknown entries due to physical laws as in mass-spring systems and electrical circuits [20]. In consensus networks, for example, the system dynamics are governed by Laplacian matrices where individual entries are uncertain, and yet the sum of the entries on each row is equal to zero. In [3], it was observed that a set of selected input nodes may not satisfy controllability if these structural properties are not taken into account. While input selection methods that incorporate system structure have been proposed for specific applications such as consensus [14], a computationally tractable general framework that guarantees controllability of structured systems remains an open problem.

1.1. Our Contributions. In this paper, we develop a submodular optimization framework for input selection based on joint consideration of performance and controllability in structured linear descriptor systems. The submodular structure implies that a variety of input selection problems can be solved up to a $(1 - 1/e)$ optimality bound, using algorithms that depend on the constraints of each problem. We first show that selecting the minimum-size set of input nodes to satisfy controllability of structured linear descriptor systems can be mapped to a maximum-cardinality matroid intersection problem, leading to the first polynomial-time algorithm for ensuring controllability of such systems. We then investigate selecting a set of up to k input nodes to maximize a performance metric while satisfying a controllability constraint, and prove that this problem is equivalent to submodular maximization with two matroid basis constraints. We develop a randomized algorithm for solving a relaxed, continuous version of the problem with a $(1 - 1/e)$ optimality bound, which can be rounded to a feasible input set that satisfies controllability. As a third problem, we relax the requirement that the system is controllable and select input nodes based on a trade-off between performance and controllability. In this case, we prove that the problem has the structure of submodular maximization subject to a cardinality constraint, leading to $(1 - 1/e)$ optimality bound.

We next study input selection when the complex network is strongly connected (i.e., there exists a directed path between any two nodes). We prove that, for almost all systems with a given structure, the controllability of the networked system can be represented as a single matroid constraint. Based on this result, we derive a linear-time algorithm for selecting input nodes for controllability in structured linear descriptor systems, and prove that this algorithm is guaranteed to select the minimum-size input set. We further show that the problem of selecting a set of up to k input nodes to optimize performance subject to a controllability constraint can be solved with optimality bound of $(1 - 1/e)$ when the network is strongly connected.

We investigate three special cases of our framework, arising from classes of structured systems that have been studied in the existing literature, namely, linear consensus [17], second-order integrator dynamics [38], and systems in which all nonzero parameters can take arbitrary values [24]. We show that our general approach achieves at least the same optimality guarantees compared to the current state of the art for each individual problem. Our results are illustrated via numerical study in the special case of a consensus network.

1.2. Related Work. Structural controllability of linear systems with given inputs was first studied by Lin in [21], in the case where all nonzero matrix entries are independent free parameters. Controllability of systems with additional relationships between the matrix entries was considered subsequently [9, 26, 37]. In [37], graph-based conditions for structural controllability of descriptor systems were introduced. The work of [26, 28, 27] provided a matroid-based framework for structural controllability of mixed-matrix descriptor systems, containing both fixed and free entries, as well as polynomial-time algorithms for verifying controllability of such systems. For a detailed survey of controllability results in linear descriptor systems, see [10]. In these works, conditions and algorithms for verifying structural controllability with a given input set are provided, but the problem of selecting the input nodes is not considered.

Selecting input nodes to satisfy controllability has been extensively studied in recent years. Necessary and sufficient conditions for a set of input nodes to guarantee controllability in leader-follower consensus dynamics were presented in [43]. Graph-based necessary conditions were derived in [36]. These works considered controllability from a given set of input nodes, but did not introduce efficient algorithms for selecting the input nodes. In [24], a polynomial-time graph matching algorithm was introduced for selecting a minimum-size set of input nodes to satisfy structural controllability. The problem of selecting input nodes for controllability was further considered in [6, 33, 40] for the case where all matrix entries are either zero or are free, independent parameters. For the case where all entries of the system matrices are fully known, input selection algorithms and optimality bounds were derived in [30]. In the present paper, we consider a broader class of system matrices that contain both free parameters and fixed entries, taking the existing works as special cases.

Input selection based on performance criteria has also received research interest. In [5, 8], submodular optimization approaches for selecting input nodes for robustness to noise and smooth convergence to a desired state were developed. A convex relaxation approach for minimizing errors due to link and state noise was proposed in [11, 22, 23]. Combinatorial algorithms for input selection to minimize the H_2-norm, using information centrality, were introduced in [12].

1.3. Organization. The paper is organized as follows. In Section 2, we present our system model, definitions and sufficient conditions for controllability, background on submodularity and matroid theory, and examples of performance metrics that can be incorporated into our framework. In Section 3, we develop our submodular optimization framework for input selection in structured systems. Section 4 discusses input selection in strongly connected network, and shows how connectivity improves the optimality bounds of our approach. Section 5 presents three special cases of our framework found in the existing literature. Section 6 contains a numerical study. Section 7 concludes the paper.

2. Model and Preliminaries. In this section, we describe the system model and definitions of controllability considered. Background on matroid theory and submodularity are given.

2.1. System Model. We consider a linear, time-invariant networked system with a total of n states, where $\mathbf{x}(t) \in \mathbb{R}^n$ is the vector of states at time t. In the absence of any control inputs, the system dynamics are described by

$$F\dot{\mathbf{x}}(t) = A\mathbf{x}(t). \tag{2.1}$$

Eq. (2.1) defines a *linear descriptor system* [10]. The matrices F and A can be further decomposed as $F = Q_F + T_F$ and $A = Q_A + T_A$. The values of Q_F and Q_A are known, fixed parameters. The matrices T_F and T_A are *structure matrices*, in which any nonzero entry can take any arbitrary real value. Dependencies between the free parameters can be enforced through rows of zeros in the F matrix and appropriate values of the fixed parameters; see Section 6 for examples. We assume that the system defined by (2.1) satisfies solvability, defined as follows.

Definition 2.1. *An LTI system of the form (2.1) is* solvable *if for any initial state* $\mathbf{x}(0) \in \mathbb{R}^n$*, there exists a unique trajectory* $\{\mathbf{x}(t) : t > 0\}$ *that satisfies (2.1).*

We now describe the effect of control inputs on the system (2.1). A subset S of states act as control inputs, i.e., for each state $i \in S$, there exists a control input $u_i(t)$ such that $x_i(t) = u_i(t)$ for all t. The states in S correspond to the states of nodes that are controlled directly by an external entity. In the following, without loss of generality, we assume that the state indices are ordered so that states $\mathbf{x}_R = (x_1, x_2, \ldots, x_{n-|S|})$ do not act as control inputs, and states $\mathbf{x}_S = \{x_{n-|S|+1}, \ldots, x_n\}$ act as control inputs.

The system dynamics are then given by

$$\begin{pmatrix} \hat{F} \\ \hline 0 \end{pmatrix} \begin{pmatrix} \dot{\mathbf{x}}_R(t) \\ \dot{\mathbf{x}}_S(t) \end{pmatrix} = \left(\begin{array}{c|c} A_{RR} & A_{RS} \\ \hline 0_{|S|\times(n-|S|)} & -I_{|S|\times|S|} \end{array}\right) \begin{pmatrix} \mathbf{x}_R(t) \\ \mathbf{x}_S(t) \end{pmatrix} + \begin{pmatrix} 0_{(n-|S|)\times|S|} \\ I_{|S|\times|S|} \end{pmatrix} \mathbf{u}(t) \tag{2.2}$$

In (2.2), $\hat{F}$ is the $(n-|S|) \times n$ matrix consisting of the first $(n-|S|)$ rows of F. The matrices A_{RR} and A_{RS} consist of the first $(n-|S|)$ and last $|S|$ columns of the first $(n-|S|)$ rows of A, respectively. The vector $\mathbf{u}(t)$ is the control input signal. The last $|S|$ rows of the equation enforce the condition that $x_i(t) = u_i(t)$ for all $i \in S$.

As a notation, for a square matrix $X \in \mathbb{R}^{n\times n}$, $N(X)$ is a graph with n vertices, where there exists an edge (i, j) if X_{ji} is nonzero.

We now present the definition of controllability considered in this work, which can be found in more detail in [37]. For general systems of the form

$$F\dot{\mathbf{x}}(t) = A\mathbf{x}(t) + B\mathbf{u}(t) \tag{2.3}$$

controllability is defined as follows.

Definition 2.2. *The system (2.2) is* controllable *if, for any admissible initial state* $\mathbf{x}_0$ *and any reachable final state* $\mathbf{x}^*$ *and time* $t > 0$*, there exists a control signal* $\{\mathbf{u}(t') : t' \in [0, t]\}$ *such that* $\mathbf{x}(0) = \mathbf{x}_0$ *and* $\mathbf{x}(t) = \mathbf{x}^*$.

Motivated by [27], we assume that every subdeterminant of $(Q_A - sQ_F|Q_B)$ is a monomial in s. Structural controllability is defined as follows.

Definition 2.3. *The system (2.3) is* structurally controllable *if there exist values for the free parameter matrices* T_F*,* T_A*, and* T_B *such that (2.3) is controllable.*

Structural controllability holds if the system (2.2) is controllable for almost any choice of the free parameters. Note that, if $Q_A = Q_F = 0$, then Definition 2.3 reduces to that of [24]. A matrix pencil interpretation is given by the following theorem.

Theorem 2.4 ([26]). *The system (2.3) is structurally controllable if and only if there exist free parameter matrices* T_F*,* T_A*, and* T_B *such that the system is solvable* *and the following conditions hold:*

$$rank(A|B) = n \tag{2.4}$$

$$rank((A - zF)|B) = n \quad \text{for all } z \in \mathbb{C} \setminus \{0\} \tag{2.5}$$

It remains to find equivalent or sufficient conditions for (2.4) and (2.5). In Section 2.3, we give a graph construction, introduced in [26], that will be used to derive sufficient conditions for (2.5).

2.2. Matroids and Submodularity. In what follows, we define the concepts of matroids and submodular functions, which will be used in our optimization framework. All definitions and lemmas in this subsection can be found in [31].

Definition 2.5. *A matroid $\mathcal{M} = (V, \mathcal{I})$ is defined by a finite set V (denoted the ground set) and a collection $\mathcal{I}$ of subsets of V such that (a) $\emptyset \in \mathcal{I}$, (b) if $X \subseteq Y$ and $Y \in \mathcal{I}$, then $X \in \mathcal{I}$, and (c) if $X, Y \in \mathcal{I}$ and $|X| < |Y|$, then there exists $v \in Y \setminus X$ such that $(X \cup \{v\}) \in \mathcal{I}$. The collection $\mathcal{I}$ is denoted as the collection of independent sets of $\mathcal{M}$.*

A maximal independent set is a *basis*; we let $\mathcal{B}(\mathcal{M})$ denote the set of bases of a matroid $\mathcal{M}$. The *rank function* ρ of a matroid is a function $\rho : 2^V \rightarrow \mathbb{Z}_{\geq 0}$, given by $\rho(X) = \max\{|Y| : Y \subseteq X, Y \in \mathcal{I}\}$. The *rank* of a matroid is equal to $\rho(V)$.

A simple example of a matroid is the uniform matroid U_k, defined by $X \in U_k$ if $|X| \leq k$ for some k. In a linear matroid, the set V is equal to a collection of vectors in $\mathbb{R}^m$, and a set of vectors is independent if the vectors are linearly independent. In this case, the rank function is equal to the column rank of the matrix defined by the vectors. The following is a method of constructing matroids that will be used in this paper.

Definition 2.6. *Let $\mathcal{M}_1 = (V_1, \mathcal{I}_1)$ and $\mathcal{M}_2 = (V_2, \mathcal{I}_2)$ be matroids. Then the* matroid union *$\mathcal{M} = \mathcal{M}_1 \vee \mathcal{M}_2$ is defined by $V = V_1 \cup V_2$ and $X \in \mathcal{I}$ if $X = X_1 \cup X_2$ with $X_1 \in \mathcal{I}_1$, $X_2 \in \mathcal{I}_2$.*

It can be shown that the matroid union $\mathcal{M}$ is a matroid, with rank function $\rho(X) = \min\{\rho_1(Y) + \rho_2(Y) + |X \setminus Y| : Y \subseteq X\}$. A second matroid construction is the *dual matroid*, described as follows.

Definition 2.7. *Let $\mathcal{M} = (V, \mathcal{I})$ be a matroid. The dual of $\mathcal{M}$, denoted $\mathcal{M}^*$, has ground set V and set of independent sets $\mathcal{I}^*$ given by $\mathcal{I}^* = \{X' \subseteq X : V - X$ is a basis of $\mathcal{M}\}$. If ρ is the rank function of $\mathcal{M}$, then the rank function ρ^* of $\mathcal{M}^*$ is given by $\rho^*(X) = \rho(V - X) + |X| - \rho(V)$.*

We now define the concept of a *submodular function*.

Definition 2.8. *Let V be a finite set. A function $f : 2^V \rightarrow \mathbb{R}_{\geq 0}$ is* submodular *if for any sets A and B with $A \subseteq B$ and any $v \in V \setminus B$,*

$$f(A \cup \{v\}) - f(A) \geq f(B \cup \{v\}) - f(B).$$

2.3. Auxiliary Graph Construction. We consider the following auxiliary graph constructed from (2.3). Define a matrix

$$\Omega = \begin{array}{r} \mathbf{w}: \\ \mathbf{x}: \\ \mathbf{u}: \end{array} \begin{pmatrix} Q_A - Q_F & Q_B \\ I & 0 \\ 0 & I \end{pmatrix} \tag{2.6}$$

Here, $\mathbf{w}$, $\mathbf{x}$, and $\mathbf{u}$ denote indices, so that the first n rows are indexed $w_1, \dots, w_n$, the second n rows are indexed $x_1, \dots, x_n$, and the third k rows are indexed $u_1, \dots, u_k$.

As an intermediate step in the construction, define a bipartite graph H with vertex set

$$V_H = \{w_1^T, \dots, w_n^T\} \cup \{x_1^Q, \dots, x_n^Q\} \cup \{w_1^Q, \dots, w_n^Q\}$$

and edge set

$$E_H = \{(w_i^T, w_i^Q) : i = 1, \dots, n\} \cup \{(w_i^T, x_j^Q) : (i,j) \in N(T_A) \cup N(T_F)\}.$$

The following lemma gives properties of matchings in this bipartite graph.

Lemma 2.9 ([26], Theorem 4.7). *If the system (2.3) is solvable, then there exists a perfect matching m on the graph H (i.e., a matching in which all nodes in $\{w_1^T, \dots, w_n^T\}$ are matched) such that the rows indexed in the set*

$$J = \{w_i : m(w_i^T) = w_i^Q\} \cup \{x_i : m(w_j^T) = x_i^Q \text{ for some } w_j\}$$

are linearly independent in Ω. This matching can be computed in polynomial time.

Let J be a set satisfying the conditions of Lemma 2.9 with matching m, and let Ω_J be a submatrix of Ω obtained from the rows in J. We have that Ω_J is an $n \times (n+k)$ matrix with full rank, and hence we can find rows J_1 such that $\Omega_{J \cup J_1}$ is a full-rank $(n+k) \times (n+k)$ matrix. Let $\tilde{\Omega} = \Omega \Omega_{J \cup J_1}^{-1}$. We index the columns of $\tilde{\Omega}$ in the set $J \cup J_1$.

We now define the auxiliary graph $\hat{G} = (\hat{V}, \hat{E})$ using the matrix $\tilde{\Omega}$. The vertex set is equal to

$$\begin{aligned}\hat{V} = \{w_i^T : i = 1, \dots, n\} \cup \{w_i^Q : i = 1, \dots, n\} \cup \{x_i^T : i = 1, \dots, n\} \cup \{x_i^Q : i = 1, \dots, n\} \\ \cup \{u_i^T : i = 1, \dots, k\} \cup \{u_i^Q : i = 1, \dots, k\},\end{aligned}$$

while the edge set is given by

$$\begin{aligned}\hat{E} = &\{(w_i^T, x_j^Q) : (i,j) \in N(T_A) \cup N(T_F), m(w_i^T) \neq x_j^Q\} \\ &\cup \{(x_j^Q, w_i^T) : (i,j) \in N(T_A) \cup N(T_F), m(w_i^T) = x_j^Q\} \\ &\cup \{(w_i^T, w_i^Q) : m(w_i^T) \neq w_i^Q\} \cup \{(w_i^Q, w_i^T) : m(w_i^T) = m(w_i^Q)\} \\ &\cup \{(w_i^T, u_j^Q) : (i,j) \in N(T_B)\} \cup \{(u_j^T, u_j^Q) : j = 1, \dots, k\} \\ &\cup \{(x^Q, y^Q) : x \in \hat{V} \setminus J, y \in J, \tilde{\Omega}_{xy} \neq 0, \tilde{\Omega}_{xz} = 0 \forall z \in J_1\}\end{aligned}$$

Based on this graph construction, the following sufficient condition for $\text{rank}(zF - A|B) = n$ for all nonzero z can be derived.

Lemma 2.10. *Let*

$$S_- = \{v^Q : \tilde{\Omega}_{vj} \neq 0 \text{ for some } j \in J_1\}.$$

Let V'' denote the set of nodes in $\hat{V}$ that are part of a cycle. If all nodes in V'' are connected to S_- in the graph $\hat{G}$, then the condition rank$(zF - A|B) = n$ holds for all nonzero $z \in \mathbb{C}$ and almost any values of the free parameters.

A proof is given in the appendix.

2.4. Performance Metrics. The optimality guarantees for the input selection algorithms presented in this work are applicable to monotone submodular performance metrics. The first metric is the *network coherence*, defined as follows.

DEFINITION 2.11. *Consider the node dynamics $\dot{x}_i(t) = -\sum_{j \in N(i)} (x_i(t) - x_j(t)) + w_i(t)$, where $w_i(t)$ is a zero-mean white process with autocorrelation function $W(\tau) = \delta(\tau)$. The network coherence $f(S)$ from input set S is the mean-square deviation in the node state from consensus in steady-state, equal to*

$$\lim_{t \rightarrow \infty} \mathbf{E}||\mathbf{x}(t)||_2^2.$$

The network coherence was defined in [34] and shown to be a supermodular function of the input set in [8]. The second metric is the *convergence error*.

DEFINITION 2.12. *Consider the node dynamics $\dot{x}_i(t) = -\sum_{j \in N(i)} W_{ij}(x_i(t) - x_j(t))$, where the edge weights W_{ij} are nonnegative. The convergence error at time t is defined as $||\mathbf{x}(t) - x^*\mathbf{1}||_p$, for $p \in [1, \infty)$, where x^* is the state of the input nodes.*

The convergence error was proven to be a supermodular function of the input set in [5]. Other examples of submodular functions include the information gathered by a set of input nodes [19] and the trace of the controllability Gramian [42].

3. Problem Formulation - Input Selection for Performance and Controllability. In this section, we present our submodular optimization framework for selecting input nodes based on performance and controllability. In order to provide computational tractability, we first map the sufficient conditions of Theorem 2.4 to matroid constraints on the input set. We present algorithms for selecting an input set to guarantee controllability. We then formulate the problem of selecting a set of up to k input nodes to maximize a performance metric while satisfying controllability. We prove that the problem is a submodular maximization problem with two matroid basis constraints, and present efficient approximation algorithms. For the case where the number of input nodes may not be sufficient to guarantee controllability, we introduce a graph controllability index and formulate the problem of selecting input nodes based on a trade-off between performance and controllability.

3.1. Mapping controllability to matroid constraints. We derive matroid constraints for the set of non-input nodes that are equivalent to or sufficient for the conditions of Theorem 2.4. As a first step, we develop an equivalent representation of the dynamics (2.2), and prove that structural controllability of (2.2) is equivalent to structural controllability of the equivalent dynamics.

LEMMA 3.1. *Define the dynamics*

$$F\dot{\mathbf{x}} = \left(\begin{array}{c|c} A_{RR} & A_{RS} \\ \hline A_{SR} & A_{SS} \end{array}\right) \left(\begin{array}{c} \mathbf{x}_R(t) \\ \mathbf{x}_S(t) \end{array}\right) + \left(\begin{array}{c} 0 \\ T_B \end{array}\right) \mathbf{u}(t) \tag{3.1}$$

where T_B is a diagonal matrix where the diagonal entries are free parameters. Then the system (2.2) is structurally controllable if and only if (3.1) is structurally controllable.

Proof. Suppose that the system (3.1) is controllable with $(T_B)_{ii} = \alpha_i$ for some real α_i's. For a given initial state $\mathbf{x}_0$ and desired state $\mathbf{x}^*$, suppose that there exists a set of control inputs $u_1(t), \ldots, u_k(t)$ such that $\mathbf{x}(t) = \mathbf{x}^*$. We have

$$\alpha_i u_i(t) + \sum_j A_{ij} x_j(t) = \sum_j F_{ij} \dot{x}_j(t),$$

which is equivalent to

$$\alpha_i u_i(t) + \sum_j A_{ij} x_j(t) - \sum_j F_{ij} \dot{x}_j(t) + x_i(t) = x_i(t). \tag{3.2}$$

Rearranging terms implies that $x_i(t) = \hat{u}_i(t)$, where $\hat{u}_i(t)$ is the left-hand side of (3.2). Hence using $\hat{u}_i(t)$ as the input signal implies that structural controllability is achieved for the dynamics (2.2) as well. The proof of the converse is similar. □

Lemma 3.1 implies that it suffices to consider the conditions of Theorem 2.4 under the equivalent system (3.1). In Lemmas 3.2–3.4, we derive a sufficient matroid constraint for Eq. (2.4). We first define a matroid constraint on the set of non-input nodes that is equivalent to (2.4), starting with a preliminary lemma.

Lemma 3.2 ([26]). *Let $\mathcal{M}(I|Q_A|Q_B)$ denote the linear matroid defined by the fixed parameter matrices Q_A and Q_B, and let $\mathcal{M}(I|T_A|T_B)$ denote the matroid defined by the free parameter matrices T_A and T_B. Then rank$(A|B) = n$ iff rank$(\mathcal{M}(I|Q_A|Q_B) \vee \mathcal{M}(I|T_A|T_B)) = 2n$.*

For the system (3.1), the condition $\text{rank}[\mathcal{M}(I|Q_A|Q_B) \vee \mathcal{M}([I|T_A|T_B])] = 2n$ of Lemma 3.2 is given by

$$\text{rank}[\mathcal{M}([I|Q_A|0]) \vee \mathcal{M}([I|T_A|T_B(S)])] = 2n,$$

where $T_B(S)$ is a diagonal matrix with a free parameter in the i-th diagonal entry for $i \in S$ and zeros elsewhere. In order to establish a matroid constraint for Lemma 3.2 we have the following lemma.

Lemma 3.3. *The function*

$$\rho_1(S) = \text{rank}(\mathcal{M}([I|Q_A|0]) \vee \mathcal{M}([I|T_A|T_B(S)])) - \text{rank}(\mathcal{M}([I|Q_A]) \vee \mathcal{M}([I|T_A]))$$

is a matroid rank function. The rank condition rank$(A|B) = n$ holds for input set S if and only if $\rho_1(S) = 2n - \text{rank}(\mathcal{M}([I|Q_A|0]) \vee \mathcal{M}([I|T_A|0]))$.

Proof. Define the set $X = \{1, \ldots, 2n\}$. The function $\rho_1(S)$ is equivalent to $\rho_1(S) = \text{rank}(S \cup X) - \text{rank}(X)$, which is equal to the rank function of the contraction of the matroid $\mathcal{M}([I|Q_A|0]) \vee \mathcal{M}([I|T_A|T_B])$ by the set X. Hence $\rho_1(X)$ defines the rank function of a matroid. By Lemma 3.2, the condition $\text{rank}([A|B]) = n$ is satisfied if and only if $\text{rank}(\mathcal{M}([I|Q_A|0]) \vee \mathcal{M}([I|T_A|T_B(S)])) = 2n$. Combining with the definition of $\rho_1(S)$ completes the proof. □

Lemma 3.3 implies that the condition (2.4) can be expressed as a constraint on a matroid rank function, $\rho_1(S)$. We let $\mathcal{M}_1$ denote the matroid with rank function defined by $\rho_1(S)$ from Lemma 3.3. Finally, we express $\text{rank}([A|B]) = n$ as a matroid constraint on the set of non-input nodes.

Lemma 3.4. *The condition rank$([A|B]) = n$ holds if and only if the set of non-input nodes R satisfies $R \in \mathcal{M}_1^*$, the dual of the matroid induced by rank function $\rho_1(S)$.*

Proof. Let ρ_1^* denote the rank function of the dual matroid $\mathcal{M}_1^*$. We have $\rho_1^*(R) = \rho_1(V \setminus R) + |R| - \rho_1(V)$, which is equivalent to $\rho_1(S) = \rho_1^*(R) - |R| + \rho_1(V)$. The constraint of Lemma 3.3 is equivalent to $\rho_1(S) \geq \rho_1(V)$, which is in turn equivalent to $\rho_1^*(R) \geq |R|$. This, however, holds only when $R \in \mathcal{M}_1^*$. □

We now turn to the constraint that $\text{rank}(A - zF|B) = n$ (Eq. (2.5)), which we map to a matroid constraint in Lemmas 3.5–3.8. We formulate the matroid constraint by identifying a sufficient graph-based condition (Lemmas 3.5 and 3.6). We then prove that this graph-based condition is equivalent to a partition matroid constraint

(Lemmas 3.7 and 3.8). The following intermediate lemma is the first step in our approach.

LEMMA 3.5. *There exists a graph $G' = (V', E')$, which can be constructed in polynomial time, such that the auxiliary graph $\hat{G} = (\hat{V}, \hat{E})$ corresponding to system (3.1) is given by*

$$\hat{V} = V' \cup \{u_1^T, \ldots, u_k^T\} \cup \{u_1^Q, \ldots, u_k^Q\} \tag{3.3}$$

$$\hat{E} = E' \cup \{(u_i^T, u_i^Q) : i = 1, \ldots, k\} \cup \{(w_i^T, u_i^T) : i \in S\} \tag{3.4}$$

In this graph, the set $J_1 = \{u_1^Q, \ldots, u_k^Q\}$ and the condition of Lemma 2.10 is satisfied if each node of V' that belongs to a cycle in G' is reachable to a node in $\{w_i^T : i \in S\}$ in the graph $\hat{G}$.

Proof. The matrix Ω corresponding to (3.1) is given by

$$\Omega = \left(\begin{array}{cc} Q_A - Q_F & 0 \\ I & 0 \\ 0 & I \end{array} \right)$$

By solvability of (3.1) and Lemma 2.9, we can select n linearly independent rows J from the first $2n$ rows of Ω. Let Ψ denote the matrix consisting of these linearly independent rows. The matrix Ω_J can be completed to a full-rank matrix by selecting $J_1 = \{u_1, \ldots, u_k\}$, giving

$$\Omega_{J \cup J_1} = \left(\begin{array}{cc} \Psi & 0 \\ 0 & I \end{array} \right), \quad \Omega_{J \cup J_1}^{-1} = \left(\begin{array}{cc} \Psi^{-1} & 0 \\ 0 & I \end{array} \right).$$

Note that the matrix $\Omega\Omega_{J \cup J_1}^{-1}$ does not depend on the input set S. Let G' denote the auxiliary graph when $S = \emptyset$. By construction, the auxiliary graph $\hat{G}$ with a non-empty input set S is given by (3.3) and (3.4), since adding nodes to S simply adds edges to $N(T_B)$.

To prove that reachability to the nodes $\{w_i^T : i \in S\}$ is sufficient, note that it suffices that each node is reachable to $J_1 = \{u_1^Q, \ldots, u_k^Q\}$ by Lemma 2.10. Since each node in $\{w_i^T : i \in S\}$ is reachable to J_1, it suffices for all other nodes to be reachable to $\{w_i^T : i \in S\}$. $\square$

As a consequence of Lemmas 2.10 and 3.5, to ensure that $\text{rank}(zF - A|B) = n$, it suffices to select an input set S such that each node in G' is reachable to $\{w_i^T : i \in S\}$. In order to select such an input set, we define an equivalence relation $\sim$ on the nodes in V' as $i \sim j$ if node i is path-connected to node j in G' and vice versa. We let $[i] = \{j : i \sim j\}$, and define $\overline{V} = \{[i] : i \in \hat{V}\}$ (so that $\overline{V}$ is the quotient set of $\hat{V}$ under the relation $\sim$).

Define the graph $\overline{G} = (\overline{V}, \overline{E})$ by $(i, j) \in \overline{E}$ if there exists $i' \in [i]$ and $j' \in [j]$ such that $(i', j') \in E$. Note that $\overline{G}$ is a directed acyclic graph, in which each node represents one connected component of the graph G'. We let $\overline{V}'$ denote the set of isolated nodes, i.e., nodes that have no incoming edges in $\overline{G}$. The following lemma gives an equivalent condition for connectivity in $\hat{G}$ based on the graph $\overline{G}$.

LEMMA 3.6. *All nodes in $\hat{V}$ are connected to S iff for each $[i] \in \overline{V}'$, $\{w_j^T : j \in S\} \cap [i] \neq \emptyset$.*

Proof. We first show that if $\{w_j^T : j \in S\} \cap [i] \neq \emptyset$ for all $[i] \in \overline{V}'$, then all nodes in $\hat{V}$ are connected to S. Let $v \in \hat{V}$. If $v \in [i]$ with $[i] \in \overline{V}'$, then there exists

$j \in \{w_j^T : j \in S\} \cap [i]$ such that j is path-connected to v. If $v \in [i]$ with $[i] \notin \overline{V}'$, then there are nodes $i' \in [i]$ and $j_1 \in [j_1]$ for some $[j_1] \in \overline{V}$ such that $(j_1, i') \in \hat{E}$. Now, either $[j_1] \in \overline{V}'$ or there exists $j_1' \in [j_1]$ and $j_2 \in [j_2]$ such that $(j_2, j_1') \in \hat{E}$. Since the graph $\overline{G}$ is acyclic, there exists a sequence of components $[i], [j_1], \ldots, [j_L]$, with $([j_{l+1}], [j_l]) \in \hat{E}$ and $[j_L] \in \overline{V}'$. This sequence of components defines a directed path from a node $v' \in S$ to v.

Now, suppose that all nodes in $\hat{V}$ are connected to S. For each $v \in [i]$, with $[i] \in \overline{V}'$, v must be connected to at least one input node. Since $[i] \in \overline{V}'$, only other nodes in $[i]$ are connected to v. Hence we must have $\{w_j^T : j \in S\} \cap [i] \neq \emptyset$. ☐

Lemma 3.6 enables us to express the connectivity criterion as a matroid constraint. First, define a function $\rho_2(S) = |\{[i] \in \overline{V}' : [i] \cap S \neq \emptyset\}|$. The following lemma describes the rank condition in terms of function $\rho_2(S)$.

Lemma 3.7. *Let $c = |\overline{V}'|$. The function $\rho_2(S)$ is a matroid rank function, and all nodes in $\hat{V}$ are connected to S iff $\rho_2(S) = c$.*

Proof. The function $\rho_2(S)$ is a matroid rank function because $\rho_2(\emptyset) = 0$ and $(\rho_2(S \cup \{v\}) - \rho_2(S)) \in \{0, 1\}$, with $\rho_2(S \cup \{v\}) - \rho_2(S) = 1$ iff there exists i such that $v \in [i]$ and $\{w_j^T : j \in S\} \cap [i] = \emptyset$. Furthermore, $\rho_2(S) = c$ if and only if for every $[i] \in \overline{V}'$, $\{w_j^T : j \in S\} \cap [i] \neq \emptyset$, which is exactly the condition of Lemma 3.6. ☐

Let $\mathcal{M}_2$ denote the matroid induced by rank function $\rho_2(S)$. Since $\overline{V}'$ defines a collection of disjoint subsets of $\hat{V}$, and a set S is independent in $\mathcal{M}_2$ if it contains at most one node in each subset, $\mathcal{M}_2$ is a partition matroid. We are now ready to state a sufficient (but not necessary) matroid constraint on R for the condition (2.5).

Lemma 3.8. *Let $\mathcal{M}_2^*$ be the dual of the matroid induced by $\rho_2(S)$. If the non-input nodes R satisfy $R \in \mathcal{M}_2^*$, then the condition rank$((zF - A)|B) = n$ is satisfied.*

Proof. The rank function $\rho_2^*(R)$ of $\mathcal{M}_2^*$ can be written as $\rho_2^*(R) = \rho_2(V \setminus R) + |R| - \rho_2(V) = \rho_2(S) + |R| - c$, or equivalently, $\rho_2(S) = \rho_2^*(R) + c - |R|$. Hence $\rho_2(S) = c$ is equivalent to $\rho_2^*(R) = |R|$, which holds if and only if $R \in \mathcal{M}_2^*$. ☐

We combine the results of Lemmas 3.4 and 3.8 to yield the following theorem, which gives sufficient but not necessary conditions for controllability from a given input set.

Theorem 3.9. *If $R \in \mathcal{M}_1^* \cap \mathcal{M}_2^*$, then the system is controllable from input set $S = V \setminus R$.*

Having defined matroid-based sufficient conditions for controllability, we will next formulate the problem of selecting an input set to guarantee structural controllability. The selected input set will be the minimum-size set that satisfies the conditions of Theorem 3.9, although there may be input sets with smaller cardinality satisfying controllability due to the non-necessity of the conditions in Theorem 3.9.

3.2. Minimum-Size Input Set Selection for Structural Controllability. Selecting a minimum-size set S to satisfy structural controllability is equivalent to selecting a maximum-size set of non-input nodes $R = V \setminus S$ such that controllability is satisfied. Based on Theorem 3.9, the problem of selecting a set of input nodes to guarantee structural controllability can be formulated as

$$\text{maximize}\, \{|R| : R \in \mathcal{M}_1^*, R \in \mathcal{M}_2^*\}\,. \tag{3.5}$$

Eq. (3.5) is a matroid intersection problem, which can be solved in time $O(n^{5/2}\tau)$, where τ is the time required to test if a set R is in $\mathcal{M}_1^*$ and $\mathcal{M}_2^*$ [41], which can be performed in polynomial time.

Algorithm 1 gives a polynomial-time procedure for solving (3.5) using the maximum cardinality matroid intersection algorithm of [41, Ch. 41]. Lemmas 3.3 and 3.8 and the above discussion generalize the main result of [24] from free matrices to systems with a mix of free and fixed parameters.

Algorithm 1 Algorithm for selecting a set of input nodes to ensure structural controllability by solving (3.5).

procedure MIN_CONTROLLABLE_SET($\mathcal{M}_1^*$, $\mathcal{M}_2^*$)
  Input: Matroids $\mathcal{M}_1^*$ and $\mathcal{M}_2^*$
  Output: Set of inputs S
  $R \leftarrow \emptyset$
  while 1 **do**
    $E_{\mathcal{M}_1^*,\mathcal{M}_2^*}(R) \leftarrow \emptyset$
    for All $i \in R$, $j \notin R$ **do**
      if $(R - \{i\} \cup \{j\}) \in \mathcal{M}_1^*$ **then**
        $E_{\mathcal{M}_1^*,\mathcal{M}_2^*}(R) \leftarrow E_{\mathcal{M}_1^*,\mathcal{M}_2^*}(R) \cup \{(i,j)\}$
      end if
      if $(R - \{i\} \cup \{j\}) \in \mathcal{M}_2^*$ **then**
        $E_{\mathcal{M}_1^*,\mathcal{M}_2^*}(R) \leftarrow E_{\mathcal{M}_1^*,\mathcal{M}_2^*}(R) \cup \{(j,i)\}$
      end if
    end for
    $D_{\mathcal{M}_1^*,\mathcal{M}_2^*}(R) \leftarrow$ directed graph with vertex set V and edge set $E_{\mathcal{M}_1^*,\mathcal{M}_2^*}(R)$
    $X_1 \leftarrow \{j \in V \setminus R : (R \cup \{j\}) \in \mathcal{M}_1^*\}$
    $X_2 \leftarrow \{j \in V \setminus R : (R \cup \{j\}) \in \mathcal{M}_2^*\}$
    if path exists from a node in X_1 to a node in X_2 **then**
      $P \leftarrow$ shortest X_1-X_2 path
      $R \leftarrow R \Delta P$
    else
      break
    end if
  end while
  $S \leftarrow V \setminus R$
  return S
end procedure

3.3. Input Selection for Joint Performance and Controllability. We now consider the problem of maximizing a monotone performance metric $f(S)$ while satisfying controllability with a set of up to k input nodes. Based on Theorem 3.9, the problem formulation is given by

$$\begin{array}{ll} \text{maximize}_{S \subseteq V} & f(S) \\ \text{s.t.} & |S| \leq k \\ & (V \setminus S) \in \mathcal{M}_1^* \\ & (V \setminus S) \in \mathcal{M}_2^* \end{array} \tag{3.6}$$

Eq. (3.6) is a combinatorial optimization problem, making it NP-hard to solve in the general case. The following lemma describes an equivalent formulation to (3.6).

LEMMA 3.10. *Define* $r_1 = \mathit{rank}(\mathcal{M}_1)$ *and* $r_2 = \mathit{rank}(\mathcal{M}_2)$. *Let* $\hat{\mathcal{M}}_1 = \mathcal{M}_1 \vee U_{k-r_1}$ *and* $\hat{\mathcal{M}}_2 = \mathcal{M}_2 \vee U_{k-r_2}$, *with* $\hat{\mathcal{B}}_1$ *and* $\hat{\mathcal{B}}_2$ *denoting the sets of bases of* $\hat{\mathcal{M}}_1$ *and* $\hat{\mathcal{M}}_2$, *respectively. Let S^* denote the optimal solution to the problem*

$$\begin{array}{lll} (3.7) & \textit{maximize}_{S \subseteq V} & f(S) \\ & \textit{s.t.} & S \in \hat{\mathcal{B}}_1 \cap \hat{\mathcal{B}}_2 \end{array}$$

The set S^* is the optimal solution to (3.6).

Proof. First, we have that the optimal solution to (3.6) satisfies $|S| = k$. If not, then since $f(S)$ is monotone, we can add elements to S and increase the value of f without violating the constraints $(V \setminus S) \in \mathcal{M}_1^*$ and $(V \setminus S) \in \mathcal{M}_2^*$. Furthermore, if $R = (V \setminus S) \in \mathcal{M}_1^*$, then R can be completed to a basis $\overline{R}$ of $\mathcal{M}_1^*$, and we have $(V \setminus \overline{R}) \subseteq (V \setminus R)$, implying that $S = (V \setminus \overline{R}) \cup \overline{S}$ for some set $\overline{S}$ with $|\overline{S}| = k - r_1$. Since $\overline{R}$ is a basis of $\mathcal{M}_1^*$, $V \setminus \overline{R}$ is a basis of $\mathcal{M}_1$, and hence $S \in \hat{\mathcal{B}}_1$. A similar result holds for $\hat{\mathcal{M}}_2$.

Now, consider $S^* \in \hat{\mathcal{B}}_1 \cap \hat{\mathcal{B}}_2$. By construction $\text{rank}(\hat{\mathcal{M}}_1) = \text{rank}(\hat{\mathcal{M}}_2) = k$. We can therefore write $S^* = S_1^* \cup \hat{S}_1^*$, where S_1^* is a basis of $\mathcal{M}_1$ and $|\hat{S}_1^*| = k - r_1$. Since S_1^* is a basis of $\mathcal{M}_1$, we have $(V \setminus S_1^*) \in \mathcal{M}_1^*$, and thus the set $R = (V \setminus S^*) \subseteq (V \setminus S_1^*) \in \mathcal{M}_1^*$, implying that $R = (V \setminus S^*) \in \mathcal{M}_1^*$. A similar argument implies that $(V \setminus S^*) \in \mathcal{M}_2^*$. ∎

By Lemma 3.10, solving (3.6) is equivalent to solving (3.7). We present a two-stage algorithm for approximating (3.7). In the first stage, the algorithm solves a relaxed, continuous version of the problem. In the second stage, the algorithm rounds the solution to an integral value satisfying the constraints of (3.7).

The algorithm is defined as Algorithm 2 below. It contains two subroutines, namely, MAX_WEIGHTED_BASIS and SWAP_ROUND. The subroutine MAX_WEIGHTED_BASIS takes as input two matroids $\mathcal{M}'$ and $\mathcal{M}''$ with the same ground set V (with $|V| = n$), as well as a weight vector $\boldsymbol{\alpha} \in \mathbb{R}^n$, and outputs a set $I \in \mathcal{B}(\mathcal{M}') \cap \mathcal{B}(\mathcal{M}'')$ such that $\sum_{i \in I} \alpha_i$ is maximized (provided at least one common basis exists). Polynomial-time algorithms for finding such sets are well-known [41, Ch. 43].

The subroutine SWAP_ROUND takes as input a vector $\mathbf{r}$ in the common base polytope of two matroids $\mathcal{M}'$ and $\mathcal{M}''$, and outputs a set $I \in \mathcal{B}(\mathcal{M}') \cap \mathcal{B}(\mathcal{M}'')$. The algorithm is randomized with the output satisfying $\mathbf{E}(f(I)) \geq F(\mathbf{r})$. The swap round algorithm was proposed in [4].

In Algorithm 2, the $\mathbf{1}(I(t))$ denotes the incidence vector of set $I(t)$, which has a 1 in the i-th entry if $i \in I$ and 0 otherwise. The following theorem describes the optimality bound of Algorithm 2.

THEOREM 3.11. *Algorithm 2 runs in polynomial time with complexity $O(\tau n^5)$. Letting S^* denote the optimal solution to (3.6), the vector $\mathbf{y}(1)$ returned by the continuous relaxation satisfies $F(\mathbf{y}(1)) \geq (1 - 1/e) f(S^*)$, where F is the multilinear relaxation of $f(S)$. The rounded solution S is a feasible solution to (3.6).*

The proof is given in the appendix. The following theorem provides additional optimality guarantees when the objective function $f(S)$ is linear.

THEOREM 3.12. *If the function $f(S)$ is of the form $f(S) = \sum_{i \in S} \tau_i$ for some real-valued weights $\tau_1, \ldots, \tau_n$, then the solution to (3.6) can be obtained in polynomial time.*

Proof. If the function $f(S)$ is of the form $f(S) = \sum_{i \in S} \tau_i$, then (3.6) is equivalent to maximizing a modular function subject to two matroid basis constraints. For problems of this form, the Edmonds weighted matroid intersection algorithm provides an optimal solution in polynomial time [41]. ∎

Algorithm 2 Input selection algorithm for joint performance and controllability.

procedure INPUT_SELECT(f, $\hat{\mathcal{M}}_1$, $\hat{\mathcal{M}}_2$, k)
  Input: Monotone submodular objective function $f : 2^V \rightarrow \mathbb{R}$
    Matroids $\hat{\mathcal{M}}_1$, $\hat{\mathcal{M}}_2$
    Maximum number of inputs k
  Output: Set of inputs S
  $\delta \leftarrow \frac{1}{9k^2}$, $t \leftarrow 0$, $\mathbf{y}(0) \leftarrow \mathbf{0}$
  while $t < 1$ **do**
    $R(t)$ contains each $j \in V$ independently with probability $y_j(t)$
    for $j \in V$ **do**
      $\omega_j(t) \leftarrow \mathbf{E}[f(R(t) \cup \{j\}) - f(R(t))]$
    end for
    $I(t) \leftarrow \text{MAX\_WEIGHTED\_BASIS}(\hat{\mathcal{M}}_1, \hat{\mathcal{M}}_2, \boldsymbol{\omega})$
    $\mathbf{y}(t+\delta) \leftarrow \mathbf{y}(t) + \delta \cdot \mathbf{1}(I(t))$
    $t \leftarrow (t + \delta)$
  end while
  $S \leftarrow \text{SWAP\_ROUND}(\mathbf{y}(1), \hat{\mathcal{M}}_1, \hat{\mathcal{M}}_2)$
  return S
end procedure

3.4. Selecting Input Nodes for Performance-Controllability Trade-Off. In this section, we study input selection based on a trade-off between performance and controllability, instead of treating controllability as a constraint that must be satisfied. This maximization may be beneficial when the number of input nodes k is insufficient to guarantee controllability.

We first introduce two graph controllability indices (GCIs) that can be traded off with a performance metric in order to maximize the level of performance and controllability. The first controllability index $c_1(S)$ is given by

$$c_1(S) = \max\{|V'| : \mathrm{rank}(A(V')|B(V')) = |V'|\}, \tag{3.8}$$

where $A(V')$ and $B(V')$ are sub-matrices of A and B consisting of the rows and columns indexed in V'. Intuitively, $c_1(S)$ is the size of the largest subgraph of V such that the zero modes of all nodes in the subgraph are controllable. The second GCI quantifies the controllability of the nonzero modes, characterized by the constraint $\mathrm{rank}(zF - A|B) = n$. We define $c_2(S)$ by

$$c_2(S) = |\{i \in V : i \text{ is reachable to } \mathbf{u} \text{ in } \hat{G}\}| \tag{3.9}$$

where $\hat{G}$ is defined as in Section 2.3. The function $c_2(S)$ quantifies the number of nodes that are reachable to the input nodes, and hence satisfy controllability of the nonzero modes. If $c_1(S) + c_2(S) = 2n$, then both the zero and nonzero modes of all nodes are controllable, and hence controllability is satisfied. Otherwise, the problem of joint maximization of performance and controllability can be formulated as

$$\begin{array}{ll} \text{maximize}_{S \subseteq V} & f(S) + \eta(c_1(S) + c_2(S)) \\ \text{s.t.} & |S| \leq k \end{array} \tag{3.10}$$

The trade-off parameter $\eta \geq 0$ is used to vary the relative weight assigned to performance or controllability criteria. When η is small, then nodes are selected for performance alone; at the other extreme, when η is large, nodes are primarily selected to maximize controllability. The following result is the first step in deriving efficient algorithms for solving (3.10).

THEOREM 3.13. *The functions $c_1(S)$ and $c_2(S)$ are submodular in S.*

Proof. The function $c_1(S)$ is equal to the maximum-size set of non-input nodes with controllable zero modes from the input nodes, plus the number of input nodes. This can be written as $c_1(S) = \rho_1(V \setminus S) + |S|$, where ρ_1 is defined as in Lemma 3.3. Since ρ_1 is a matroid rank function, ρ_1 is submodular and hence $\rho_1(V \setminus S)$ is submodular as well. Since the sum of submodular functions is submodular, $\rho_1(V \setminus S) + |S|$ is submodular as a function of S.

It remains to show submodularity of $c_2(S)$. Let $S \subseteq T$, and suppose that $v \notin T$. We have that $c_2(T \cup \{v\}) - c_2(T)$ is equal to the number of nodes that are reachable to v, but not to any node in T. Since $S \subseteq T$, any node that is not reachable to T is automatically not reachable to S. Hence, any node that is reachable to v but not any node in T is also reachable to v but not any node in S, implying that $c_2(T \cup \{v\}) - c_2(T) \geq c_2(S \cup \{v\}) - c_2(S)$. □

A greedy algorithm for approximating (3.10) is as follows. The set S is initialized to be empty, and the algorithm proceeds over k iterations. At the i-th iteration, the element $v \in V$ maximizing $f(S \cup \{v\}) + \eta(c_1(S \cup \{v\}) + c_2(S \cup \{v\}))$ is selected and added to S, terminating after k iterations. The following theorem gives an optimality bound for this algorithm.

THEOREM 3.14. *The solution S obtained by the greedy algorithm satisfies $(f(S) + \eta(c_1(S) + c_2(S))) \geq (1 - 1/e)(f(S^*) + \eta(c_1(S^*) + c_2(S^*)))$, where S^* is the optimal solution to (3.10).*

Proof. Since $f(S)$, $c_1(S)$, and $c_2(S)$ are submodular and monotone as functions of S, the function $f(S) + \eta(c_1(S) + c_2(S))$ is monotone and submodular. Hence, Theorem 4.1 of [29] implies that the greedy algorithm returns a set S satisfying a $(1 - 1/e)$-optimality bound with the optimal set S^*, thus completing the proof. □

3.5. Controllability of Impulsive Dynamics. Descriptor system dynamics often have impulsive dynamics, leading to system trajectories that may not be physically realizable. The following lemma provides sufficient conditions for removing impulsive dynamics that can be incorporated into our framework.

LEMMA 3.15 ([16]). *There exists a feedback matrix K such that the pair $(F, A + BK)$ is impulse-free for any admissible initial state if and only if*

$$rank((\lambda F - A)|B) = n \quad \forall \lambda \in \mathbb{C} \setminus \{0\} \tag{3.11}$$

$$rank\left(\begin{array}{ccc} F & A & B \\ 0 & F & 0 \end{array}\right) = rank(F) + rank((\lambda F - A)|B). \tag{3.12}$$

Condition (3.11) is already guaranteed if the matroid constraint of Lemma 3.8 is satisfied. Eq. (3.12) reduces to

$$\text{rank}\left(\begin{array}{ccc} F & A & B \\ 0 & F & 0 \end{array}\right) = \text{rank}(F) + n. \tag{3.13}$$

The following result provides a matroid constraint for (3.13).

LEMMA 3.16. *There exists a matroid $\mathcal{M}_3$ such that the condition (3.13) holds for almost all choices of the free parameters iff $R = V \setminus S$ is independent in $\mathcal{M}_3$.*

Proof. By [26, Lemma 2.5], we have

$$\operatorname{rank}\begin{pmatrix} F & A & B \\ 0 & F & 0 \end{pmatrix} =$$
$$\operatorname{rank}\left[\mathcal{M}\left(\left[\begin{array}{cc|cc|c} I & 0 & Q_F & Q_A & 0 \\ 0 & I & 0 & Q_F & 0 \end{array}\right]\right) \vee \mathcal{M}\left(\left[\begin{array}{cc|cc|c} I & 0 & T_F & T_A & T_B \\ 0 & I & 0 & T_F & 0 \end{array}\right]\right)\right] \triangleq Z.$$

Let $\Theta = \{1, \ldots, 5n\}$ and $X = \{1, \ldots, 4n\}$, and let ρ denote the rank function of the matroid induced by Z. We have that (3.13) holds iff $\operatorname{rank}(X \cup S) = \operatorname{rank}(F) + n$. This is equivalent to $\rho(X \cup S) - \rho(X) = \operatorname{rank}(F) - \rho(X) + n$. The function $\hat{\rho}(S) \triangleq \rho(X \cup S) - \rho(X)$ is the rank function of the contraction of the matroid induced by Z around X. Hence we have the condition $\hat{\rho}(S) \geq \hat{n}$. Furthermore, $\hat{n}$ is the maximum possible value of $\hat{\rho}(S)$, and so $\hat{\rho}(S) \geq \hat{n}$ is equivalent to $R \in (Z \setminus X)^*$, i.e., the dual of the contraction of Z by X. Letting $\mathcal{M}_3 = (Z/X)^*$ completes the proof. ◻

Lemma 3.16 implies that controllability of impulsive modes can also be mapped to matroid constraints, and hence that impulsive dynamics can be incorporated into the submodular input selection approach.

4. Input Selection under Strong Connectivity. The input selection algorithms of the previous section hold for any arbitrary structured linear descriptor system. In this section, we investigate the case where the graph induced by the matrices T_F and T_A is strongly connected, i.e., there exists a directed path from any node i to any node j. In this case, there is additional problem structure that reduces the complexity and improves the optimality bounds of our input selection algorithms. The following lemma gives system properties that hold with high probability for strongly connected networks.

LEMMA 4.1. *If the graph induced by T_F and T_A is strongly connected and rank$(A|B) = n$, then the condition rank$((zF - A)|B) = n$ holds for all nonzero $z \in \mathbb{C}$.*

Proof. If $\operatorname{rank}(A|B) = n$, then $\det(zF - A) \neq 0$ for all $z \in \mathbb{C}$, except for some nonzero complex numbers $z_1, \ldots, z_n$. Consider z_i such that $\det(z_i F - A) = 0$. We have that

$$\det(z_i F - A) = \sum_{\sigma \in S_n} \prod_{j=1}^{n} (z_i F - A)_{j\sigma(j)}.$$

We observe that each σ corresponding to a nonzero term of the summation induces a decomposition of the graph into cycles $j_1, \ldots, j_m$, in which $j_{l+1} = \sigma(j_l)$ and $j_1 = \sigma(j_m)$. If the determinant is zero, then there exist at least two such decompositions, corresponding to distinct permutations σ and σ', with products of weights that sum to zero.

Suppose that the l-th column of the matrix $(z_i F - A)$ is linearly dependent on the other columns. Suppose that the column is replaced by one of the input columns from B. Since the graph is strongly connected, a new cycle is induced by adding the input column. For almost all values of the free parameters of B, the cycle will not cancel out with the other cycles in the graph, and hence the determinant will be nonzero. ◻

If $\operatorname{rank}((zF - A)|B) = n$ for all nonzero z, then only the constraint $(V \setminus S) \in \mathcal{M}_1^*$ must hold. We now describe how this additional problem structure improves the runtime and optimality gaps of each of the input selection problems considered.

4.1. Minimum-Size Input Set Selection in Strongly Connected Networks. In the strongly connected case, the minimum-size input set selection problem reduces to

$$
\begin{array}{ll}
\text{maximize} & |R| \\
\text{s.t.} & R \in \mathcal{M}_1^*
\end{array}
\tag{4.1}
$$

The following algorithm can be used to compute the solution to (4.1). Initialize the set $R = \emptyset$, and let $V = \{1, \dots, n\}$ be the set of possible input nodes. The algorithm iterates over all nodes in V, starting with the node indexed 1. For each node i, test if $(R \cup \{i\}) \in \mathcal{M}_1^*$. If so, set $R = R \cup \{i\}$. The algorithm terminates after all n nodes have been tested.

Lemma 4.2. *When the network is strongly connected and the condition of Lemma 4.1 holds, the greedy algorithm returns the minimum-size input set to guarantee controllability within $O(n)$ computations of the matroid independence condition $R \in \mathcal{M}_1^*$.*

Proof. Let $\mathcal{B}$ be the set of bases of $\mathcal{M}_1^*$. We define a lexicographic ordering on the sets $R \in \mathcal{B}$ as follows. Let R_1, $R_2 \in \mathcal{B}$, and write $R_1 = \{a_1, \dots, a_m\}$ and $R_2 = \{b_1, \dots, b_m\}$, where $a_1 < a_2 < \cdots < a_m$ and $b_1 < b_2 < \cdots < b_m$. We let $R_1 \prec R_2$ if there exists i such that $a_j = b_j$ for $j < i$ and $a_i < b_i$. We have that $\prec$ induces a total ordering on $\mathcal{B}$, since for any R_1 and R_2 with $R_1 \neq R_2$, we have $R_1 \prec R_2$ or $R_2 \prec R_1$.

Let R^* denote the set in $\mathcal{B}$ that is minimal under the ordering $\prec$. We show that the algorithm described above outputs R^*. Let $R^* = \{a_1, \dots, a_m\}$, and let $R_i^* = R^* \cap \{1, \dots, i\}$. Finally, let R_i denote the set computed by our algorithm at iteration i. We prove by induction that $R_i^* = R_i$ for each i.

We have $R_0 = R_0^*$ trivially. Now, suppose $R_i = R_{i-1}^*$. We have two cases. First, suppose that $i \in R^*$. Since $R_{i-1} \cup \{i\} = R_{i-1}^* \cup \{i\} = R_i^*$ and $R_i^* \subseteq R^*$, we have that $(R_{i-1} \cup \{i\}) \in \mathcal{M}_1^*$. Hence the algorithm will add i to the set, and $R_i = R_i^*$.

Now, suppose that $i \notin R^*$, and suppose that $R_i^* \neq R_i$. By inductive hypothesis, we must have that $i \in R_i$, which occurs if and only if $R_i = (R_{i-1} \cup \{i\})$ is independent in $\mathcal{M}_1^*$. Since $\mathcal{M}_1^*$ is a matroid, we can complete R_i to a basis $\hat{R} \in \mathcal{B}$. By definition of the $\prec$ ordering, $\hat{R} \prec R^*$, contradicting the assumption that R^* is minimal under the ordering $\prec$. This contradiction implies that $i \notin R_i$, and so $R_i = R_i^*$.

Continuing inductively until $i = n$, we have that $R_n = R_n^* = R^*$. Since R_n is equal to $R^* \in \mathcal{B}$, R_n is a basis of $\mathcal{M}_1^*$, and hence is a solution to (4.1). The $O(n)$ runtime follows from the fact that the algorithm makes one independence check per iteration over a total of n iterations. ∎

The additional problem structure in the strongly connected case leads to a simplified algorithm with reduced runtime compared to Algorithm 1.

4.2. Joint Performance and Controllability Input Selection in Strongly Connected Networks. When the system graph is strongly connected and the conditions of Lemma 4.1 hold, the complexity and optimality bounds of input selection for joint performance and controllability are improved. With this additional structure, the problem formulation is given by

$$
\begin{array}{ll}
\text{maximize} & f(S) \\
\text{s.t.} & |S| \leq k \\
 & (V \setminus S) \in \mathcal{M}_1^*
\end{array}
\tag{4.2}
$$

The following lemma gives an equivalent formulation to (4.2).

Lemma 4.3. *Let $r_1 = rank(\mathcal{M}_1)$ and define $\hat{\mathcal{M}}_1 = \mathcal{M}_1 \vee U_{k-r_1}$. If the objective function $f(S)$ is monotone, then the optimization problem (4.2) has the same solution* *as*

$$\begin{array}{cc} \textit{maximize} & f(S) \\ \textit{s.t.} & S \in \hat{\mathcal{M}}_1 \end{array} \tag{4.3}$$

Proof. Let S^* and $\hat{S}$ denote the optimal solutions to (4.2) and (4.3), respectively. We observe that both $|S^*| = |\hat{S}| = k$ by monotonicity of $f(S)$. We show that if $|S| = k$, then the conditions $(V \setminus S) \in \mathcal{M}_1^*$ and $S \in \hat{\mathcal{M}}_1$ are equivalent. First, suppose that $(V \setminus S) \in \mathcal{M}_1^*$. Let $R^* = (V \setminus S) \in \mathcal{M}_1^*$. The set R^* can be completed to a basis of $\mathcal{M}_1^*$, denoted $\hat{R} = R^* \cup R'$, and so we have $S = (V \setminus \hat{R}) \cup (V \setminus R')$. Now, $(V \setminus \hat{R}) \in \mathcal{M}_1$ and $|V \setminus R'| = k - r_1$, and so $S \in \hat{\mathcal{M}}_1$.

Suppose that $S \in \hat{\mathcal{M}}_1$ and $|S| = k$. Since $|S| = k$, S is a basis of $\hat{\mathcal{M}}_1$, and so S can be written as $S = S_1 \cup S_2$ where $S_1 \in \mathcal{M}_1$. Hence $(V \setminus S) \subseteq (V \setminus S_1) \in \mathcal{M}_1^*$, and so the second constraint of (4.2) is satisfied. ◻

Convex relaxation approaches have been proposed for solving matroid-constrained monotone submodular maximization problems [2, 4]. One such approach is to replace Line 12 in Algorithm 2 with a subroutine that computes the maximum-weighted basis of a matroid (such a basis can be computed efficiently using a greedy algorithm). It was shown in [2] that this algorithm achieves a $(1 - 1/e)$ optimality bound.

4.3. Performance-Controllability Trade-Off in Strongly Connected Networks. In the strongly connected network case, we define the graph controllability index (GCI)

$$c(S) = \max\{|V'| : V' \text{ controllable from } S\}.$$

The following lemma provides additional structure on $c(S)$.

Lemma 4.4. *The function $c(S) = \tilde{c}(S) + \zeta$, where ζ is a constant and $\tilde{c}(S)$ is a matroid rank function.*

Proof. The largest controllable subgraph of G corresponds to a subset of states such that, for the matrix A' with columns indexed in V', we have $\text{rank}(\mathcal{M}([I|Q_{A'}|0] \vee [I|T_{A'}|T_B(S)]) = 2|V'|$. This subset of columns, however, is exactly the maximum-size independent set in $(\mathcal{M}([I|Q_A|0]) \vee \mathcal{M}([I|T_A|T_B(S)]))$, and the value of $c(S)$ is $\text{rank}(\mathcal{M}([I|Q_A|0]) \vee \mathcal{M}([I|T_A|T_B(S)]))$. By Lemma 3.3, the rank is equal to the rank of $\mathcal{M}([I|Q_A]) \vee \mathcal{M}([I|T_A])$ plus a matroid rank function of S. ◻

The problem of selecting a set of up to k input nodes to maximize both a performance metric $f(S)$ and the GCI $c(S)$ is formulated as

$$\begin{array}{ll} \text{maximize} & f(S) + \eta c(S) \\ \text{s.t.} & |S| \leq k \end{array} \tag{4.4}$$

As in the general case, a greedy algorithm for maximizing $f(S) + \eta c(S)$ is guaranteed to return an input set S^* such that $f(S^*) + \eta c(S^*)$ is within a $(1 - 1/e)$ factor of the optimum. Moreover, when the performance metric $f(S)$ is identically zero, so that only controllability is optimized, we have the following result.

Lemma 4.5. *If $f(S) = 0$, then the greedy algorithm returns the optimal solution to (4.4).*

Proof. For the problem of maximizing a matroid rank function subject to a cardinality constraint, the greedy algorithm is known to return an optimal solution [31]. If $f(S) = 0$, then by Lemma 4.4, Eq. (4.4) is equivalent to maximizing a matroid rank function subject to a cardinality constraint, and hence the greedy algorithm returns the optimal input set S. ◻

5. Special Cases of Our Approach. In this section, we consider three special cases of our framework, namely systems with linear consensus dynamics, networked systems where each node has second integrator dynamics, and systems where all parameters are free.

5.1. Linear Consensus Dynamics. We first consider a network of N nodes where each node $i \in \{1,\ldots,N\}$ has a state $x_i(t) \in \mathbb{R}$. The state dynamics of the non-input nodes are given by $\dot{x}_i(t) = -\sum_{j \in N(i)} W_{ij}(x_i(t) - x_j(t))$, where W_{ij} are nonnegative weights. In [14], it was shown that, by introducing a set of states $\{x_j^e : j = 1,\ldots,M\}$, where M is the number of edges in the network, the system can be written in the form (2.3) as

$$\left(\begin{array}{cc} I & 0 \\ 0 & 0 \end{array}\right)\left(\begin{array}{c} \dot{\mathbf{x}}(t) \\ \dot{\mathbf{x}}^e(t) \end{array}\right) = \left(\begin{array}{cc} 0 & K \\ K_I & W \end{array}\right)\left(\begin{array}{c} \mathbf{x}(t) \\ \mathbf{x}^e(t) \end{array}\right). \tag{5.1}$$

In (5.1), K_I is the incidence matrix of the graph and K is the transpose of the incidence matrix. W is a diagonal matrix with e-th entry equal to the weight on edge e. We assume that the weights W are free parameters, so that Q_F, T_F, Q_A, and T_A are given by

$$Q_F = \left(\begin{array}{cc} I & 0 \\ 0 & 0 \end{array}\right), \quad Q_A = \left(\begin{array}{cc} 0 & K \\ K_I & 0 \end{array}\right), \quad T_A = \left(\begin{array}{cc} 0 & 0 \\ 0 & W \end{array}\right). \tag{5.2}$$

As a first step towards analyzing this class of system dynamics under our framework, we consider the condition of Lemma 2.10. For this system, the matrix Ω of Section 2.3 is equal to

$$\Omega = \begin{array}{c} \mathbf{w} \\ \mathbf{w}^e \\ \mathbf{x} \\ \mathbf{x}^e \\ \mathbf{u} \end{array} \left(\begin{array}{ccc} -I & K & 0 \\ K_I & 0 & 0 \\ I & 0 & 0 \\ 0 & I & 0 \\ 0 & 0 & I \end{array}\right) \tag{5.3}$$

We use $\mathbf{w}^e$ and $\mathbf{x}^e$ to denote the auxiliary graph nodes corresponding to the states $\mathbf{x}^e(t)$. We observe that in the augmented graph, since the weight matrix is diagonal, there is a directed edge from $w_j^{e,T}$ to $x_j^{e,T}$ for all edges indexed $j = 1,\ldots,M$. We have the following intermediate result.

LEMMA 5.1. *The matching m with $m(w_j^{e,T}) = x_j^{e,T}$ for all $j = 1,\ldots,M$ and $m(w_i^T) = w_i^Q$ for all $i = 1,\ldots,N$ satisfies the conditions of Lemma 2.9.*

Proof. The matching m is valid under the construction of Section 2.3. The rows of Ω from (5.3) indexed in $J = \{x_j^{e,Q} : j = 1,\ldots,M\} \cup \{w_i^Q : i = 1,\ldots,N\}$ form the matrix

$$\Omega_J = \left(\begin{array}{cc} -I & K \\ 0 & I \end{array}\right),$$

which has full rank. □

We can then compute $\Omega\Omega_{J \cup J_1}^{-1}$ as

$$\Omega\Omega_{J \cup J_1}^{-1} = \left(\begin{array}{ccc} I & 0 & 0 \\ -K_I & K_I K & 0 \\ -I & K & 0 \\ 0 & I & 0 \\ 0 & 0 & I \end{array}\right) \tag{5.4}$$

The graph $G' = (V', E')$ of Lemma 3.5 defined by Eq. (5.4) is described by the following lemma.

LEMMA 5.2. *For the system (5.1), the edge set E' of Lemma 3.5 is defined by*

$$\begin{aligned}E' = &\{(w_e^Q, w_i^Q) : e = (i,j) \text{ for some } j \in V\} \cup \{(x_i^Q, x_e^Q) : e = (i,j) \text{for some } j \in V\}\\ &\cup\{(w_e^Q, x_{e'}^Q) : \text{edges } e \text{ and } e' \text{ have a node in common}\}\\ &\cup\{(x_j^{e,Q}, x_j^{e,T}) : j = 1,\ldots,M\} \cup \{(w_j^{e,T}, w_j^{e,Q}) : j = 1,\ldots,M\}\\ &\cup\{(w_i^Q, w_i^T) : i = 1,\ldots,N\}.\end{aligned}$$

Proof. Considering the edge set of the auxiliary graph defined in Section 2.3, the set $\{(w_i^T, x_j^Q) : (i,j) \in N(T_A) \cup N(T_F), m(w_i^T) \neq x_j^Q\}$ is empty, while the set $\{(x_j^Q, w_i^T) : (i,j) \in N(T_A) \cup N(T_F), m(w_i^T) = x_j^Q\}$ is equal to $\{(x_j^{e,Q}, x_j^{e,T}) : j = 1,\ldots,M\} \cup \{(w_j^{e,T}, w_j^{e,Q}) : j = 1,\ldots,M\}$. The set $\{(w_i^T, w_i^Q) : m(w_i^T) \neq w_i^Q\} \cup \{(w_i^Q, w_i^T) : m(w_i^T) = m(w_i^Q)\}$ is equal to $\{(w_j^{e,T}, w_j^{e,Q}) : j = 1,\ldots,M\} \cup \{(w_i^Q, w_i^T) : i = 1,\ldots,N\}$.

It remains to compute the set $\{(x^Q, y^Q) : x \in \hat{V} \setminus J, y \in J, \tilde{\Omega}_{xy} \neq 0, \tilde{\Omega}_{xz} = 0 \ \forall z \in J_1\}$. This set is defined by the off-diagonal entries of $\Omega\Omega_{J\cup J_1}^{-1}$ from (5.4). The entries from $w_j^{e,Q}$ to w_i^Q correspond to the entries of the incidence matrix, and hence there is a nonzero entry if and only if edge j is incident to node i. A similar argument holds for the $(x_i^Q, x_j^{e,Q})$ edges. Finally, an edge $(w_j^{e,Q}, w_{j'}^{e,Q})$ is formed if $(K_I K)_{jj'} \neq 0$. This matrix is the *edge Laplacian*, which has a nonzero entry if and only if either $j = j'$, or edges j and j' are incident to the same node. □

This description of the graph G' enables characterization of the input-connected nodes in $\hat{V}$.

LEMMA 5.3. *The nodes x_i^T, w_i^T and w_i^Q do not belong to any cycle in G'. A node x_i^Q is input-connected in G' if and only if i is input-connected in the graph G induced by the consensus dynamics. A node $x_j^{e,T}$, $w_j^{e,T}$, $x_j^{e,Q}$, or $w_j^{e,Q}$ is input-connected in G' if and only if edge j is incident on a node that is input-connected in the consensus network G.*

Proof. By Lemma 5.2, w_i^T and w_i^Q are only connected to each other, and hence are not part of any cycle since the link is directional. Similarly, the nodes x_i^T cannot belong to any cycle, since they have no incoming edges.

Now, suppose that there is a path from node i to an input i' in G. Let $(i, i_1), \ldots, (i_r, i')$ denote one such path, and let $j_0, \ldots, j_r$ denote the indices of the edges on the path. Then there is a path π from x_i^Q to $w_{i'}^T$, given by

$$\pi = (x_i^Q, x_{j_0}^{e,Q}) \cup \bigcup_{l=0}^{r-1} \left\{(x_{j_l}^{e,Q}, x_{j_l}^{e,T}), (x_{j_l}^{e,T}, w_{j_l}^{e,T}), (w_{j_l}^{e,T}, w_{j_l}^{e,Q}), (w_{j_l}^{e,Q}, x_{j_{l+1}}^{e,Q})\right\} \cup \{(x_{j_r}^{e,Q}, w_{i'}^Q), (w_{i'}^Q, w_{i'}^T)\} \tag{5.5}$$

For the other direction, we have that any path from x_i^Q to $w_{i'}^T$ has the form of (5.5), and hence defines a path from i to i' in G. Finally, suppose that edge j is incident on node i and that there is a path from a node i to an input i' in G. Let $(i, i_1), \ldots, (i_r, i')$ denote one such path, and let $j_0, \ldots, j_r$ denote the indices of the edges on the path. For node $w_j^{e,Q}$, there is a path given by

$$\pi' = (w_j^{e,Q}, x_{j_0}^{e,Q}) \cup \bigcup_{l=0}^{r-1} \{(x_{j_l}^{e,Q}, x_{j_l}^{e,T}), (x_{j_l}^{e,T}, w_{j_l}^{e,T}), (w_{j_l}^{e,T}, w_{j_l}^{e,Q}), (w_{j_l}^{e,Q}, x_{j_{l+1}}^{e,Q})\} \cup \{(x_{j_r}^{e,Q}, w_{i'}^{Q}), (w_{i'}^{Q}, w_{i'}^{T})\}$$

A path can also be found for nodes $x_j^{e,Q}$, $x_j^{e,T}$, and $w_j^{e,T}$ by using the path $(x_j^{e,Q}, x_j^{e,T})$, $(x_j^{e,T}, w_j^{e,T})$, $(w_j^{e,T}, w_j^{e,Q})$. ∎

Based on Lemma 5.3, we can characterize exactly when the condition of Lemma 2.10 holds, based on the connectivity of the network graph G.

LEMMA 5.4. *The condition of Lemma 2.10 holds for the system (5.1) if and only if each node is input-connected in the graph G.*

The proof follows directly from Lemma 5.3. Lemma 5.4 enables us to improve the optimality bounds for a class of metrics with a certain structure. Suppose that the connected components of the graph are equal to $G_1, \ldots, G_r$, with $G_i = (V_i, E_i)$ for $i = 0, \ldots, r$. We consider metrics of the form

$$f(S) = \sum_{i=1}^{r} f_i(S \cap V_i). \tag{5.6}$$

Eq. (5.6) has the interpretation that the performance of nodes in connected component V_i only depends on the set of input nodes for component V_i, instead of the overall input set. This structure holds for, e.g., the metrics of [34, 8, 5]. For these metrics, we have the following optimality result.

THEOREM 5.5. *For the consensus system (5.1), the problem of maximizing a performance metric of the form (5.6) subject to controllability as a constraint and $|S| \leq k$, formulated as*

$$\begin{array}{ll} \textit{maximize} & f(S) \\ \textit{s.t.} & (V \setminus S) \in \mathcal{M}_1^* \cap \mathcal{M}_2^* \\ & |S| \leq k \end{array} \tag{5.7}$$

can be approximated up to an optimality bound of $(1 - 1/e)$ in polynomial time. As a special case, if the graph G is strongly connected, then any monotone submodular performance metric can be approximated up to an optimality bound of $(1 - 1/e)$ in polynomial time.

Proof. The proof is by showing that the optimality bounds of Algorithm 2 are improved in this case. By Theorem 3.11, the continuous relaxation phase of Algorithm 2 returns a vector $y(1)$ such that $F(y(1)) \geq (1 - 1/e)f(S^*)$, where S^* is the optimal solution to (5.7). Now, Theorem II.3 of [4] implies that the SWAP_ROUND subroutine satisfies $\mathbf{E}(f(S \cap Q)) \geq F(x_i : i \in Q)$ for any set Q of equivalent elements of one of the matroids $\hat{\mathcal{M}}_1$ or $\hat{\mathcal{M}}_2$. For the matroid $\hat{\mathcal{M}}_2$, each set of elements V_s is equivalent, and so $\mathbf{E}(f(S \cap V_s)) \geq F(x_i : i \in V_s)$. Summing over s yields the desired result.

For the special case, we have that when the graph is strongly connected, input connectivity holds provided there is at least one input. Hence we can obtain a $(1-1/e)$-bound on the optimal input set. ∎

5.2. Double Integrator Dynamics. We study networked systems where the second derivative of each node's state $\xi_i(t)$ is a function of its neighbor states, so that

$\ddot{\xi}_i(t) = \sum_{j \in N(i)} W_{ij}\xi_j(t) + \Gamma_{ij}\dot{\xi}_j(t)$, where $N(i)$ is the neighbor set of node i, and Γ_{ij} and W_{ij} are free parameter values. The double integrator model is applicable for larger vehicles that have inertial components in their dynamics [38]. We write this system in the form (2.3) by introducing variables $\zeta_i(t) = \dot{\xi}_i(t)$, resulting in dynamics

$$\begin{pmatrix} \dot{\xi}(t) \\ \dot{\zeta}(t) \end{pmatrix} = \begin{pmatrix} 0 & I \\ W & \Gamma \end{pmatrix} \begin{pmatrix} \xi(t) \\ \zeta(t) \end{pmatrix}, \tag{5.8}$$

where $F = I$. In analyzing this system, we observe that it is not possible to independently control the states $\xi_i(t)$ and $\zeta_i(t)$ to any arbitrary trajectories, since $\zeta_i(t) = \dot{\xi}_i(t)$. Hence we assume that the state $\zeta_i(t)$ (the velocity) is controlled in the input nodes, while the state $\xi_i(t)$ continues to follow the dynamics (5.8).

We first investigate the auxiliary graph condition of Lemma 2.10. The matrix Ω is given by

$$\Omega = \begin{array}{c} \mathbf{w}^{\xi} \\ \mathbf{w}^{\zeta} \\ \mathbf{x}^{\xi} \\ \mathbf{x}^{\zeta} \\ \mathbf{u} \end{array} \begin{pmatrix} -I & I & 0 \\ 0 & -I & 0 \\ I & 0 & 0 \\ 0 & I & 0 \\ 0 & 0 & I \end{pmatrix} \tag{5.9}$$

We have that the rows indexed in $\mathbf{w}$ have full rank, and hence the matching $m(w_i^{\xi,T}) = w_i^{\xi,Q}$ and $m(w_i^{\zeta,T}) = w_i^{\zeta,Q}$ for $i = 1, \ldots, N$ satisfies the conditions of Lemma 2.9. This gives $J = \{w_i^{\zeta,Q} : i = 1, \ldots, N\} \cup \{w_i^{\xi,Q} : i = 1, \ldots, N\}$. We then have $\Omega_{J \cup J_1}$, $\Omega_{J \cup J_1}^{-1}$, and $\Omega\Omega_{J \cup J_1}^{-1}$ as

$$\Omega_{J \cup J_1} = \begin{pmatrix} -I & I & 0 \\ 0 & -I & 0 \\ 0 & 0 & I \end{pmatrix}, \quad \Omega_{J \cup J_1}^{-1} = \begin{pmatrix} -I & -I & 0 \\ 0 & -I & 0 \\ 0 & 0 & I \end{pmatrix}, \tag{5.10}$$

$$\Omega\Omega_{J \cup J_1}^{-1} = \begin{array}{c} \mathbf{w}_{\xi} \\ \mathbf{w}_{\zeta} \\ \mathbf{x}^{\xi} \\ \mathbf{x}^{\zeta} \\ \mathbf{u} \end{array} \begin{pmatrix} I & 0 & 0 \\ 0 & I & 0 \\ -I & -I & 0 \\ 0 & -I & 0 \\ 0 & 0 & I \end{pmatrix}.$$

These matrix values lead to the following description of the auxiliary graph.

Lemma 5.6. *The graph G' of Lemma 3.5 has edge set E' given by*

$$\begin{aligned} E' = &\{(w_i^{\xi,Q}, w_i^{\xi,T}) : i = 1, \ldots, N\} \cup \{(w_i^{\zeta,Q}, w_i^{\zeta,T}) : i = 1, \ldots, N\} && (5.11) \\ &\cup \{(x_i^{\xi,Q}, x_i^{\xi,T}) : i = 1, \ldots, N\} \cup \{(x_i^{\zeta,Q}, x_i^{\zeta,T}) : i = 1, \ldots, N\} \\ &\cup \{(w_i^{\xi,T}, x_j^{\xi,T}) : j \in N(i)\} \cup \{(w_i^{\xi,T}, x_j^{\zeta,T}) : j \in N(i)\} && (5.12) \\ &\cup \{(w_i^{\zeta,T}, w_j^{\xi,T}) : j \in N(i)\} \cup \{(w_i^{\zeta,T}, w_j^{\xi,T}) : j \in N(i)\} \\ &\cup \{(x_i^{\xi,Q}, w_i^{\xi,Q}) : i = 1, \ldots, N\} \cup \{(x_i^{\xi,Q}, w_i^{\zeta,Q}) : i = 1, \ldots, N\} && (5.13) \\ &\cup \{(x_i^{\zeta,Q}, w_i^{\zeta,Q}) : i = 1, \ldots, N\} \end{aligned}$$

Proof. The edges (5.11) correspond to the edges $\{(w_i^Q, w_i^T) : m(w_i^T) = w_i^Q\}$ in the definition of $\hat{E}$. The edges enumerated in (5.12) correspond to the edges

$\{(w_i^T, x_j^T) : (i,j) \in N(T_A)$. Finally, the value of $\Omega\Omega_{J\cup J_1}^{-1}$ from (5.4) implies that the edges enumerated in (5.13) correspond to the edges $\{(x^Q, y^Q) : x \in \hat{V} \setminus J, y \in J, \tilde{\Omega}_{xy} \neq 0, \tilde{\Omega}_{xz} = 0 \forall z \in J_1\}$. ☐

Lemma 5.6 leads to the following result, which relates the connectivity of the auxiliary graph and the graph G induced by the node dynamics. This is analogous to Lemma 5.3.

LEMMA 5.7. *For any node i, the nodes $w_i^{\xi,Q}$, $w_i^{\zeta,Q}$, $x_i^{\xi,Q}$, $x_i^{\zeta,Q}$, $w_i^{\xi,T}$, $w_i^{\zeta,T}$, $x_i^{\xi,T}$, and $x_i^{\zeta,T}$ are input-connected in the auxiliary graph G' if and only if node i is input-connected in the graph G.*

Proof. Suppose that node i is connected to an input node i' in G, with path $(i,i_1),(i_1,i_2),\ldots,(i_r,i')$. Now, consider the node $w_i^{\zeta,T}$. We can construct a path π from $w_i^{\zeta,T}$ to $w_{i'}^{\xi,T}$ as

$$\begin{aligned}\pi = (w_i^{\zeta,T}, x_{i_0}^{\xi,T}) \cup \bigcup_{l=0}^{r} \{&(x_{i_l}^{\xi,T}, x_{i_l}^{\xi,Q}), (x_{i_l}^{\xi,Q}, w_{i_l}^{\xi,Q}), (w_{i_l}^{\xi,Q}, w_{i_l}^{\xi,T}), (w_{i_l}^{\xi,T}, x_{i_{l+1}}^{\xi,T})\} \\ &\cup \{(x_{i'}^{\xi,T}, x_{i'}^{\xi,Q}), (x_{i'}^{\xi,Q}, w_{i'}^{\xi,Q}), (w_{i'}^{\xi,Q}, w_{i'}^{\xi,T})\}.\end{aligned} \tag{5.14}$$

Paths for the other types of nodes in the auxiliary graph can be constructed in a similar fashion. Conversely, any path to an input node in the auxiliary graph will have the form (5.14), and hence can be used to construct a path to an input node in the graph G. ☐

Lemma 5.7 implies that, for performance metrics satisfying (5.6), Algorithm 2 returns a set S satisfying $f(S) \geq (1-1/e)f(S^*)$, where S^* is the optimal solution. The proof is analogous to Theorem 5.5.

5.3. Input Selection in Networks of Free Parameters. We now investigate systems where all of the matrix entries are free parameters, as in the models of [7, 24, 40]. We consider systems of the form $\dot{\mathbf{x}}(t) = A\mathbf{x}(t)$, where A is a free matrix and $F = I$. In such systems, we have the following known result.

LEMMA 5.8 ([21]). *If A is a free matrix and $F = I$, then the condition rank$(zF - A|B) = n$ holds if and only if each node is connected to at least one input node in the graph induced by A.*

Lemma 5.8 implies that, if a system with free parameters is strongly connected, then it suffices to find an input set such that the matroid of Lemma 3.3 is full rank. A minimum-size input set such that $\text{rank}(\mathcal{M}([I|T_A|T_B(S)])) = n$ can be found efficiently using the greedy algorithm. This enables efficient computation of an input set with the same size as in [24], but through the matroid optimization framework.

Finally, we observe that, by Lemma 5.8, Algorithm 2 returns a set S that satisfies a $(1-1/e)$ optimality bound, by the argument of Theorem 5.5.

6. Numerical Study. We numerically evaluated our framework using Matlab^{TM}. We simulated two systems. First, we considered a descriptor system analyzed in [26]. Next, we studied a descriptor system defined by an electric circuit. We then simulated the consensus network case of Section 5.1.

In the example of [26], the system dynamics are defined by $F\dot{\mathbf{x}}(t) = A\mathbf{x}(t)$, where

$$F = \begin{pmatrix} 0 & 0 & * \\ 1 & 1 & * \\ 0 & 0 & 0 \end{pmatrix}, \quad A = \begin{pmatrix} 1 & * & 0 \\ 0 & 0 & 1 \\ -1 & -1 & * \end{pmatrix}$$

where $*$ indicates a free parameter. The constraint of Lemma 3.3 is defined by

$$\mathrm{rank}\left[\mathcal{M}\left(\left[\begin{array}{ccc|ccc|ccc} 1 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & -1 & -1 & 0 & 0 & 0 & 0 \end{array}\right]\right) \vee \mathcal{M}\left(\left[\begin{array}{ccc|ccc|c} 1 & 0 & 0 & 0 & * & 0 & \\ 0 & 1 & 0 & 0 & 0 & 0 & \cdot \\ 0 & 0 & 1 & 0 & 0 & * & \end{array}\right]\right)\right] = 6.$$

Since the matrix

$$\begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ -1 & -1 & 0 \end{pmatrix}$$

is full-rank, the constraint is satisfied automatically. Turning to the constraint $(V \setminus S) \in \mathcal{M}_2^*$, we have that $\overline{V}$ contains two equivalence classes, namely $\{w_3^T\}$ and $\{w_1^T\}$. Hence, the only independent sets in $\mathcal{M}_2^*$ are $\emptyset$ and $\{2\}$, and the input set must contain states 1 and 3 for controllability to hold. This agrees with the results of [26].

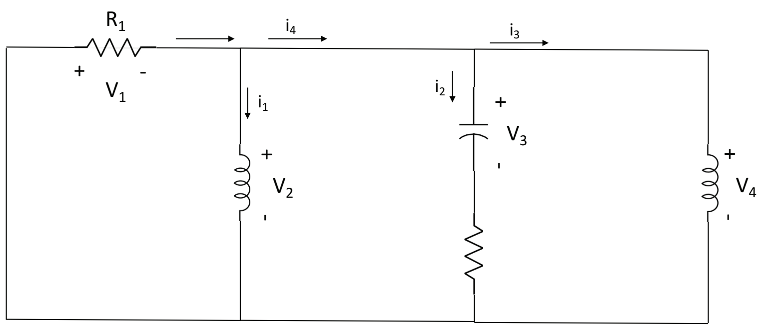

FIG. 1. *Circuit diagram. The goal is to select a location to place a voltage or current source in order to ensure that all voltages and currents in the circuit are controllable.*

Next, consider the circuit network of Figure 1. The system state $\mathbf{x}(t)$ is defined by the voltage and current at each element of the circuit, so that $\mathbf{x}(t) = [v_1(t)\ v_2(t)\ v_3(t)\ v_4(t)\ i_1(t)\ i_2(t)\ i_3(t)\ i_4(t)]^T$. The goal is to select a state to act as control input, interpreted as placing a voltage or current source at a particular location in order to satisfy controllability. We have that

$$F = \begin{pmatrix} 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{pmatrix}, \quad A = \begin{pmatrix} 0 & * & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & * & 0 & 0 \\ 0 & 0 & 0 & * & 0 & 0 & 0 & 0 \\ * & 0 & 0 & 0 & 1 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 & 1 & 1 & -1 \\ -1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 1 & 0 & * & 0 & 0 \\ 0 & 1 & 1 & 0 & 0 & * & 0 & 0 \end{pmatrix}.$$

By examination, the constraint $\mathrm{rank}(zF - A|B) = n$ holds for all nonzero z and any choice of the input set. Structural controllability can therefore be guaranteed using the techniques of Section 4. It suffices to control the state v_1 (corresponding to introducing a voltage source in series with the resistor R_1) to satisfy controllability.

The consensus network simulation is defined as follows. Network topologies were generated by placing nodes at uniform random positions within a square region, and creating a link (i,j) if node i is within the communication range of node j. The range of each node was chosen uniformly at random from the interval $[0,600]$. We investigated the minimum-size set of input nodes for structural controllability, as well as selection of input nodes for joint performance and controllability.

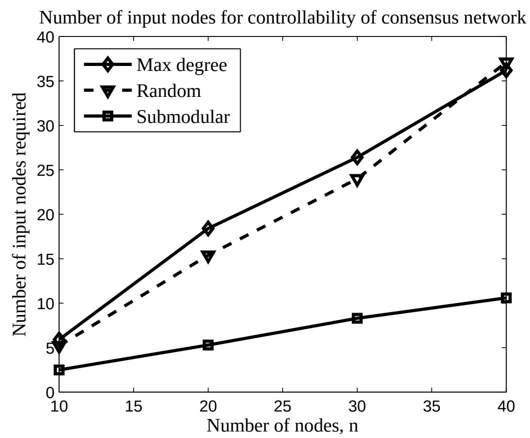

Fig. 2. *Minimum-size input set for structural controllability in a consensus network. The submodular optimization approach is compared to max degree-based and random input selection. The submodular optimization approach typically requires roughly one-quarter of the network to be controlled, while the random and degree-based heuristics select nearly all network nodes before controllability is satisfied.*

Since each node has a single state, the total number of states n is equal to the total number of nodes. We considered networks of size $n = \{10, 20, 30, 40\}$. The deployment area was selected to yield average node degrees $d = 3$. We compared our submodular optimization approach with selecting high degree nodes as inputs, as well as selecting random nodes to act as inputs. The submodular optimization approach required fewer input nodes to satisfy controllability, with the other heuristics selecting nearly all network nodes before controllability is satisfied. For all schemes, the number of input nodes was increasing in the network size.

We evaluated selection of input nodes in order to minimize the convergence error with controllability as a constraint. The convergence error was defined as $||\mathbf{x}(t) - x^{*}\mathbf{1}||_2$, where x^* is the state of the input nodes, $t = 1$, and the initial state and edge weights were chosen uniformly at random. The number of nodes was equal to 20, while the deployment area was chosen to achieve an average degree of 2. The submodular optimization approach provided lower convergence error than the degree-based and random heuristics, while also satisfying controllability from the input set. As the number of input nodes increased, the gap between the submodular optimization approach and the other heuristics increased.

7. Conclusions. In this paper, we studied the problem of input selection for joint performance and controllability. Our main contributions were two-fold. First, we proposed polynomial-time algorithms for input selection in structured linear descriptor systems. Second, we developed a submodular optimization framework for

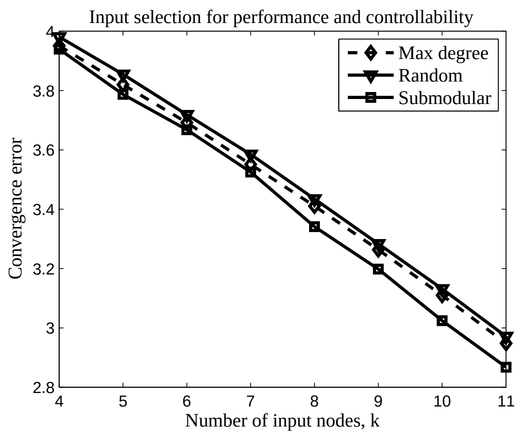

Fig. 3. *Convergence error when input nodes are selected to minimize convergence error while satisfying controllability. The number of nodes n was equal to* 20. *The submodular optimization approach provided lower convergence error than degree-based and random selection algorithms, especially as the number of input nodes increased.*

joint input selection based on both performance and controllability. Our approach was to prove that structural controllability of linear descriptor systems can be mapped to two matroid constraints, representing controllability of the zero and nonzero modes of the system. We demonstrated that, by exploiting this matroid structure, a minimum-size set of input nodes to guarantee structural controllability of such systems can be selected in polynomial time via matroid intersection algorithms. We further showed that selection of input nodes for joint performance and controllability can be formulated as a submodular maximization problem subject to two matroid basis constraints. We presented polynomial-time algorithms for obtaining a continuous solution to the input selection problem, providing a $(1-1/e)$ optimality bound, which can then be rounded to obtain a feasible input set. We demonstrated that, when the objective function is modular, the optimal input selection for performance and controllability can be computed in polynomial time.

We investigated input selection in systems where the graph representation of the system is strongly connected, and found that for almost all systems of this type the number of matroid constraints can be reduced from two to one. This led to an $O(n)$ algorithm for selecting a minimum-size set of input nodes for structural controllability, as well as more efficient polynomial-time algorithms for approximating the optimal input set up to a factor of $(1-1/e)$. We studied linear consensus systems, double integrator systems, and systems consisting of free parameters within our framework, and showed that the additional structure of each system provided a provable $(1-1/e)$ optimality bound for input selection based on performance and controllability.

REFERENCES

[1] B. Bamieh, M.R. Jovanovic, P. Mitra, and S. Patterson, *Coherence in large-scale networks: Dimension-dependent limitations of local feedback*, IEEE Trans. Automat. Control, 57 (2012), pp. 2235 –2249.

[2] G. Calinescu, C. Chekuri, M. Pal, and Jan Vondrak, *Maximizing a submodular set function subject to a matroid constraint*, SIAM J. Comput., 40 (2011), pp. 1740–1766.

[3] A. Chapman and M. Mesbahi, *On strong structural controllability of networked systems: a*

constrained matching approach, American Control Conference (ACC), (2013), pp. 6126–6131.

[4] C. Chekuri, J. Vondrak, and R. Zenklusen, *Dependent randomized rounding via exchange properties of combinatorial structures*, 51st Annual IEEE Symposium on Foundations of Computer Science (FOCS), (2010), pp. 575–584.

[5] A. Clark, B. Alomair, L. Bushnell, and R. Poovendran, *Minimizing convergence error in multi-agent systems via leader selection: A supermodular optimization approach*, IEEE Trans. Automat. Control, 59 (2014), pp. 1480–1494.

[6] A. Clark, L. Bushnell, and R. Poovendran, *Leader selection for minimizing convergence error in leader-follower systems: A supermodular optimization approach*, 10th International Symposium on Modeling and Optimization in Mobile, Ad Hoc and Wireless Networks (WiOpt), (2012), pp. 111–115.

[7] ———, *On leader selection for performance and controllability in multi-agent systems*, 51st IEEE Conference on Decision and Control (CDC), (2012), pp. 86–93.

[8] ———, *A supermodular optimization framework for leader selection under link noise in linear multi-agent systems*, IEEE Trans. Automat. Control, 59 (2014), pp. 283–297.

[9] J Corfmat and A.S. Morse, *Structurally controllable and structurally canonical systems*, IEEE Trans. Automat. Control, 21 (1976), pp. 129–131.

[10] J.-M. Dion, C. Commault, and J. Van Der Woude, *Generic properties and control of linear structured systems: a survey*, Automatica J. IFAC, 39 (2003), pp. 1125–1144.

[11] M. Fardad, F. Lin, and M. Jovanovic, *Algorithms for leader selection in large dynamical networks: Noise-free leaders*, 50th IEEE Conference on Decision and Control and European Control Conference (CDC-ECC), (2011), pp. 7188–7193.

[12] K. Fitch and N.E. Leonard, *Information centrality and optimal leader selection in noisy networks*, 52nd IEEE Conference on Decision and Control (CDC), (2013), pp. 7510–7515.

[13] J. Ghaderi and R. Srikant, *Opinion dynamics in social networks: A local interaction game with stubborn agents*, arXiv preprint arXiv:1208.5076, (2012).

[14] D. Goldin and J. Raisch, *On the weight controllability of consensus algorithms*, IEEE European Control Conference (ECC), (2013), pp. 233–238.

[15] Y. Hong, J. Hu, and L. Gao, *Tracking control for multi-agent consensus with an active leader and variable topology*, Automatica J. IFAC, 42 (2006), pp. 1177–1182.

[16] Ming Hou, *Controllability and elimination of impulsive modes in descriptor systems*, IEEE transactions on automatic control, 49 (2004), pp. 1723–1729.

[17] A. Jadbabaie, J. Lin, and A. S. Morse, *Coordination of groups of mobile autonomous agents using nearest neighbor rules*, IEEE Trans. on Automat. Control, 48 (2003), pp. 988–1001.

[18] M. Ji, A. Muhammad, and M. Egerstedt, *Leader-based multi-agent coordination: Controllability and optimal control*, American Control Conference (ACC), (2006), pp. 1358–1363.

[19] A. Krause, A. Singh, and C. Guestrin, *Near-optimal sensor placements in Gaussian processes: Theory, efficient algorithms and empirical studies*, J. Mach. Learn. Res., 9 (2008), pp. 235–284.

[20] Peter Kunkel and Volker Mehrmann, *The linear quadratic optimal control problem for linear descriptor systems with variable coefficients*, Mathematics of Control, Signals and Systems, 10 (1997), pp. 247–264.

[21] C.T. Lin, *Structural controllability*, IEEE Trans. Automat. Control, 19 (1974), pp. 201–208.

[22] F. Lin, M. Fardad, and M. Jovanovic, *Algorithms for leader selection in large dynamical networks: Noise-corrupted leaders*, in Proceedings of the 50th IEEE Conference on Decision and Control and European Control Conference (CDC-ECC), IEEE, 2011.

[23] F. Lin, M. Fardad, and M.R. Jovanovic, *Algorithms for leader selection in stochastically forced consensus networks*, IEEE Trans. Automat. Control, 59 (2014), pp. 1789–1802.

[24] Y.Y. Liu, J.J. Slotine, and A.L. Barabási, *Controllability of complex networks*, Nature, 473 (2011), pp. 167–173.

[25] S. Mangan and U. Alon, *Structure and function of the feed-forward loop network motif*, Proc. Natl. Acad. Sci. USA, 100 (2003), pp. 11980–11985.

[26] K. Murota, *Refined study on structural controllability of descriptor systems by means of matroids*, SIAM J. Control Optim., 25 (1987), pp. 967–989.

[27] ———, *Systems Analysis by Graphs and Matroids: Structural Solvability and Controllability*, vol. 3, Springer Science & Business Media, 1987.

[28] ———, *Matrices and Matroids for Systems Analysis*, vol. 20, Springer Science & Business Media, 2000.

[29] G.L. Nemhauser, L.A. Wolsey, and M.L. Fisher, *An analysis of approximations for maximizing submodular set functions - I*, Math. Program., 14 (1978), pp. 265–294.

[30] A. Olshevsky, *Minimum input selection for structural controllability*, arXiv preprint

arXiv:1407.2884, (2014).
[31] J.G. Oxley, *Matroid Theory*, Oxford University Press, 1992.
[32] F. Pasqualetti, S. Martini, and A. Bicchi, *Steering a leader-follower team via linear consensus*, Hybrid Systems: Computation and Control, (2008), pp. 642–645.
[33] F. Pasqualetti, S. Zampieri, and F. Bullo, *Controllability metrics and algorithms for complex networks*, arXiv preprint arXiv:1308.1201, (2013).
[34] S. Patterson and B. Bamieh, *Leader selection for optimal network coherence*, in Decision and Control (CDC), 2010 49th IEEE Conference on, IEEE, 2010, pp. 2692–2697.
[35] J. R. T. Lawton, R. W. Beard, and B. J. Young, *A decentralized approach to formation maneuvers*, IEEE Transactions on Robotics and Automation, 19 (2003), pp. 933–941.
[36] A. Rahmani, M. Ji, M. Mesbahi, and M. Egerstedt, *Controllability of multi-agent systems from a graph-theoretic perspective*, SIAM J. Control Optim., 48 (2009), pp. 162–186.
[37] K.J. Reinschke and G. Wiedemann, *Digraph characterization of structural controllability for linear descriptor systems*, Linear Algebra App., 266 (1997), pp. 199–217.
[38] W. Ren, *On consensus algorithms for double-integrator dynamics*, IEEE Trans. Automat. Control, 53 (2008), pp. 1503–1509.
[39] M. Rohden, A. Sorge, M. Timme, and D. Witthaut, *Self-organized synchronization in decentralized power grids*, Phys. Rev. Lett., 109 (2012), p. 064101.
[40] J. Ruths and D. Ruths, *Control profiles of complex networks*, Science, 343 (2014), pp. 1373–1376.
[41] A. Schrijver, *Combinatorial Optimization: Polyhedra and Efficiency*, Springer, 2003.
[42] T. H. Summers, F. L. Cortesi, and J. Lygeros, *On submodularity and controllability in complex dynamical networks*, arXiv preprint arXiv:1404.7665, (2014).
[43] H.G. Tanner, *On the controllability of nearest neighbor interconnections*, 43rd IEEE Conference on Decision and Control (CDC), 3 (2004), pp. 2467–2472.
[44] P.K.C. Wang and F.Y. Hadaegh, *Coordination and control of multiple microspacecraft moving in formation*, Journal Astronaut. Sci., 44 (1996), pp. 315–355.

8. Appendix. In this appendix, we provide proofs of Lemma 2.10 and Theorem 3.11.

The steps in the proof of Lemma 2.10 follow those in [26]. First, define a set of coefficients on the edges $\hat{E}$ of graph $\hat{G}$. The coefficient $\gamma(e)$ on edge $e \in \hat{E}$ is defined by

$$\gamma(e) = \begin{cases} r_i, & e = (w_i^T, w_i^Q), m(w_i^T) \neq w_i^Q \\ -r_i, & e = (w_i^T, w_i^Q), m(w_i^T) = w_i^Q \\ c_i, & e = (x_i^T, x_i^Q), i \notin J \\ -c_i, & e = (x_i^T, x_i^Q), i \in J \\ 1, & e = (w_i^T, x_j^T), m(w_i^T) \neq x_j^Q \\ -1, & e = (w_i^T, x_j^T), m(w_i^T) = x_j^Q \\ 0, & \text{else} \end{cases} \tag{8.1}$$

where r_i and c_j are integers satisfying

$$\begin{aligned} r_i - c_j &= 1 \qquad \text{if}(Q_F)_{ij} \neq 0 \\ r_i - c_j &= 0 \qquad \text{if}(Q_A)_{ij} \neq 0 \\ r_i - c_{n+j} &= 0 \qquad \text{if}(Q_B)_{ij} \neq 0 \end{aligned}$$

It is shown in [26] that the r_i and c_j are well-defined under the assumption that each nonvanishing subdeterminant of $(Q_A - sQ_F|Q_B)$ is of the form αs^p with a rational α and integer p.

Now, let $\tilde{\Omega} = \Omega\Omega_{J\cup J_1}^{-1}$ and $V^- = \{v^Q : \tilde{\Omega}_{vj} \neq 0 \text{ for some } j \in J_1\}$. The following appears as Theorem 4.7 of [26], albeit with slightly modified notation.

Theorem 8.1. *Let $\tilde{V}$ denote the set of vertices that are not reachable to V^-, and let $\tilde{G}$ denote the subgraph of $\hat{G}$ induced by $\tilde{V}$. Then the condition rank$((zF-A)|B) = n$*

holds for almost any choice of the free parameters if and only if the sum of the $\gamma(e)$ *along any directed cycle in* $\tilde{G}$ *is zero.*

We now prove Lemma 2.10.

Proof. [Proof of Lemma 2.10] If the condition of Lemma 2.10 holds, then $\tilde{V}$ does not contain any cycle. Hence the condition of Theorem 8.1 holds automatically, and we have $\text{rank}((zF-A)|B) = n$ for almost any free parameters. ◻

In order to prove Theorem 3.11, we first prove a sequence of lemmas, which will establish that $F(\mathbf{y}(1)) \geq (1-1/e)f(S^*)$. The lemmas follow Lemmas 3.1–3.3 of [2], however, the results of [2] are for a single matroid constraint, instead of two matroid basis constraints as in Theorem 3.11.

Lemma 8.2. *Let* $y \in [0,1]^n$ *and let* $R \subseteq V$ *denote a random set such that* $j \in R$ *with probability* y_j. *Then*

$$f(S^*) \leq F(y) + \max_{I \in \hat{\mathcal{B}}_1 \cap \hat{\mathcal{B}}_2} \sum_{j \in I} \mathbf{E}(f_R(j)),$$

where $f_R(j) = f(R \cup \{j\}) - f(R)$.

Proof. By submodularity, we have that $f(S^*) \leq f(R) + \sum_{j \in S^*} f_R(j)$. Taking expectation over R yields

$$f(S^*) \leq \mathbf{E}(f(R)) + \sum_{j \in S^*} \mathbf{E}(f_R(j)) \leq F(y) + \max_{I \in \mathcal{I}} \sum_{j \in I} \mathbf{E}(f_R(j)),$$

as desired. ◻

Lemma 8.2 establishes an optimality result for an idealized version of Algorithm 2 where computation is performed over the actual values of $\mathbf{E}(f_R(j))$ instead of estimates computed via random sampling. The following lemma introduces bounds on the errors introduced by sampling.

Lemma 8.3. *With high probability, at each time* t *the algorithm finds a set* $I(t)$ *such that*

$$\sum_{j \in I(t)} \mathbf{E}(f_{R(t)}(j)) \geq (1-2k\delta) f(S^*) - F(y(t)).$$

The proof is identical to Lemma 3.2 of [2] and is omitted. Finally, we prove the optimality bound on the solution to the continuous relaxation.

Lemma 8.4. *With high probability, the fractional solution* $y(1)$ *found by solving the continuous problem satisfies*

$$F(y) \geq \left(1 - \frac{1}{e} - \frac{1}{3d}\right) f(S^*).$$

The proof follows that of Lemma 3.3 of [2]. We now prove Theorem 3.11.

Proof. [Proof of Theorem 3.11] The fact that $F(y) \geq \left(1 - \frac{1}{3}\right) f(S^*)$ follows directly from Lemma 8.4. The feasibility of S follows from the SWAP_ROUND algorithm, which preserves membership in the bases of both matroids at each iteration and returns an integral solution, corresponding to a common basis of $\hat{\mathcal{B}}_1$ and $\hat{\mathcal{B}}_2$. The complexity of the procedure is dominated by the solution of the continuous problem, which is in turn determined by the cost of computing the maximum weight matroid intersection at each iteration. Since there are $O(n^2)$ iterations and the maximum weight matroid intersection has complexity $O(\tau n^3)$ [41], where τ is the cost of testing independence in $\hat{\mathcal{M}}_1$ and $\hat{\mathcal{M}}_2$, the overall complexity is $O(\tau n^5)$. ◻