

Received December 3, 2019, accepted December 27, 2019, date of publication January 1, 2020, date of current version January 10, 2020.

Digital Object Identifier 10.1109/ACCESS.2019.2963541

# Sparsest Input Selection for Controllability of Singular Systems via a Two-Step Greedy Algorithm

YAN ZHANG AND WANHONG ZHANG<sup>✉</sup>

Department of Chemical Machinery, Qinghai University, Xining 810016, China

Corresponding author: Wanhong Zhang (zwh0511@163.com)

This work was supported in part by the National Natural Science Foundation of China under Grant 61741115, and in part by the Qinghai Department of Science and Technology under Grant 2017-ZJ-953Q.

**ABSTRACT** In this paper, the problem of determining the sparsest input matrices to ensure controllability of linear singular systems is investigated. Firstly, it is shown that, determining the sparsest input matrices to ensure reachable controllability or complete controllability is NP-hard, even when the system ‘singularity’ is arbitrarily large. Secondly, submodular functions for singular systems are built, upon which greedy algorithms are developed to approximate the sparsest input matrices with guaranteed performance bounds for the case where there is no restriction on the number of independent inputs. Thirdly, a two-step greedy algorithm is proposed for determining the sparsest input matrices with a given number of inputs to ensure controllability. Compared with the existing algorithms for sparsest input selections, the proposed algorithm achieves better trade-off between the approximation performances and computation efficiency, which are demonstrated by two simulation examples.

**INDEX TERMS** Controllability, computational complexity, networked control systems, optimization.

## I. INTRODUCTION

Input selections under the objectives to meet/optimize certain system performances have long been active but challenging issues in control community [1]. In a recent paper [2], it is first shown that, given an autonomous system as (1), it is NP-hard to determine the minimal number of state variables that need to be actuated by an input to ensure system controllability,

$$\dot{x}(t) = Ax(t), \quad (1)$$

where  $x(t)$  is the state vector, and  $A$  is state transition matrix. A more general proposition is that, given a collection of possible input matrix columns, it is NP-hard to choose the minimal number of columns to form an input matrix so that the resulting system can be controllable [2], [3].

An input selection problem for a more general class of linear systems is considered in this paper, which is known as the singular system or the descriptor system. Analysis and synthesis of descriptor systems have received significant attention due to their widespread application in modeling

and control of variously actual systems, such as electrical systems, biological systems, economical systems, hybrid systems and intelligent transportation systems [4]–[8], [10], [11]. Additionally, the descriptor system involves the usual state-space formulation as a special form, and can establish some dynamical systems without state-space formulation. This paper focuses on determining the sparsest input matrices with a given number of inputs that ensure controllability of singular systems. Two points make this problem worthy of extra attention. One is that, criteria for complete controllability of singular systems are a little more complicated than that of the nonsingular systems described by the standard state-space model. Some related concepts, like modes, controllable subspaces, etc., are often defined upon the standard canonical forms of singular systems, which make extending input selection methods from nonsingular systems to singular ones not so straightforward. The other one is that, as argued in [9], the problem of determining the sparsest input matrices with guaranteed performance bounds to ensure controllability is far from completely settled for nonsingular systems with multiple eigenvalues. It is left for us to find algorithms with guaranteed bounds, acceptable

The associate editor coordinating the review of this manuscript and approving it for publication was Donatella Darsena<sup>✉</sup>.

approximation performances and relatively low computational complexity.

*Related Work:* Controllability and observability of large-scale networked dynamical systems have drawn the attention of control scientists [12]–[14]. As mentioned earlier, determining the minimal number of actuated states to ensure controllability for nonsingular systems has been proven to be NP-hard [2]. However, there is a simple greedy algorithm which can maximize the rank increase of controllability matrix in each iteration. It can achieve a logarithmic approximation factor and can obtain the best performance in polynomial time. Moreover, it has a closed form solution with the minimum input number to ensure controllability, which is equal to the maximum geometric multiplicity of the state transition matrix [15]. Note that, this is similar to actuator and sensor deployment problems investigated under structural framework in [12], [16]. And more specifically, system matrices are fixed with zeros or free parameters. For example, the matching theory [12] is used to provide the minimum input number required for the controllability of structure, which enable the problem of determining the minimum actuated states with structural controllability guarantee in polynomial time [16].

Apart from focusing on the binary concept of controllability, researchers also proposed some energy related metrics to quantify controllability and found appropriate actuator selection strategies to satisfy certain performance criteria [17]. Approaches for input selection based on this concept, including submodular optimization [3], a method of the relaxed control energy metric [18], have been developed in the literature. These investigations extend the binary concept of controllability to quantitative one, which deepens our insight into controllability of singular system. However, these methods either have higher computing complexities or have lower approximation performance.

Concepts of reachable controllability and complete controllability for descriptor systems seem to be first developed in [19]. Afterwards, various criteria have been proposed, including some rank based criteria [19], graph characterization for structural controllability [20], and matroid theory based criteria for mixed descriptor systems which entries of system matrices are either fixed constants, or free parameters [21]. Input selections for controllability of mixed descriptor systems have been considered in [22], under a dimensionless assumption defined therein. The main difference between problems considered in that paper and this one is that, the objective in [22] is to determine the minimal number of dedicated inputs, that is, an input is dedicated if it actuates only one state, whereas we will show that for a given singular system with fixed system matrices, selecting the minimal number of dedicated inputs is less complicated than selecting the sparsest input matrices with limited number of inputs.

*Main Contributions:* In this paper, we consider the problem of determining the sparsest input matrices with given number of inputs to ensure controllability of singular systems.

The contributions of this paper are in three aspects. Firstly, it is proven that determining the sparsest input matrices to ensure reachable controllability or complete controllability is NP-hard, even when the system ‘singularity’ is arbitrarily large. Secondly, submodular functions for singular systems are built, upon which greedy algorithms are developed to approximate the sparsest input matrices with guaranteed performance bounds for the case where there is no restriction on the number of independent inputs. Thirdly, a two-step greedy algorithm is proposed for determining the sparsest input matrices with a given number of inputs to ensure (reachable or complete) controllability. This algorithm has much less search space than the simple greedy one, and remarkably, has a submodularity-ratio like provable approximation bound. Compared to the existing methods, numerous numerical simulations show that this algorithm achieves better trade-off between the approximation performances and computation efficiency.

The organization of this paper is as follows. Section II gives the problem formulation, and Section III presents some preliminaries. Section IV gives computational complexities of the problems considered in this paper. Afterwards, Sections V and VI deal with algorithms for problems considered in this paper. Some numerical examples are provided in Section VII. Section VIII finally concludes this paper.

*Notations:*  $\mathbb{R}$ ,  $\mathbb{C}$ ,  $\mathbb{Z}$  and  $\mathbb{N}$  denote the set of real, complex, integral and nonnegative integral numbers, respectively. For a matrix  $M$ ,  $M_{ij}$  denotes its  $(i, j)$ -th entry if no confusion is made.  $\|M\|_0$  denotes the number of nonzero entries in a matrix  $M$ .  $M^\dagger$  denotes the Moore-Penrose pseudo-inverse of  $M$ . Let  $J, J_1, J_2$  be a set of integers, then  $M_J$  represents the submatrix of  $M$  comprising the columns with indices given by  $J$ , and  $M_{J_1, J_2}$  the submatrix of  $M$  comprising the rows indexed by  $J_1$  and columns indexed by  $J_2$ .  $|S|$  denotes the cardinality of a set  $S$ . By  $\text{diag}\{M_i\}_{i=1}^N$  ( $\text{col}\{M_i\}_{i=1}^N$ ) we denote the block diagonal matrix (compact matrix) with the  $i$ -th diagonal block (row block) being  $M_i$ .

## II. PROBLEM FORMULATION

Consider the following linear time invariant (LTI) system

$$E\dot{x}(t) = Ax(t) + Bu(t), \quad (2)$$

where  $x(t) \in \mathbb{R}^n$  is the state vector and  $u(t) \in \mathbb{R}^l$  is the input vector at time  $t$ ,  $A \in \mathbb{R}^{n \times n}$ ,  $B \in \mathbb{R}^{n \times l}$  are respectively the state transition matrix and input matrix. The constant matrix  $E \in \mathbb{R}^{n \times n}$  may be singular. System (2) or  $(E, A, B)$  is called the linear singular system or descriptor system. Obviously, when  $E$  is nonsingular, this system collapses to the standard state-space system [23].

If  $\det(sE - A) \neq 0$  for at least one  $s \in \mathbb{C}$ , where  $E$  may be singular, the pair  $(E, A)$  is called regular. System (2) is said to be solvable, if for any admissible inputs  $u(t) \in \mathbb{R}^l$ , there is a unique state solution  $x(t)$  satisfying Equation (2). It is known that system (2) is solvable, if and only if  $\det(sE - A) \neq 0$  for at least one  $s \in \mathbb{C}$ . In this paper, we always assume that system (2) is solvable.

In this work, we are interested in determining the sparsest input matrices to ensure controllability of the linear singular systems. Before proceeding that, the notion of reachable controllability (R-controllability) and complete controllability (C-controllability) for singular systems is introduced.

The admissible initial state is the initial state vector  $x(0) \in \mathbb{R}^n$  such that Equation (2) admits a differentiable solution of  $x(t)$  and  $u(t)$ . The set of all admissible states is denoted by  $R_0$ , which is called the reachable set.

**Definition 1 (R-controllability [19]):** System (2) is said to be R-controllable, if for any two states  $x_0$  and  $x_1$  in the set of admissible initial states  $R_0$ , there exists a finite time  $T$  and a differentiable control input  $u(t)|_{0 \leq t \leq T}$ , such that  $x(0) = x_0$ ,  $x(T) = x_1$ .

**Definition 2 (C-controllability [19]):** System (2) is said to be C-controllable, if for any two states  $x_0, x_1 \in \mathbb{R}^n$ , there exists a finite time  $T$  and a differentiable control input  $u(t)|_{0 \leq t \leq T}$ , such that  $x(0) = x_0, x(T) = x_1$ .

For notation simplicity, we just say  $(E, A, B)$  is R-controllable (C-controllable) if system (2) is so.

**Lemma 1 (Conditions for R-Controllability):** Suppose that system (2) is solvable. System (2) is R-controllable, if and only if

$$\text{rank}([sE - A, B]) = n, \forall s \in \mathbb{C}. \quad (3)$$

**Lemma 2 (Conditions for C-Controllability):** Suppose that system (2) is solvable. System (2) is C-controllable, if and only if

$$\text{rank}([sE - A, B]) = n, \quad \forall s \in \mathbb{C}, \quad (4)$$

$$\text{rank}([E, B]) = n. \quad (5)$$

The problems considered in this paper can be formulated as following.

**Problem 1:** Given the pair  $(E, A)$  in (2), determine

$$\min_{B \in \mathbb{R}^{n \times n}} \|B\|_0$$

$(E, A, B)$  is C-controllable (resp. R-controllable)

**Problem 2:** Given the pair  $(E, A)$  in (2) and an integer  $l \leq n$  such that there exists a  $B \in \mathbb{R}^{n \times l}$  making  $(E, A, B)$  C-controllable (resp. R-controllable). Determine

$$\min_{B \in \mathbb{R}^{n \times l}} \|B\|_0$$

$(E, A, B)$  is C-controllable (resp. R-controllable)

Problem 2 differs from Problem 1 in that, there is a restriction on the available number of independent inputs. Such scenario occurs, for example, in a leader-follower multi-agent system where the number of leaders is limited, or in a circuit system where the number of independent voltage sources is limited. It is shown in [9] that, these two problems are not necessarily equivalent for a nonsingular system where  $A$  has multiple eigenvalues. It is straightforward to see that, Problem 1 is always feasible, as the input matrix  $B = I_n$  always makes  $(E, A, B)$  C-controllable and R-controllable. Besides, Problem 1 can be seen as a special case of Problem 2 by setting  $l = n$ .

### III. PRELIMINARIES

This section presents some preliminaries for our further derivations.

#### A. GENERALIZED EIGENVALUES

The generalized eigenvalue problem of two matrices  $M, N \in \mathbb{R}^{n \times n}$  is to find a scalar  $\lambda$  and the corresponding vector  $\phi$  such that  $\phi M = \lambda \phi N$ . We call  $\lambda$  the generalized eigenvalue and  $\phi$  the corresponding left eigenvector of the pair  $(M, N)$ . There are many numerical algorithms to determine the generalized eigenvalues. See [24] and the references therein.

Using the generalized eigenvalue arguments, the following lemma states an equivalent condition for (3) to hold.

**Lemma 3:** Suppose that the solvability condition holds for system (2), i.e.,  $\det(sE - A) \neq 0$  for at least one  $s \in \mathbb{C}$ . Assume that there are  $m$  distinct generalized eigenvalues  $\{s_i\}_{i=1}^m$  for  $(E, A)$ . For each  $s_i$ , let  $X_i$  be the matrix consisting of the maximum number of linearly independent row vectors of  $\phi$ , where  $\phi$  satisfies  $\phi E = s_i \phi A$  (i.e.,  $X_i$  is a basis matrix of the left null space of  $s_i E - A$ ). Then, (3) holds, if and only if  $X_i B$  is of full row rank for each  $i = 1, \dots, m$ .

*Proof:* See the appendix. ■

#### B. SUBMODULARITY

Let  $\Omega$  be a finite set and  $2^\Omega$  be its the power set. A set function  $f : 2^\Omega \rightarrow \mathbb{R}$  assigns a real scalar to each subset of  $\Omega$ . A nonincreasing function  $f$  is a set function such that for all  $\mathcal{F}_1 \subseteq \mathcal{F}_2 \subseteq \Omega$ , it holds that  $f(\mathcal{F}_1) \geq f(\mathcal{F}_2)$ .

**Definition 3 (Submodularity [25]):** A set function  $f : 2^\Omega \rightarrow \mathbb{R}$  is submodular if for any sets  $\mathcal{F}_1, \mathcal{F}_2$  with  $\mathcal{F}_1 \subseteq \mathcal{F}_2$  and any element  $\omega \in \Omega \setminus \mathcal{F}_2$ , we have that

$$f(\mathcal{F}_1 \cup \{\omega\}) - f(\mathcal{F}_1) \geq f(\mathcal{F}_2 \cup \{\omega\}) - f(\mathcal{F}_2), \quad (6)$$

and supermodular if the reversed inequalities in (6) hold.

**Lemma 4 [25]:** Suppose that the set function  $f : 2^\Omega \rightarrow \mathbb{R}$  is submodular and nonincreasing. For the optimization problem  $\min_{\mathcal{F} \subseteq \Omega} |\mathcal{F}|$  under the constraint  $f(\mathcal{F}) \geq \alpha$ , where  $\alpha$  is a given threshold, the simple greedy algorithm (i.e., in each iteration choosing the element with the maximum increase in  $f(\mathcal{F})$ ) achieves a solution  $\mathcal{F}_{gre}$  that satisfies the following inequality

$$\frac{|\mathcal{F}_{gre}|}{|\mathcal{F}^*|} \leq 1 + \log\left(\frac{f(\Omega) - f(\emptyset)}{f(\Omega) - f(\mathcal{F}_{T-1})}\right),$$

where  $\mathcal{F}^*$  is the optimal solution,  $\mathcal{F}_{T-1}$  is the set returned at the second-last iteration of the greedy algorithm.

#### C. MATROID

Matroid is a structure that captures and generalizes the notion of linear independence in vector spaces. Given a finite set  $\Omega$  and elements together with a family  $\mathcal{F} = \{\mathcal{F}_1, \mathcal{F}_2, \dots\}$  of subsets of  $\Omega$ , the pair  $(\Omega, \mathcal{F})$  is a matroid if:

- $\emptyset \in \mathcal{F}$ ;
- if  $\mathcal{F}_1 \in \mathcal{F}$  and  $\mathcal{F}_2 \subseteq \mathcal{F}_1$ , then  $\mathcal{F}_2 \in \mathcal{F}$ ;
- if  $\mathcal{F}_1, \mathcal{F}_2 \in \mathcal{F}$  and  $|\mathcal{F}_1| = |\mathcal{F}_2| + 1$ , then there exist some  $\omega \in \mathcal{F}_1 \setminus \mathcal{F}_2$  satisfying  $\mathcal{F}_2 \cup \omega \in \mathcal{F}$ .

From the above definition,  $\Omega$  is also named as the ground set, and a member of  $\mathcal{F}$  is referred to as an independent set. For a matrix  $E$ , a linear matroid can be defined as  $\mathcal{N}(E) = (\Omega, \mathcal{F})$ , where  $\Omega$  is the set of indices of columns of  $E$ ,  $\mathcal{F}$  is the collection of indices of columns of  $E$  which are linearly independent, i.e.,  $\mathcal{F} = \{J : \text{rank}(E_J) = |J|\}$ . Given two matroids  $\mathcal{N}_1 = (\Omega, \mathcal{F}_1)$  and  $\mathcal{N}_2 = (\Omega, \mathcal{F}_2)$ , the matroid intersection  $\mathcal{N}_1 \cap \mathcal{N}_2$  is defined as the collection of all common independent sets of  $\mathcal{N}_1$  and  $\mathcal{N}_2$ . The cardinality of  $\mathcal{N}_1 \cap \mathcal{N}_2$ , represented by  $\rho(\mathcal{N}_1 \cap \mathcal{N}_2)$ , is defined as

$$\rho(\mathcal{N}_1 \cap \mathcal{N}_2) = \max\{|J| : J \in \mathcal{F}_1, J \in \mathcal{F}_2\}.$$

A structured matrix is a matrix whose entry is either a fixed zero or can take an arbitrary real value independently. The set of  $n \times m$  structured matrices is denoted by  $\{0, *\}^{n \times m}$ , where  $*$  denotes the entry which can take values independently. For a matrix  $M$  whose entries are functions of some free parameters, its generic rank, indicated by  $\text{grank}(M)$ , is the maximum rank it can achieve as the function of its free parameters.

**Lemma 5** [26]: Given  $X \in \mathbb{R}^{r \times n}$ ,  $\bar{B} \in \{0, *\}^{n \times l}$ , it is valid that

$$\text{grank}(X\bar{B}) = \max\{|J| : J \subseteq \{1, \dots, n\}, \text{rank}(X_{R,J}) = |J|, \text{grank}(\bar{B}_{J,C}) = |J|\},$$

in which  $R = \{1, \dots, r\}$ ,  $C = \{1, \dots, l\}$ .

The generic rank of  $X\bar{B}$  can be computed via matroid intersection algorithm, which has polynomial time complexity. From Lemma 5, it is easy to see that  $\text{grank}(X\bar{B}) = \rho(\mathcal{N}(X) \cap \mathcal{N}(\bar{B}^T))$ .

#### IV. COMPLEXITY ANALYSIS

In this section we prove that both Problem 1 and Problem 2 are NP-hard for the singular systems. We will first give the computational complexities and then present their proofs.

**Theorem 1:** When the purpose is to make the system C-controllability, Problem 1 and Problem 2 are NP-hard, even if the rank deficiency of  $E$  is arbitrarily large.

**Theorem 2:** When the purpose is to make the system R-controllability, Problem 1 and Problem 2 are NP-hard, even if the rank deficiency of  $E$  is arbitrarily large.

Since by setting  $l = n$ , Problems 1 and 2 are equivalent, which are both equivalent to finding the sparsest diagonal input matrix to ensure system controllability,<sup>1</sup> for the proofs of these two theorems it suffices to prove the NP-hardness of the latter problem. In the following, we will use  $e_{p \times q}$  to denote a  $p \times q$  matrix whose entries are all one,  $0_{n \times n}$  the  $n \times n$  zero matrix, and  $e_i^n$  the  $i$ th standard basis vector in the  $n$ -dimensional vector space. We will say  $(E, A)$  is  $k$  C-controllable (R-controllable), if there exists a diagonal matrix  $B$  with no more than  $k$  nonzero entries such that  $(E, A, B)$  is C-controllable (R-controllable).

<sup>1</sup>From Lemma 3, for any  $B_*$  which is an optimal solution to Problem 1, the  $n \times n$  diagonal matrix  $B_d$ , which is constructed by letting  $B_{d,i} = 1$  whenever the  $i$ th row of  $B_*$  is nonzero, is always a feasible solution to Problem 1, and  $\|B_d\|_0 \leq \|B_*\|_0$ .

Our proofs adopt some similar ideas from [2]. The reductions start from the set cover problem. Let  $\mathcal{C} = \{C_1, \dots, C_h\}$  be a collection of subsets of  $\{1, \dots, p\}$ . A set cover of  $\mathcal{C}$  is a set  $\{i_1, \dots, i_k\} \subseteq \{1, \dots, h\}$ , such that  $C_{i_1} \cup \dots \cup C_{i_k} = \{1, \dots, p\}$ , where  $k$  is called the cardinality or size of this set cover. It is known that, given an integer  $k$ , determining whether there exists a set cover with size  $k$  is NP-complete. The complete proofs are given in the following.

**Proof of Theorem 1:** Let  $\mathcal{C} = \{C_1, \dots, C_h\}$  be a collection of subsets of  $\{1, \dots, p\}$ . Without losing of generality, assume that  $C_1 \cup \dots \cup C_h = \{1, \dots, p\}$ . Construct an  $p \times h$  matrix  $C$  as  $C_{ij} = 1$  if the  $j$ th set contains  $i$ . Construct an  $(h+p) \times (h+p+q)$  matrix

$$X = \begin{bmatrix} I_h & 0_{h \times p} & e_{h \times q} \\ C_{p \times h} & I_p & 0_{p \times q} \end{bmatrix}.$$

From this construction, matrix  $X_{\mathcal{I}, \mathcal{I}}$  is block triangular with nonzero diagonal entries, and hence invertible, where  $\mathcal{I} = \{1, \dots, h+p\}$ . Therefore,  $X$  is of full row rank. Define matrices  $E = \text{diag}\{I_{h+p}, 0_{q \times q}\}$ , and  $\Lambda = \{1, 2, \dots, h+p\}$ . Now we will construct matrix  $A$  satisfying  $\Lambda X E = X A$  and  $\det(sE - A) \neq 0$ .

To this end, let  $a_i \in \mathbb{R}^{h+p+q}$  satisfy  $X a_i = \Lambda X e_i^{h+p+q}$ , for  $i = 1, \dots, h+p$ . Since  $\text{rank}(X) = \text{rank}([X, \Lambda X e_i^{h+p+q}])$ , such  $a_i$  always exists. Let  $X^\perp \in \mathbb{R}^{(h+p+q) \times q}$  be a basis matrix that spans the right null space of  $X$ . Since the first  $h+p$  columns of  $X$  are linearly independent, it can be validated that  $X_{\mathcal{J}_1, \mathcal{J}_2}^\perp$  is invertible, where  $\mathcal{J}_1 = \{h+p+1, \dots, h+p+q\}$ ,  $\mathcal{J}_2 = \{1, \dots, q\}$ . Construct  $A$  as  $A = [a_1, \dots, a_{h+p}, X^\perp]$ .

Based on these constructions, it can be validated that  $\Lambda X E = X A$ . Therefore,  $X_i$ , which is the  $i$ th row of  $X$ , is the left generalized eigenvector of  $(E, A)$  associated with the generalized eigenvalue  $i$ . Moreover, there is a monomial  $s^{p+h} \det X_{\mathcal{J}_1, \mathcal{J}_2}^\perp$  in  $\det(sE - A)$  that cannot be zeroed out by other monomials. Hence,  $(E, A)$  is regular. As  $\text{rank}(E) = h+p$ ,  $E$  is singular and its rank deficiency can be arbitrarily large. Obviously, the above mentioned  $A$  can be constructed in polynomial time by solving linear matrix equations. In the following, we will prove that the system associated with  $(E, A)$  is  $k+q$  C-controllable, if and only if the collection  $\mathcal{C}$  has a set cover with size  $k$ .

For the one direction, assume that  $\mathcal{C}$  has a set cover  $\{i_1, \dots, i_k\}$ . Let the input matrix  $B \in \mathbb{R}^{(h+p+q) \times (h+p+q)}$  be such that  $B_{jj} = 1$  for  $j = i_1, \dots, i_k$ , and  $B_{jj} = 1$  for  $j = h+p+1, \dots, h+p+q$ , and other entries are zero. Then, from Lemma 2, it can be seen that  $(E, A, B)$  is C-controllable. Thus,  $(E, A)$  is  $k+q$  C-controllable.

For the other direction, assume that  $(E, A)$  is  $k+q$  C-controllable. This means that, there exists a diagonal matrix  $B \in \mathbb{R}^{(h+p+q) \times (h+p+q)}$  with no more than  $k+q$  nonzero entries, so that  $(E, A, B)$  is C-controllable. According to Lemma 2, to make  $\text{rank}([E, B]) = h+p+q$ , the  $(h+p+1)$ th,  $(h+p+2)$ th,  $\dots$ ,  $(h+p+q)$ th diagonal entries of  $B$  must be nonzero. On the basis of that, to make  $(E, A, B)$  C-controllable for a diagonal  $B$ , it is necessary and sufficient

to make the support <sup>2</sup> of the  $(h + i)$ th row of  $X$  intersects with  $\mathcal{J}$ , for each  $i \in \{1, \dots, p\}$ , where  $\mathcal{J} = \{j : B_{jj} \neq 0, j \in \{1, \dots, p+h\}\}$  and  $|\mathcal{J}| \leq k$ . As every element of  $\{1, \dots, p\}$  is contained in  $\mathcal{C}_1 \cup \dots \cup \mathcal{C}_h$ , for each  $j \in \mathcal{J} \cap \{h+1, \dots, h+p\}$ , we can find a replacement  $j^* \in \{1, \dots, h\}$ , such that  $C_{jj^*} = 1$  and thus the support of the  $j$ th row of  $X$  intersects with  $\mathcal{J} \setminus \{j\} \cup \{j^*\}$ . As  $|\mathcal{J}| \leq k$ , we finally have that the collection  $\mathcal{C}$  has a cover set with size no more than  $k$ . The NP-hardness of Problem 1 for regular singular systems then follows the NP-completeness of the set cover problem. ■

*Proof of Theorem 2:* Again we use the reduction from the set cover problem. The reduction is almost the same as that of the C-controllability case. The only difference is that, there is no need to consider the condition (5). A little bit modifications are sufficient to fit for this purpose. To this end, let  $E = \text{diag}\{I_{h+p+1}, 0_{q \times q}\}$ , where  $q > 1$  and the ‘singularity’  $q$  can be arbitrarily large. Parameter matrix  $X \in \mathbb{R}^{(h+p+1) \times (h+p+1+q)}$  becomes

$$X = \begin{bmatrix} I_h & 0_{h \times p} & e_{h \times 1} & e_{h \times q} \\ C_{p \times h} & I_p & 0 & 0_{p \times q} \\ 0 & 0 & e_{1 \times 1} & 0 \end{bmatrix}.$$

It can be verified that the first  $(h + p + 1)$  columns of  $X$  are linearly independent. Hence,  $X$  is of full row rank. Let  $E = \text{diag}\{I_{h+p+1}, 0_{q \times q}\}$ . Afterwards, construct matrix  $A \in \mathbb{R}^{(h+p+q+1) \times (h+p+q+1)}$  in the similar way to the C-controllability case. It can be verified that,  $(E, A)$  is regular. Then, following similar arguments to the proof of Theorem 1, it can be claimed that  $(E, A)$  is  $k$  R-controllable, if and only if the collection  $\mathcal{C}$  has a set cover with size no more than  $k + 1$ . ■

*Remark 1:* As mentioned above, our main ideas follow those of [2], but with some technical differences. More specific, to generalize their proofs to the singular case, as the rank deficiency of  $E$  (which can be seen as the ‘singularity’ of the system) can be arbitrarily large, it takes extra attention to guarantee that  $(E, A)$  is regular in the constructions. Besides, the conditions for C-controllability are more complicated than those of the Popov-Belevitch-Hautus (PBH) test for the nonsingular state-space systems.

**V. SUBMODULARITIES FOR SINGULAR SYSTEMS AND SOLUTIONS TO PROBLEM 1**

In this section, we identify some submodular functions for linear singular systems that facilitate approximation algorithms for Problem 1, as well as the two-step greedy algorithm in the next section for Problem 2.

First, we consider the dimension of reachable subspace. Unlike the linear nonsingular case, the reachable state set of linear singular systems is often hard to be represented in the original forms of their system parameters [19]. A transformation is utilized to obtain the standard canonical form of system (2). To be specific, there exist two nonsingular

<sup>2</sup>The support of a vector is the set of indices of nonzero entries of this vector.

matrices  $P$  and  $Q$ , such that

$$PEQ = \text{diag}\{I_{n_1}, E_2\}, PAQ = \text{diag}\{A_1, I_{n_2}\}, PB = \text{col}\{B_1, B_2\},$$

where  $E_2 \in \mathbb{R}^{n_2 \times n_2}$  is nilpotent, i.e.,  $E_2^k = 0$  for some integer  $k$ , and integers  $n_1 + n_2 = n$ . The standard canonical form of system (2) is given as follows:

$$\begin{aligned} \dot{x}_1(t) &= A_1 x_1(t) + B_1 u(t) \\ E_2 \dot{x}_2(t) &= x_2(t) + B_2 u(t) \end{aligned}$$

with state transformation  $\text{col}\{x_1(t), x_2(t)\} = Q^{-1}x(t)$ , where  $x_1(t) \in \mathbb{R}^{n_1}$  and  $x_2(t) \in \mathbb{R}^{n_2}$ . Given  $A_i \in \mathbb{R}^{n_i \times n_i}$  and  $B_i \in \mathbb{R}^{n_i \times l}$ , let  $\langle A_i | B_i \rangle$  denote the space spanned by  $[B_i, A_i B_i, \dots, A_i^{n_i-1} B_i]$ ,  $i = 1, 2$ . According to [17], the reachable subspace  $R_{rs}$  of system (2) is given by

$$R_{rs} = \langle A_1 | B_1 \rangle \oplus \langle E_2 | B_2 \rangle,$$

where  $\oplus$  denotes direct sum.

*Proposition 1:* Let  $\dim R_{rs}(B)$  be the dimension of reachable subspaces of system (2) with input matrix  $B$ . For any given  $B \in \mathbb{R}^{n \times l}$ ,  $\dim R_{rs}(B_S)$  is submodular on  $S \subseteq \{1, \dots, l\}$ .

*Proof:* See the appendix. ■

However, utilizing the dimension of reachable subspaces based on the standard canonical form to approximate Problem 1, may be not natural and cause extra computational burdens. In the following, we give another submodular function based on the generalized eigenvectors.

Let  $m, s_i |_{i=1}^m$  and  $X_i |_{i=1}^m$  be defined as in Lemma 3. Moreover, let  $r_i$  be the number of rows of  $X_i$ . It is valid that  $\sum_{i=1}^m r_i \leq \text{rank}(E)$ . Define a set function  $f(S)$  on  $S \subseteq \{1, \dots, n\}$  as

$$f(S) = \sum_{i=1}^m \text{rank}(X_i I_S) + \text{rank}([E, I_S]).$$

By Lemma 3, condition (3) is satisfied, if and only if  $\text{rank}(X_i B) = r_i$  for each  $i \in \{1, \dots, m\}$ . Hence,  $(E, A, B)$  is C-controllable, if and only if  $\sum_{i=1}^m \text{rank}(X_i B) + \text{rank}([E, B]) = \sum_{i=1}^m r_i + n$ . Based on this fact and the analysis below Theorem 2, Problem 1 has the same optimal value as the following problem:

$$\begin{aligned} \min_{S \subseteq \{1, \dots, n\}} & |S| \\ f(S) &= \sum_{i=1}^m r_i + n. \end{aligned}$$

*Theorem 3:* The greedy algorithm based on  $f(S)$  (see Algorithm 1) achieves an  $O(\log(n))$ -approximation for Problem 1 when the objective is to ensure C-controllability.

*Proof:* See the appendix. ■

*Corollary 1:* Define a set function  $\bar{f}(S)$  on  $S \subseteq \{1, \dots, n\}$  as

$$\bar{f}(S) = \sum_{i=1}^m \text{rank}(X_i I_S)$$

The greedy algorithm based on  $\bar{f}(S)$  (similar to Algorithm 1, the termination condition is  $\bar{f}(S) = \sum_{i=1}^m r_i$ ) achieves an  $O(\log \text{rank}(E))$ -approximation for Problem 1 when the objective is to ensure R-controllability.

*Proof:* See the appendix. ■

**Algorithm 1** Approximation Algorithm for the Problem 1

**Require:**  $E, A$

**Ensure:** Approximation of the minimal set  $S$  such that  $(E, A, I_S)$  is C-controllable

- 1: Calculate  $X_i$  of  $(E, A)$ , for  $i = 1, \dots, m$ ;
- 2: Initialize  $V \leftarrow \{1, \dots, n\}, S \leftarrow \emptyset$ ;
- 3: **while**  $f(S) < \sum_{i=1}^m r_i + n$  **do**
- 4:    $s \leftarrow \arg \max_{\omega \in V \setminus S} f(S \cup \{\omega\}) - f(S)$ ;
- 5:    $S \leftarrow S \cup \{s\}$ ;
- 6: **end while**
- 7: **return**  $S$ .

This section has identified some submodular functions which facilitate approximation algorithms for Problem 1 on singular systems. In fact, these results are consistent with nonsingular systems. Algorithms in this section are the first step towards the two-step greedy algorithm developed in the next section.

**VI. TWO-STEP GREEDY ALGORITHM FOR PROBLEM 2**

We now proceed with Problem 2 when there is a fixed number of available inputs. A two-step greedy algorithm with provable performance bounds will be proposed.

**A. FEASIBILITY CONDITION AND NONSUBMODULARITY**

For the ease of derivations, let  $r_{m+1}$  be the dimension of the left null space of  $E$  (i.e.,  $r_{m+1} = n - \text{rank}(E)$ ), and  $X_{m+1}$  be the matrix consisting of a collection of linearly independent row vectors which span the left null space of  $E$ . Then, from Lemma 3, system (2) is C-controllable, if and only if  $X_i B$  is of full row rank for  $i = 1, \dots, m + 1$ .

The following proposition gives the feasibility condition for Problem 2.

*Proposition 2:* Let  $r_{\max} = \max_{1 \leq i \leq m} r_i$ . The minimal number of inputs to ensure the C-controllability equals  $\max\{r_{\max}, n - \text{rank}(E)\}$ .

*Proof:* The necessity is obvious. To show the sufficiency, let  $\Gamma$  be the vector consisting of all entries of  $B$ , and write  $B$  as a function  $B(\Gamma)$  of  $\Gamma$ . Denote  $l_{\max} = \max\{r_{\max}, n - \text{rank}(E)\}$ . Let  $\mathcal{P}(i) = \{\Gamma \in \mathbb{R}^{n \times l_{\max}} : X_i B(\Gamma) \text{ is not of full row rank}\}$ . It is easy to see that  $\mathcal{P}(i)$  has zero Lebesgue measure in  $\mathbb{R}^{n \times l_{\max}}$ . Define a set  $\mathcal{P}$  by  $\mathcal{P} = \mathbb{R}^{n \times l_{\max}} \setminus \bigcup_{i=1}^{m+1} \mathcal{P}(i)$ . As  $m+1$  is countable,  $\mathcal{P}$  is open and dense in  $\mathbb{R}^{n \times l_{\max}}$ . As a result, a real valued matrix always exists in  $\mathcal{P}$ , which makes the associated system C-controllable. ■

Proposition 2 generalizes the known result that, the minimal number of inputs to guarantee controllability of a standard state-space described system equals the maximum geometric multiplicity of the system state transition

matrix [13]. As a result of the above mentioned proposition, Problem 2 is feasible if and only if  $l \geq \max\{r_{\max}, n - \text{rank}(E)\}$ . We denote  $l_{\max} \triangleq \max\{r_{\max}, n - \text{rank}(E)\}$ . In the following, we always assume that Problem 2 is feasible, i.e.,  $l \geq l_{\max}$ .

It is natural to consider the following function:

$$g(\bar{B}) = \sum_{i=1}^m \text{grank}(X_i \bar{B}) + \text{grank}([E, \bar{B}])$$

where  $\bar{B} \in \{0, *\}^{n \times l}$ , as a performance function<sup>3</sup> for Problem 2. Since  $\text{grank}([E, \bar{B}]) = n - r_{m+1} + \text{grank}(X_{m+1} \bar{B})$ , an equivalent expression of  $g(\bar{B})$  is  $g(\bar{B}) = n - r_{m+1} + \sum_{i=1}^{m+1} \text{grank}(X_i \bar{B})$ . It can be verified that, Problem 2 has the same optimal value as the following problem:

$$\begin{aligned} & \min_{\bar{B} \in \{0, *\}^{n \times l}} \|\bar{B}\|_0 \\ & \text{s.t. } g(\bar{B}) = \sum_{i=1}^{m+1} r_i. \end{aligned}$$

In fact, although solutions to the above mentioned optimization problem are structured matrices, which characterize the zero-nonzero patterns of the input matrices, it has been proven in [11] that once we have the feasible zero-nonzero patterns  $\bar{B}$  of the input matrices, a numerical realization of  $\bar{B}$  making each  $X_i \bar{B}$  full of row rank, thus making the system C-controllable,  $i = 1, \dots, m + 1$ , can be determined in polynomial time. Hence, in what follows, we mainly focus on the zero-nonzero patterns of the input matrix  $B$ , which is denoted by  $\bar{B}$ .

With a little abuse of notation, given  $\bar{B}_1, \bar{B}_2 \in \{0, *\}^{n \times l}$ , we say  $\bar{B}_1 \subseteq \bar{B}_2$ , if  $\bar{B}_{1ij} \neq 0$  implies  $\bar{B}_{2ij} \neq 0$ . By  $\bar{B}_2 \setminus \bar{B}_1$  we denote a structured matrix whose  $(i, j)$ th entry is nonzero only if  $\bar{B}_{2ij} = *$  and  $\bar{B}_{1ij} = 0$ . An element  $\bar{e}_{i,j} \in \bar{B}_1$  is such that  $\bar{e} \in \{0, *\}^{n \times l}$ , and there is only one nonzero entry in  $\bar{e}$ , which is the  $(i, j)$ th entry with  $(i, j)$  satisfying  $\bar{B}_{1ij} \neq 0$ . Let  $\bar{B}_{full}$  be the  $n \times l$  structured matrix whose entries are all nonzero. It is easy to see that  $g(\bar{B})$  is a set function on  $\bar{B} \subseteq \bar{B}_{full}$ .<sup>4</sup> However,  $g(\bar{B})$  is usually neither submodular nor supermodular. This can be seen from the following example.

*Example 1:* Consider  $E = I_2, A = \begin{bmatrix} 2 & 0 \\ 0 & 2 \end{bmatrix}$ . Let  $\bar{B}_1 = \begin{bmatrix} * & 0 \\ 0 & 0 \end{bmatrix}, \bar{e}_{1,2} = \begin{bmatrix} 0 & * \\ 0 & 0 \end{bmatrix}$  and  $\bar{B}_2 = \begin{bmatrix} * & 0 \\ * & 0 \end{bmatrix}$ . It can be validated that

$$g(\bar{B}_1 \cup \bar{e}_{1,2}) - g(\bar{B}_1) = 0, \quad g(\bar{B}_2 \cup \bar{e}_{1,2}) - g(\bar{B}_2) = 1.$$

Consequently,  $g(\bar{B}_2 \cup \bar{e}_{1,2}) - g(\bar{B}_2) > g(\bar{B}_1 \cup \bar{e}_{1,2}) - g(\bar{B}_1)$  while  $\bar{B}_1 \subseteq \bar{B}_2$ , which shows that  $g(\bar{B})$  is nonsubmodular. On the other hand, when  $r_i = 1$  for  $i = 1, \dots, m + 1$ , it is easy to see that  $g(\bar{B})$  is submodular (but nonmodular), which means that  $g(\bar{B})$  is usually not supermodular.

<sup>3</sup>In this paper, by performance function, we simply mean that whether the constraint conditions of the associated optimization problem are satisfied can be judged by checking the value of this function.

<sup>4</sup>For simplicity of notations,  $\bar{B}$  can either be seen as a set (of nonzero entries), or a structured matrix.

**B. SUBMODULARITY-RATIO FAILS TO GIVE NON-TRIVIAL PERFORMANCE BOUNDS FOR  $G(\bar{B})$**

Recently, submodularity-ratio, which is used to measure how far a set function is from submodularity, has been adopted to give some performance bounds for greedy algorithms for nonsubmodular functions [27].

*Definition 4 (Submodularity-Ratio, [27]):* Let  $h$  be a non-negative set function on  $V$ . For any  $S, T \subseteq V$ , define  $h_S(T) = h(T \cup S) - h(S)$ . The submodularity-ratio of  $h$  is defined as

$$\gamma_h = \min_{S, T \subseteq V} \frac{\sum_{t \in T \setminus S} h_S(\{t\})}{h_S(T)}.$$

The submodularity-ratio  $\gamma_h$  is such that  $0 \leq \gamma_h \leq 1$ , and is 1 if and only if  $h$  is submodular. The  $\gamma_h$  usually gives some nontrivial performance bounds for the greedy algorithm of a nonsubmodular function, but fails to do so when  $\gamma_h = 0$ .<sup>5</sup> However, the following example shows that,  $g(\bar{B})$  indeed has a zero submodularity-ratio.

*Example 2 Example 1 Continued, Zero Submodularity-Ratio of  $g(\bar{B})$ :* Consider the parameters in Example 1. It holds that  $g(\bar{B}_1 \cup \{\bar{e}_{1,2}\}) - g(\bar{B}_1) = 0$ ,  $g(\bar{B}_1 \cup \{\bar{e}_{2,1}\}) - g(\bar{B}_1) = 0$ , but  $g(\bar{B}_1 \cup \{\bar{e}_{1,2} \cup \bar{e}_{2,1}\}) - g(\bar{B}_1) = 1$ , which means that  $\gamma_g = 0$ .

**C. A TWO-STEP GREEDY ALGORITHM**

We propose an alternative for the simple greedy algorithm, which is a two-step greedy algorithm. This algorithm is an overall greedy algorithm, but consists of two steps. The first step is to approximate a minimal set  $S$  such that  $(E, A, I_S)$  is C-controllable. Here  $S$  is the set of directly actuated states. The second step is to select elements from  $(\bar{B}_{full})_{S,C}$ , where  $C = \{1, \dots, l\}$ , that is, to select elements from the rows of  $B_{full}$  indexed by  $S$ . The detailed description is given in Algorithm 2. Remarkably, this algorithm has a submodularity-ratio like provable approximation bound, as shown in the following theorem.

*Theorem 4:* The two-step greedy algorithm achieves an  $O(l_{max} \log(n))$ -approximation for Problem 2.

To prove Theorem 4, we first prove the following two intermediate results.

*Lemma 6:* If  $g(\bar{B}) < \sum_{i=1}^{m+1} r_i$ , then there exists  $\bar{e} \in \bar{B}_0 \setminus \bar{B}$ , such that  $g(\bar{B} \cup \{\bar{e}\}) - g(\bar{B}) \geq 1$ , where  $\bar{B}_0$  is defined in Algorithm 2.

*Proof:* Suppose that  $g(\bar{B}) < \sum_{i=1}^{m+1} r_i$ . Then, there exists at least one  $i \in \{1, \dots, m+1\}$ , such that  $grank(X_i \bar{B}) < r_i$ . Assume that  $grank(X_i \bar{B}) = r'_i < r_i$  for one integer  $i$ . Let  $R = \{1, \dots, r_i\}$ ,  $C = \{1, \dots, l\}$ . According to Lemma 5, there is  $J \subseteq S$ , such that  $rank(X_{iR,J}) = |J| = r'_i$ , and  $grank(\bar{B}_{J,C}) = |J| = r'_i$ . Since  $(E, A, I_S)$  is C-controllable, one has that  $rank(X_i I_S) = r_i$ . Therefore, there exists at least one  $j \in S \setminus J$ , such that  $grank(\bar{X}_{iR,J \cup \{j\}}) = r'_i + 1$ . Moreover,  $grank(\bar{B}_{J,C}) = |J|$  implies that there is a  $C' \subseteq C$  satisfying

$grank(\bar{B}_{J,C'}) = |J|$ ,  $|C'| = |J|$ . Hence, let  $\bar{e}_{j,i} \in \{0, *\}^{n \times l}$  be the structured matrix with its only nonzero entry being the  $(j, i)$ th entry, where  $i \in C \setminus C'$ . Hence,  $grank(\bar{B} \cup \{\bar{e}_{j,i}\}) = r'_i + 1$ . From Lemma 5, it turns out that  $grank(X_i(\bar{B} \cup \{\bar{e}_{j,i}\})) = r'_i + 1$ , meanwhile  $j \in S$ . As a result, there exists  $\bar{e} \in \bar{B}_0 \setminus \bar{B}$ , such that  $g(\bar{B} \cup \{\bar{e}\}) - g(\bar{B}) \geq 1$ . ■

The above lemma means that, before the termination of Algorithm 2, in each iteration when an element is added to  $\bar{B}$ , the increase of the associated performance function (i.e.,  $g(\bar{B})$ ) is at least one.

**Algorithm 2** A Two-Step Algorithm for Problem 2

**Require:**  $E, A$ , a fixed number of inputs  $l$  ( $l \geq l_{max}$ )

**Ensure:** Zero-nonzero patterns of the approximated sparsest input matrices for Problem 2

- 1: Calculate  $X_i$  of  $(E, A)$ , for  $i = 1, \dots, m+1$ ;

**Step 1:**

- 2: Approximate a minimal set  $S$  such that  $(E, A, I_S)$  is C-controllable using Algorithm 1;

**Step 2:**

- 3: Initialize  $\bar{B}_0 \leftarrow (\bar{B}_{full})_{S,C}$ ,  $\bar{B}_{alg} \leftarrow \emptyset$ , where  $C = \{1, \dots, l\}$ ;
- 4: **while**  $g(\bar{B}_{alg}) < \sum_{i=1}^{m+1} r_i$  **do**
- 5:    $\bar{e} \leftarrow \arg \max_{\bar{e} \in \bar{B}_0 \setminus \bar{B}_{alg}} g(\bar{B}_{alg} \cup \{\bar{e}\}) - g(\bar{B}_{alg})$ ;
- 6:    $\bar{B}_{alg} \leftarrow \bar{B}_{alg} \cup \{\bar{e}\}$ ;
- 7: **end while**
- 8: **return**  $\bar{B}_{alg}$

*Lemma 7:* Let  $\bar{B}_{alg}$  be the structured input matrix obtained from Algorithm 2. Then, there is at most  $l_{max}$  nonzero entries in each row of  $\bar{B}_{alg}$ .

*Proof:* Let  $C = \{1, \dots, l\}$ ,  $R = \{1, \dots, n\}$ . Suppose that there are  $l_{max}$  nonzero entries in the  $i$ th row of  $\bar{B}$ ,  $i \in R$ . Denote the set of column indices of these nonzero entries by  $I^*$ ,  $I^* \subseteq C$ . We will prove that, for each  $j^+ \in C \setminus I^*$ ,  $g(\bar{B} \cup \{\bar{e}_{i,j^+}\}) - g(\bar{B}) = 0$ . That is, adding  $\bar{e}_{i,j^+}$  to  $\bar{B}$  cannot increase the value of the performance function  $g(\bar{B})$ . Assume that there exists one  $j \in \{1, \dots, m+1\}$ , such that  $grank(X_j(\bar{B} \cup \{\bar{e}_{i,j^+}\})) - grank(X_j \bar{B}) = 1$ . Denote  $\bar{B}_e \triangleq \bar{B} \cup \{\bar{e}_{i,j^+}\}$ . Suppose that  $grank(X_j \bar{B}_e) = r^*$ . According to Lemma 5, it indicates that there are  $J^* \subseteq R$  and  $C^* \subseteq C$  with  $|J^*| = |R^*| = r^*$ ,  $j^+ \in C^*$  and  $i \in J^*$ , such that  $rank(X_{jR^*,J^*}) = r^*$  and  $grank(\bar{B}_{eJ^*,C^*}) = r^*$ . If  $grank(\bar{B}_{eJ^* \setminus \{i\}, C^* \setminus \{j^+\}}) < r^* - 1$ , to make  $grank(\bar{B}_{eJ^*,C^*}) = r^*$ , it must hold that  $grank(\bar{B}_{J^*,C^*}) = r^*$ . Then immediately  $grank(X_j \bar{B}) = r^*$ , which means that  $grank(X_j \bar{B}_e) - grank(X_j \bar{B}) = 0$ , causing a contradiction. On the other hand, assume  $grank(\bar{B}_{eJ^* \setminus \{i\}, C^* \setminus \{j^+\}}) = r^* - 1$ . Since  $|C^* \setminus \{j^+\}| < l_{max}$ , and  $|J^*| = l_{max}$ , there always exists a  $j^* \in I^* \setminus (C^* \setminus \{j^+\})$ , such that  $grank(\bar{B}_{J^*, \{j^*\} \cup (C^* \setminus \{j^+\})}) = r^*$ . Recalling that  $rank(X_{jR^*,J^*}) = r^*$ , it holds that  $grank(X_j \bar{B}) = r^*$  by Lemma 5. This immediately leads to that  $grank(X_j \bar{B}_e) - grank(X_j \bar{B}) = 0$ , causing a contradiction again. Hence, adding any extra element to the rows of  $\bar{B}$  with  $l_{max}$  nonzero entries cannot increase the performance

<sup>5</sup>The approximation ratio is  $O(\gamma_h^{-1} \log \frac{h(\emptyset)}{h(S_{T-1})})$ , where  $S_{T-1}$  is the selected set at the second-last iteration of the greedy algorithm.

function. Combined with Lemma 6, this proves the proposed claim. ■

*Proof of Theorem 4:* Let  $\bar{B}_{alg}$  be the structured input matrix obtained from Algorithm 2, and  $S_*$  and  $\bar{B}_*$  respectively the optimal solutions to Problems 1 and 2. Recall that  $S$  is the set returned from Algorithm 1. First, from Theorem 3, the number of nonzero rows of  $\bar{B}_{alg}$  is at least  $|S|$ , which is no more than  $(1 + \log(n))$  times of  $|S_*|$ . Second, from Lemma 5, it turns out that  $\|\bar{B}_{alg}\|_0 \leq l_{\max}|S|$ . Obviously  $\|\bar{B}_*\|_0 \geq |S_*|$ . Hence, it holds that  $\|\bar{B}_{alg}\|_0 \leq l_{\max}|S| \leq l_{\max}(1 + \log(n))|S_*| \leq l_{\max}(1 + \log(n))\|\bar{B}_*\|_0$ . ■

Let us have a brief comparison between the two-step greedy algorithm and the simple greedy one. The worst-case approximation ratio  $O(l_{\max} \log(n))$  for the former algorithm has a form similar to the performance bound of using non-trivial submodularity-ratios for nonsubmodular functions. By contrast, currently we have not obtained a non-trivial performance bound for the simple greedy algorithm. Suppose that the second step of the two-step algorithm and the simple greedy algorithm both terminate after  $N_{ite}$  iterations. Then, the ratio of size of total search spaces of these two algorithms is  $\frac{n|S|+|S|N_{ite}}{nN_{ite}}$ , where  $S$  denotes the set returned in the first step of Algorithm 2. Usually  $|S|$  is much less than  $n$  [2]. Hence, this ratio is generally much smaller than 1. It is expected that, the two-step greedy algorithm is more computationally efficient than the simple greedy one. Our numerical experiments in the next section will show that, the two-step greedy algorithm almost has the same approximation performance as that of the simple greedy one, while is more computationally efficient.

*Remark 2:* By changing  $g(\bar{B})$  to be  $\bar{g}(\bar{B}) = \sum_{i=1}^m r_i$ , Algorithm 2 can be directly modified to fit for selecting the sparsest input matrices with a given number of inputs to ensure R-controllability. Details are omitted due to its obviousness.

**VII. NUMERICAL RESULTS**

We present two numerical examples to show the main results of this paper.

First, we calculate a descriptor system defined by an electric circuit in [20]. In this example, the system dynamics are defined by  $E\dot{x}(t) = Ax(t) + Bu(t)$ , where  $Bu(t)$  are the alternative input options, and the system state  $x(t) = [v_1(t), v_2(t), v_3(t), v_4(t), i_1(t), i_2(t), i_3(t), i_4(t)]$  is defined by the voltage and current at each element of the circuit network of Fig.1. We have

$$E = \begin{bmatrix} 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix},$$

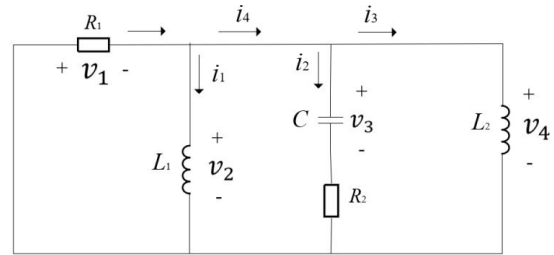
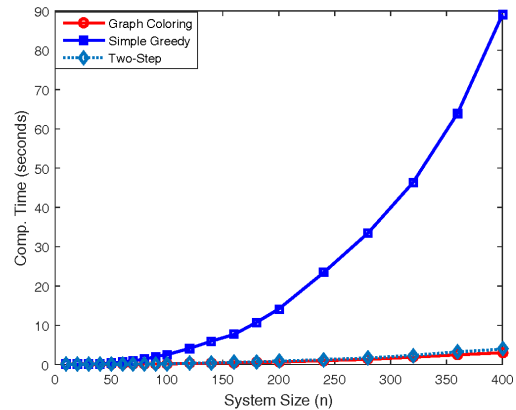
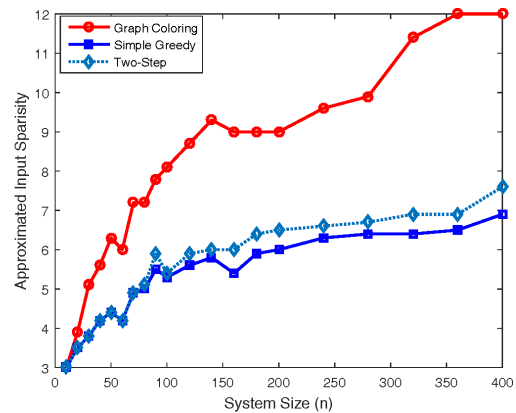


FIGURE 1. An electric circuit diagram.



(a)



(b)

FIGURE 2. Averaged running time (a) and input sparsity (b) when  $l = l_{\max} = 3$ .

$$A = \begin{bmatrix} 0 & 1/L_1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1/C & 0 & 0 \\ 0 & 0 & 0 & 1/L_2 & 0 & 0 & 0 & 0 \\ -1/R_1 & 0 & 0 & 0 & 1 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 & 1 & 1 & -1 \\ -1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 1 & 0 & R_2 & 0 & 0 \\ 0 & 1 & 1 & 0 & 0 & R_2 & 0 & 0 \end{bmatrix}.$$

For each element of the series circuit,  $R_1(R_2)$ ,  $L_1(L_2)$ ,  $C$  are the resistance, inductance and the capacitance respectively. Let the parameters be  $R_1 = R_2 = 100$ ,  $L_1 = L_2 = 10$ ,  $C = 200$ , where their physical units are standard



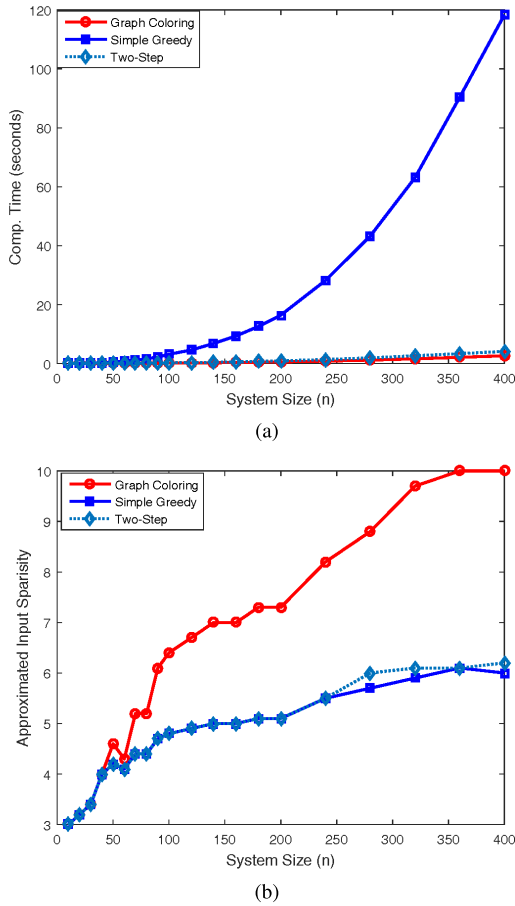


FIGURE 3. Averaged running time (a) and input sparsity (b) when  $l = l_{\max} + 1 = 4$ .

units and here omitted. The goal is to select the sparsest input in order to ensure that all states in the circuit network are of the R-controllability and C-controllability. For the pair  $(E, A)$ , there are three generalized eigenvalues  $s_1 = 0.0, s_2 = 10.0, s_3 = -5.0 \times 10^{-5}$ , and one collection of left eigenvectors is given by  $\{X_i\}_{i=1}^3$ , where  $X_1 = [-10.0, 0.0, 10.0, 0.0, 0.0, 0.0, -1.0, 1.0]$ ,  $X_2 = [100.0, -1.0, 100.0, 1000.0, 1000.0, -10.0, -10.0, 2.5 \times 10^{-5}]$  and  $X_3 = [-2.5 \times 10^{-4}, -1.0, -2.5 \times 10^{-4}, 0.0, 0.0, 0.0, 2.5 \times 10^{-5}, 2.5 \times 10^{-5}]$ . Implementing Algorithm 2, we obtain the input matrices as following:

$$B_1 = [1 \ 0 \ 0 \ 0 \ 0 \ 0 \ 0 \ 0]^T$$

for R-controllability, and

$$B_2 = \begin{bmatrix} 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{bmatrix}$$

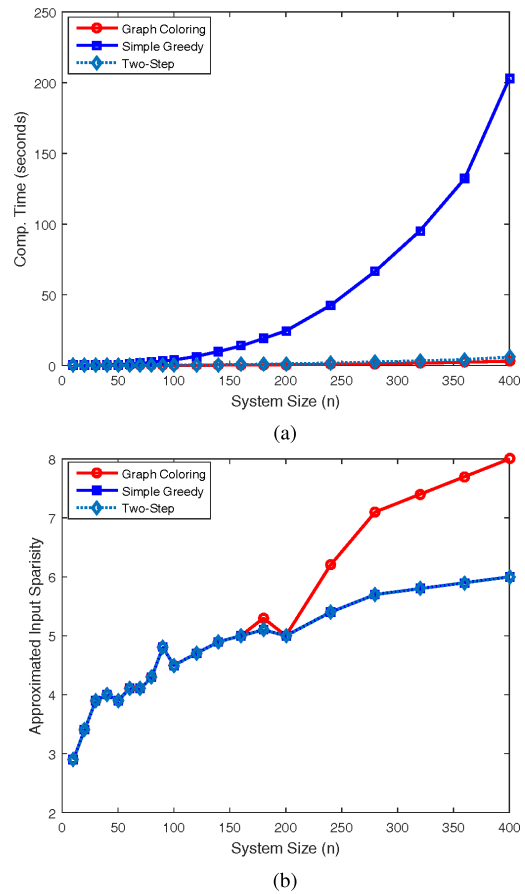


FIGURE 4. Averaged running time (a) and input sparsity (b) when  $l = l_{\max} + 2 = 5$ .

for C-controllability. Substitute the above  $B_1$  and  $B_2$  into Lemma 1 and Lemma 2 respectively, by Lemma 3, it can be found that  $\text{rank}(s_j E - A, B_i) = 8$  for  $j = 1, 2, 3, i = 1, 2$ , and  $\text{rank}[E, B_2] = 8$ . This means that, the associated system is R-controllable under the input matrix  $B_1$  (corresponding to putting a voltage source on the resistor  $R_1$ ), and C-controllable under the input matrix  $B_2$  (corresponding to putting a voltage source on the inductance  $L_2$  and four current sources to control the currents  $i_1, i_2, i_3$  and  $i_4$ ).

Next, we present some numerical results to show the efficiency and effectiveness of the proposed algorithms in this paper. To compare our two-step greedy algorithm with algorithms in the existing literature, consider the problem of constructing the sparsest input matrices for controllability of nonsingular systems, i.e.,  $E = I$ . We generate system state transition matrices using the same method as in [9], which can generate square matrices with multiple eigenvalues.

More precisely, the maximum geometric multiplicity of eigenvalues is set to be 3, which means  $l_{\max} = 3$ . For each dimension of states  $n$ , ranging from 10 to 400, 10 independent plants are generated. For each plant, set the number of inputs  $l$  to be 3, 4 and 5 respectively, and use the graph coloring based algorithm proposed in [9], Algorithm 2 and the simple greedy algorithm (based on  $g(\bar{B})$ ), respectively, to determine

the zero-nonzero structure of input matrices that ensure controllability. The averaged computation time, and the averaged sparsity of input matrices are given in Figs. 2, 3 and 4 respectively for  $l = 3, 4$  and  $5$ . All computations are implemented with a personal notebook computer with Intel(R) Core(TM) i7-5500U CUP and 12.0 GB RAM. The corresponding codes for numerical experiments could be found in the Github [28].

From Figs. 2, 3 and 4, the two-step greedy algorithm achieves comparable approximation performances to the simple greedy algorithms, while costs much less time than the latter algorithm. The graph coloring based algorithm in [9] has almost the same time consumption as that of the two-step greedy algorithm. However, when the system dimension increases or the number of independent inputs  $l$  gets closer to  $l_{\max}$ , its approximation performances seem not as good as those of the two-step greedy algorithm. It can be declared that, the two-step greedy algorithm may be preferable, as it achieves better trade-off between the approximation performances and time complexity.

## VIII. CONCLUSION

The problem of determining the sparsest input matrices with given number of inputs to ensure controllability of singular systems is considered. It is proven that determining the sparsest input matrices to ensure R-controllability and C-controllability is NP-hard, even when the system 'singularity' is arbitrarily large. Because of lack of submodularity and a nonzero submodularity-ratio, it is hard to give a non-trivial approximation bound of the simple greedy algorithm for this problem. Therefore, a two-step greedy algorithm is proposed. This algorithm has much less search space than the simple greedy one, and remarkably, has a submodularity-ratio like provable approximation bound. Through numerous numerical simulations, it is shown that this algorithm achieves better trade-off between the approximation performances and computation efficiency compared to the existing methods. As for future work, it is interesting to derive some non-trivial approximation bounds for the simple greedy algorithm in solving the problems in this paper.

## APPENDIX SOME TECHNICAL PROOFS

*Proof of Lemma 3:* Since the solvability condition holds, there are only a finite number of complex values making  $sE - A$  row rank deficient. According to Lemma 1 of [29], a composite matrix  $[M_1, M_2]$  is of full row rank, if and only if  $TM_2$  is of full row rank, where  $T$  consists of the maximum number of linearly independent row vectors spanning the left null space of  $M_1$ . Applying this lemma on  $[s_iE - A, B]$  and noting that  $X_i$  is a basis matrix of the left null space of  $s_iE - A$ , the required result then follows immediately. ■

*Proof of Proposition 1:* Partition  $P$  as  $P = \text{col}\{P_1, P_2\}$  with  $P_1 \in \mathbb{R}^{n_1 \times n}$ ,  $P_2 \in \mathbb{R}^{n_2 \times n}$ . Then,  $B_1 = P_1B_S$ ,  $B_2 = P_2B_S$ .

Hence,

$$\begin{aligned} \dim R_S(B_S) &= \dim(A_1|B_1) + \dim(E_2|B_2) \\ &= \dim(A_1|P_1B_S) + \dim(E_2|P_2B_S). \end{aligned}$$

According to Theorem 7 of [3],  $\dim \langle A_1|P_1B_S \rangle$  and  $\dim \langle E_2|P_2B_S \rangle$  are both submodular on  $S \subseteq \{1, \dots, l\}$ . Since the summation of two submodular functions is still submodular,  $\dim R_S(B_S)$  is submodular on  $S \subseteq \{1, \dots, l\}$ . ■

*Proof of Theorem 3 and Corollary 1:* From Proposition 2.1.9 of [26], given a constant matrix  $M$  with  $l$  columns,  $\text{rank} M_S$  is submodular on  $S \subseteq \{1, \dots, l\}$ . Notice that the summation of two submodular functions is still submodular. Based on these facts,  $f(S)$  and  $\bar{f}(S)$  are both submodular functions. Noticing that  $f(S)$  and  $\bar{f}(S)$  are both non-decreasing, the performance bounds in Theorem 3 and Corollary 1 then follow from Lemma 4. ■

## REFERENCES

- [1] M. Van De Wal and B. De Jager, "A review of methods for input/output selection," *Automatica*, vol. 37, no. 4, pp. 487–510, Apr. 2001.
- [2] A. Olshevsky, "Minimal controllability problems," *IEEE Trans. Control Netw. Syst.*, vol. 1, no. 3, pp. 249–258, Sep. 2014.
- [3] T. H. Summers, F. L. Cortesi, and J. Lygeros, "On submodularity and controllability in complex dynamical networks," *IEEE Trans. Control Netw. Syst.*, vol. 3, no. 1, pp. 91–101, Mar. 2016.
- [4] F. L. Lewis, "A survey of linear singular systems," *Circuits Syst. Signal Process*, vol. 5, no. 1, pp. 3–36, Mar. 1986.
- [5] J. Ishihara, M. Terra, and J. Campos, "Robust Kalman filter for descriptor systems," *IEEE Trans. Automat. Contr.*, vol. 51, no. 8, p. 1354, Aug. 2006.
- [6] L. Dai, "The difference between regularity and irregularity in singular systems," *Circuits Syst. Signal Process*, vol. 8, no. 4, pp. 435–444, Dec. 1989.
- [7] J.-E. Feng, J. Yao, and P. Cui, "Singular Boolean networks: Semi-tensor product approach," *Sci. Chin. Inf. Sci.*, vol. 56, no. 11, pp. 1–14, Nov. 2013.
- [8] M. Meng and J.-E. Feng, "Optimal control problem of singular Boolean control networks," *Int. J. Control Autom. Syst.*, vol. 13, no. 2, pp. 266–273, Apr. 2015.
- [9] Y. Zhang and T. Zhou, "Input matrix construction and approximation using a graphic approach," *Int. J. Control*, pp. 1–14, Sep. 2018.
- [10] G. Zhuang, J. Xia, J.-E. Feng, B. Zhang, J. Lu, and Z. Wang, "Admissibility analysis and stabilization for neutral descriptor hybrid systems with time-varying delays," *Nonlinear Anal., Hybrid Syst.*, vol. 33, pp. 311–321, Aug. 2019.
- [11] M. Meng and J.-E. Feng, "Topological structure and the disturbance decoupling problem of singular Boolean networks," *IET Control Theory Appl.*, vol. 8, no. 13, pp. 1247–1255, Sep. 2014.
- [12] Y. Y. Liu, J. J. Slotine, and A. L. Barabasi, "Controllability of complex networks," *Nature*, vol. 48, no. 7346, pp. 167–173, 2011.
- [13] G. Yan, G. Tsekenis, B. Barzel, J.-J. Slotine, Y.-Y. Liu, and A.-L. Barabási, "Spectrum of controlling and observing complex networks," *Nature Phys.*, vol. 11, no. 9, pp. 779–786, Sep. 2015.
- [14] Y. Zhang and T. Zhou, "Structural controllability of an NDS with LFT parameterized subsystems," *IEEE Trans. Automat. Contr.*, vol. 64, no. 12, pp. 4920–4935, Dec. 2019.
- [15] E. D. Sontag, *Mathematical Control Theory: Deterministic Finite Dimensional Systems*, vol. 6. New York, NY, USA: Springer, 2013.
- [16] S. Pequito, S. Kar, and A. P. Aguiar, "A framework for structural input/output and control configuration selection in large-scale systems," *IEEE Trans. Automat. Contr.*, vol. 61, no. 2, pp. 303–318, Feb. 2016.
- [17] F. Pasqualetti, S. Zampieri, and F. Bullo, "Controllability metrics, limitations and algorithms for complex networks," *IEEE Trans. Control Netw. Syst.*, vol. 1, no. 1, pp. 40–52, Mar. 2014.
- [18] V. Tzoumas, M. A. Rahimian, G. J. Pappas, and A. Jadbabaie, "Minimal actuator placement with bounds on control effort," *IEEE Trans. Control Netw. Syst.*, vol. 3, no. 1, pp. 67–78, Mar. 2016.

- [19] E. Yip and R. Sincovec, "Solvability, controllability, and observability of continuous descriptor systems," *IEEE Trans. Automat. Contr.*, vol. 26, no. 3, pp. 702–707, Jun. 1981.
- [20] K. J. Reinschke and G. Wiedemann, "Digraph characterization of structural controllability for linear descriptor systems," *Linear Algebra Appl.*, vol. 266, pp. 199–217, Nov. 1997.
- [21] K. Murota, "Refined study on structural controllability of descriptor systems by means of matroids," *SIAM J. Control Optim.*, vol. 25, no. 4, pp. 967–989, Jul. 1987.
- [22] A. Clark, B. Alomair, L. Bushnell, and R. Poovendran, "Input selection for performance and controllability of structured linear descriptor systems," *SIAM J. Control Optim.*, vol. 55, no. 1, pp. 457–485, Jan. 2017.
- [23] V. Kumar Mishra and N. Kumar Tomar, "Alternate checking criteria for reachable controllability of rectangular descriptor systems," *Kybernetika*, vol. 53, pp. 820–837, Dec. 2017.
- [24] G. Boutry, M. Elad, G. H. Golub, and P. Milanfar, "The generalized eigenvalue problem for nonsquare pencils using a minimal perturbation approach," *SIAM J. Matrix Anal. Appl.*, vol. 27, no. 2, pp. 582–601, Jan. 2005.
- [25] L. A. Wolsey, "An analysis of the greedy algorithm for the submodular set covering problem," *Combinatorica*, vol. 2, no. 4, pp. 385–393, Dec. 1982.
- [26] K. Murota, *Matrices and Matroids for Systems Analysis*. Heidelberg, Germany: Springer, 2009.
- [27] A. Das and D. Kempe, "Submodular meets spectral: Greedy algorithms for subset selection, sparse approximation and dictionary selection," in *Proc. 28th Int. Conf. Int. Conf. Mach. Learn.* Madison, WI, USA: Omnipress, 2011, pp. 1057–1064.
- [28] W. H. Zhang. Codes of a Two-Step Algorithm for Sparsest Input Selection for Singular Systems. Github. Accessed: 2019. [Online]. Available: <https://github.com/WanhonZhang/ZWH-Sparest-input-selection>
- [29] Y. Zhang and T. Zhou, "Controllability analysis for a networked dynamic system with autonomous subsystems," *IEEE Trans. Automat. Control*, vol. 62, no. 7, pp. 3408–3415, Jul. 2017.



**YAN ZHANG** was born in Tanghe, Henan, China. She received the B.S. degree in mechanical engineering from the Heilongjiang University of Science and Technology, Harbin, China, in 2006, and the M.S. degree in mechanical engineering from the Lanzhou University of Technology, Lanzhou, in 2009.

From 2009 to 2012, she was a Research Assistant with the Department of Chemical Machinery, Qinghai University, China. Her research interests include the intelligent manufacturing and mechanical electrical integration systems design.



**WANHONG ZHANG** was born in Huining, Gansu, China. He received the B.S. and M.S. degrees in control theory and control engineering from the Lanzhou University of Technology, Lanzhou, in 2007, and the Ph.D. degree in control science and engineering from Tsinghua University, Beijing, China, in 2016.

From 2007 to 2011, he was a Research Assistant with the Department of Chemical Machinery, Qinghai University. Since 2018, he has been an Assistant Professor with the Department of Chemical Machinery, Qinghai University. He is the author of more than 15 articles. His research interests include control theory and applications, systems identification, complex systems, and artificial intelligence.

Dr. Zhang was a recipient of the Qinghai province government paper Award for Excellence, in 2018, and the Small Island Award, in 2018.

• • •