Minimal controllability problem on linear structural descriptor systems with forbidden nodes

Shun Terasaki and Kazuhiro Sato

***Abstract*—We consider a minimal controllability problem (MCP), which determines the minimum number of input nodes for a descriptor system to be structurally controllable. We investigate the "forbidden nodes" in descriptor systems, denoting nodes that are unable to establish connections with input components. The three main results of this work are as follows. First, we show a solvability condition for the MCP with forbidden nodes using graph theory such as a bipartite graph and its Dulmage–Mendelsohn decomposition. Next, we derive the optimal value of the MCP with forbidden nodes. The optimal value is determined by an optimal solution for constrained maximum matching, and this result includes that of the standard MCP in the previous work. Finally, we provide an efficient algorithm for solving the MCP with forbidden nodes based on an alternating path algorithm.**

***Index Terms*—structural controllability, large-scale system, descriptor system, bipartite graph, DM decomposition**

I. Introduction

Controllability analysis for large-scale network systems, such as multi-agent systems [1], [2], brain networks [3], [4], and power networks [5], [6] has received a great deal of interest in recent years, because it can be used to find important nodes [7]. Controllability analysis problems include:

- quantitative problems; the maximization problems of controllability metrics [8]–[12].
- qualitative problems; selecting input problems that render the system controllable [7], [13]–[16].

The quantitative problems require the system parameters, which are not precisely determined in practical systems. In addition, quantitative problems often become computationally intractable when the state dimension becomes large. Conversely, the structural information of a system, *i.e.*, the nonzero patterns of system parameters, is usually known. This is an advantage for qualitative problems that deal only with nonzero patterns. Also, it is known that the structural controllability [17] of a structural system can be checked efficiently using graph algorithms. Thus, for large-scale network systems, it is more appropriate to consider qualitative problems.

Therefore, we consider a Minimal Controllability Problem (MCP) of the following structural descriptor network system with state $x(t) \in \mathbb{R}^n$ and input $u(t) \in \mathbb{R}^m$, where $\mathbb{R}$ is the set of real numbers:

$$F\dot{x}(t) = Ax(t) + Bu(t), \tag{1}$$

where $F, A \in \mathbb{R}^{n\times n}, B \in \mathbb{R}^{n\times m}$, and $F \in \mathbb{R}^{n\times n}$ can be a singular matrix. That is, we assume that although the specific elements of F, A, and B remain unknown, their nonzero patterns are known. The descriptor formulation aptly models practical systems featuring algebraic constraints, such as those found in electric circuit systems [18]–[20]. MCPs can be divided into two main problems [21]:

- MCP0: A problem that finds an $n \times m$ matrix B for system (1) to be structurally controllable where m is minimum. For $F = I_n$ or $F \neq I_n$, efficient algorithms exist for solving MCP0 [7], [13], [14], [16].
- MCP1: A problem that finds an $n \times n$ diagonal matrix B for system (1) to be structurally controllable and the number of nonzero elements in B is minimum [14]. Although a polynomial time algorithm exists [15] for a special case of (1) with $F \neq I_n$, MCP1 is known to be NP-hard [16] in general.

It should be noted that as mentioned in Section 3.2 in [22], there are only a few papers on MCP0 or MCP1 for system (1) with $F \neq I_n$ although just structural controllability analysis under the assumption of a given (F, A, B) has been studied in [19], [23], [24].

The previous work in [16] on structural descriptor system (1) has not considered constraints on the input destination. There is a gap between practical situations since most physical systems have state variables to which inputs cannot be directly connected. For instance, consider a system in which the position $x(t)$ and velocity $v(t)$ of an object are the state variables. In this case, the state equation involves $\dot{x}(t) = v(t)$, but it does not make practical sense to add an input to this equation. Thus, specifying forbidden targets that cannot be connected to inputs is an important practical constraint. Therefore, [13] introduced forbidden nodes to MCP1 for system (1) with $F = I_n$. However, no work applies MCPs with $F \neq I_n$.

In this paper, we address MCP0 with forbidden nodes for structural descriptor system (1), because, in general, MCP1 for system (1) with $F \neq I_n$ is NP-hard, as shown in [16]. Here, the forbidden nodes correspond to the indices of "equations", while those in [13], which studied (1) with $F = I_n$, correspond to the indices of "variables". This naturally generalizes to $F \neq I_n$. In fact,

- for $F = I_n$, the time evolution of x_i is characterized by the i-th equation of (1). Thus, in this case, we can regard the index of equations as that of variables.
- for $F \neq I_n$, the time evolution of x_i is not characterized by the i-th equation of (1). Thus, in contrast to the case of $F = I_n$, here we cannot regard the index of equations as that of variables.

The contributions of this study can be summarized as follows.

S. Terasaki and K. Sato are with the Department of Mathematical Informatics, Graduate School of Information Science and Technology, The University of Tokyo, Tokyo 113-8656, Japan, email:terasaki@g.ecc.u-tokyo.ac.jp (S. Terasaki), kazuhiro@mist.i.u-tokyo.ac.jp (K. Sato)

- We show a necessary and sufficient condition for the existence of the optimal solution of MCP0 with forbidden equations for structural descriptor system (1), which is described in the language of graph theory. This result is also useful in constructing the optimal solution.
- We provide the optimal value of MCP0 with forbidden equations for structural descriptor system (1), by employing the graph-theoretic properties of the system, such as a bipartite graph and its Dulmage–Mendelsohn (DM) decomposition. The optimal value shows that the minimum number of input nodes is determined by a variant of the maximum matching problem with constraints on the matched nodes. This result includes that of the MCP0 without forbidden equations for descriptor system (1) [16].
- We also provide an efficient algorithm for solving the above special matching problem by using the alternating path algorithm. The time complexity of this algorithm is $O(|V|+|E|\sqrt{|V|})$, which is on par with the algorithm for the MCP0 [7], [16] without forbidden equations, where $|V|$ and $|E|$ are the numbers of nodes and edges of the bipartite graph corresponding to descriptor system (1), respectively.

The remainder of this paper is organized as follows. The basic concepts of graph theory are summarized in Section II. The formulation of MCP0 with forbidden equations for structural descriptor systems is described in Section III. In Section IV, we provide the analysis and the algorithm of MCP0 with forbidden equations for the descriptor system (1). The conclusions are presented in Section V.

II. Basic concepts of graph theory

In this section, we present a comprehensive overview of the fundamental concepts of graph theory that are used in this paper.

A strongly connected component (SCC) of a directed graph with node set V is a maximal subset $C \subseteq V$ whose nodes $u, v \in C$ can be connected by a directed path on the graph.

Let $G = (V^+, V^-; E)$ be a bipartite graph. For an edge $e = (v^+, v^-) \in E$, $\partial^+ e$ and $\partial^- e$ denote the nodes of $v^+ \in V^+$ and $v- \in V^-$, respectively. That is, $\partial^+ : E \to V^+$ and $\partial^- : E \to V^-$. The edge set $M \subseteq E$ is a matching if it does not share nodes of each edge. A matching M is termed maximum matching if M contains the largest possible number of edges. We define symbol $\nu(G)$ as the size of a maximum matching of G.

We introduce the DM decomposition for a bipartite graph $G = (V^+, V^-; E)$, which is the unique decomposition algorithm for bipartite graphs (see [19] for details). Algorithm 1 describes DM decomposition. We define M as a maximum matching of G and an auxiliary directed graph $\tilde{G}_M$. The edges of $\tilde{G}_M$ are oriented from V^+ to V^- except for M. The decomposition is illustrated in Fig. 1. Subgraphs G_k $(k = 0, \ldots, b, \infty)$ in Step 6 of Algorithm 1 are called DM components of G. G_k $(k = 1, \ldots, b)$ are called consistent DM components; G_0 and G_∞ are called inconsistent DM components. The order $G_i \preceq G_j$ for the consistent DM components G_i and G_j is defined as

Algorithm 1 Algorithm for DM decomposition.

1: Find a maximum matching M on G.
2: Construct an auxiliary directed graph $\tilde{G}_M$.
3: $V_0 := \{v \ \in V^+ \cup V^- \mid \exists u \in V^+ \setminus \partial^+ M \ u \to_{\tilde{G}_M} v\}$, where $u \to_{\tilde{G}_M} v$ indicates the existence of a directed path on $\tilde{G}_M$ from u to v.
4: $V_\infty := \{v \ \in V^+ \cup V^- \mid \exists u \in V^- \setminus \partial^- M \ v \to_{\tilde{G}_M} u\}$.
5: Let G' be the subgraph of $\tilde{G}_M$ defined by deleting all nodes of $V_0 \cup V_\infty$ and all edges adjacent to the nodes.
6: Let V_k $(k = 1, \ldots, b)$ be the SCCs of G'. Let undirected graph G_k $(k = 0, 1, \ldots, b, \infty)$ be the subgraph of G induced on V_k $(k = 0, 1, \ldots, b, \infty)$.

"There is a directed path on $\tilde{G}_M$ from G_j to G_i,"

and the order between the consistent DM component G_i and inconsistent DM components G_0 and G_∞ are defined as $G_0 \preceq G_i$ and $G_i \preceq G_\infty$, respectively. Then, $\preceq$ is a partial order. Moreover, the decomposition constructed by Algorithm 1 does not depend on an initially chosen maximum matching M of G in step 1), as shown in Lemma 2.3.35 in [19]. For example, in Fig. 1, there is a directed edge from G_2 to G_1, and thus

$$G_0 \preceq G_1 \preceq G_2 \preceq G_\infty.$$

The greatest computational bottleneck in the construction of the DM decomposition is to find the maximum matching of G. This can be achieved in $O(|E|\sqrt{|V|})$ by using the augmentation path algorithm [25]. Thus, the computational complexity of DM decomposition is $O(|E|\sqrt{|V|})$.

III. Problem Settings

In this section, we formulate MCP0 with forbidden equations for the descriptor system (1).

First, we assume that system (1) is solvable, *i.e.*, for any initial state $x(0)$ with an admissible input $u(t)$, there exists a unique solution $x(t)$ to Eq. (1). This condition is equivalent to

$$\mathrm{rank}(A - sF) = n, \tag{2}$$

where s is an indeterminant [19] [26].

In this paper, we call system (1) controllable if for any admissible initial state $x(0)$ that satisfies equation (1), there exists an input $u(t)$ and a final time $T \geq 0$ such that $x(T) = 0$. This controllability is usually referred to as R-controllability. Additional controllability concepts can be found within the framework of descriptor system theory [27].

An algebraic characterization of controllability can be described as follows [26], [27]:

Proposition 1: Descriptor system (1) is controllable *if and only if*

$$\mathrm{rank}\,[A - zF \mid B] = n, \tag{3}$$

where z is any complex number.

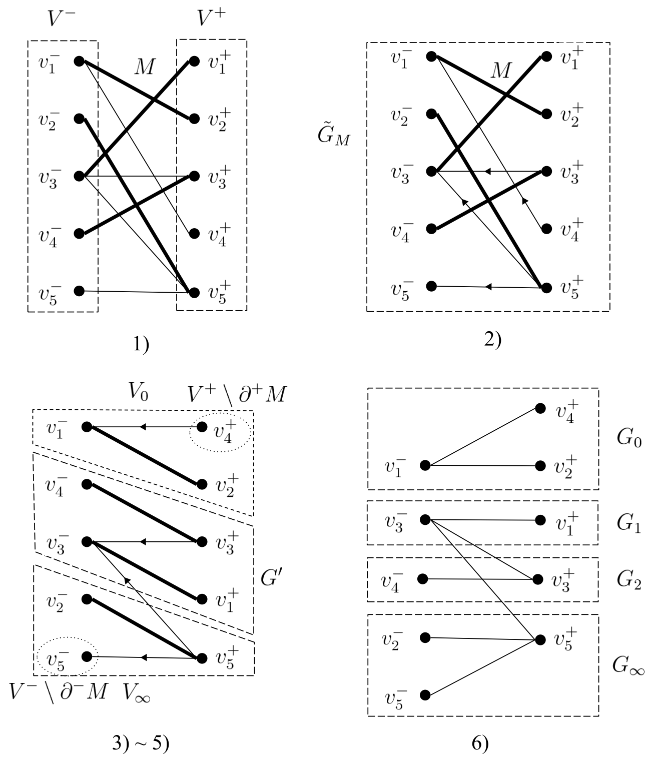


Fig. 1: Construction of DM decomposition. The bold edges represent maximum matching M.

System (1) is termed structurally controllable if condition (3) in Proposition 1 holds for (1) with generic matrices F, A, and B. It should be noted that a matrix is considered generic if each nonzero element is an independent parameter. For a more precise definition of the generic matrix, see [19].

We now introduce MCP0 with forbidden equations for descriptor system (1). Let $R = \{e_1, \dots, e_n\}$ be a set of equation indices and $\mathcal{F}$ be a subset of R that denotes forbidden equations. Then, MCP0 with forbidden equations for descriptor system (1) can be formalized as

$$\begin{cases} \underset{B \in \mathcal{G}^{n \times m}}{\text{minimize}} \quad m \\ \text{subject to} \quad \text{I) system (1) is structurally controllable,} \\ \qquad\qquad\quad \text{II) indices of nonzero rows of } B \text{ are in } R \setminus \mathcal{F}, \end{cases} \tag{4}$$

where $\mathcal{G}^{n \times m}$ denote the set of all $n \times m$ generic matrices. The significant difference between the standard MCP0 and Problem (4) is II) in (4). This constraint limits the input destination.

For instance, consider descriptor system (1) with

$$F = \begin{bmatrix} 0 & f_1 & 0 & 0 & f_2 \\ f_3 & 0 & 0 & 0 & 0 \\ 0 & f_4 & 0 & 0 & 0 \\ 0 & 0 & 0 & f_5 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix}, \ A = \begin{bmatrix} a_1 & 0 & 0 & 0 & 0 \\ a_2 & 0 & 0 & 0 & 0 \\ a_3 & 0 & 0 & 0 & 0 \\ 0 & 0 & a_4 & 0 & a_5 \\ 0 & 0 & a_6 & a_7 & 0 \end{bmatrix}. \tag{5}$$

Then, $R = \{e_1, e_2, \dots, e_5\}$. Let $\mathcal{F} = \{e_3, e_4\}$ be the set of indices of forbidden equations. The matrices

$$B_1 = \begin{bmatrix} b_1 & 0 & 0 & 0 & 0 \end{bmatrix}^\top, \ B_2 = \begin{bmatrix} 0 & 0 & 0 & 0 & b_2 \\ 0 & b_1 & 0 & 0 & 0 \end{bmatrix}^\top \tag{6}$$

satisfy condition II).

IV. Analysis and Algorithm

In this section, we provide the analysis and the algorithm of MCP0 with forbidden equations for descriptor system (1) with generic matrices using a graph-theoretic approach.

To this end, we first describe the graph representation of the descriptor system (1) with generic matrices and graph-theoretical controllability. Although there are several graph representations for descriptor system (1) [19], [24], [28], we used the bipartite graph representation [19] in this study. While the directed graph representation is common for system (1) with $F = I_n$, the bipartite graph representation is often used for $F \neq I_n$ [23].

The bipartite graph $G = (V^+, V^-; E)$ associated with descriptor system (1) is defined as follows: the node sets V^+ and V^- are defined as

$$\begin{cases} V^+ := X \cup U, \\ V^- := \{e_1, \dots, e_n\}, \end{cases}$$

where the state node set X and the input node set U are defined as

$$X := \{x_1, \dots, x_n\}, \quad U := \{u_1, \dots, u_m\},$$

respectively, and e_i in V^- corresponds to the i-th equation of system (1). That is, V^+ consists of state variables and inputs, and V^- is the set of indices of equations. Then, the edge set E is defined as

$$E := E_A \cup E_F \cup E_B, \tag{7}$$

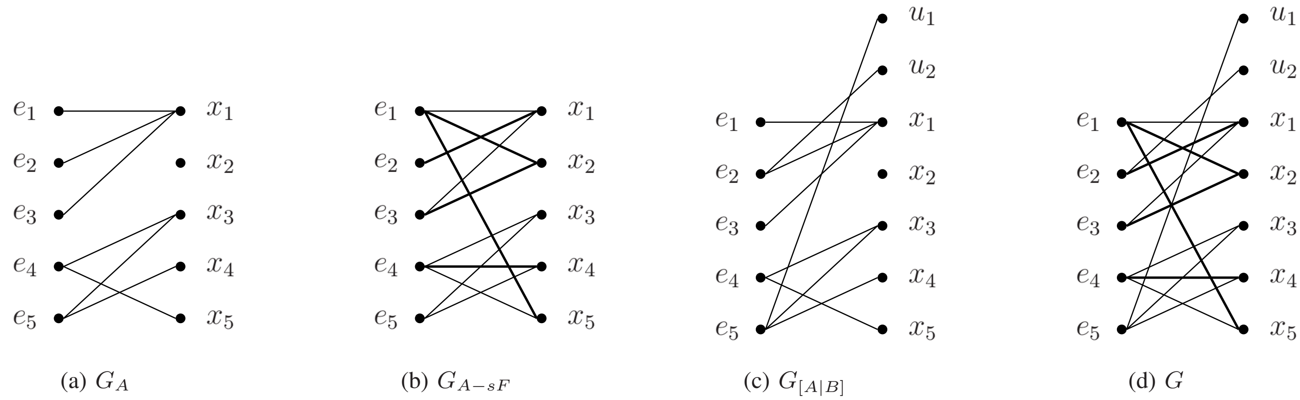
with $E_A := \{(e_i, x_j) \mid A_{ij} \neq 0\}$, $E_F := \{(e_i, x_j) \mid F_{ij} \neq 0\}$ and $E_B = \{(e_i, u_j) \mid B_{ij} \neq 0\}$. An edge belonging to E_F is termed an s-arc. We also define important subgraphs of G as $G_{A-sF} = (X, V^-; E_A \cup E_F)$, $G_{[A|B]} = (V^+, V^-; E_A \cup E_B)$, and $G_A = (X, V^-; E_A)$. For example, the bipartite representation of descriptor system (1) with (5) and B_2 in (6), and its subgraphs are illustrated in Fig. 2.

Furthermore, for a DM component of $\mathcal{G}$ that has s-arcs [16], we call it a DM s-component, where $\mathcal{G}$ is G or G_{A-sF}. Let G_k $(k = 0, \dots, b, \infty)$ be DM components of $\mathcal{G}$, and G_k $(k = 1, \dots, b)$ are called consistent DM components; G_0 and G_∞ are termed inconsistent DM components. The maximal consistent DM s-component is a consistent DM s-component G_k of $\mathcal{G}$ such that no other DM s-components are greater than G_k related to the partial order $\preceq$. Note that there can be multiple maximal consistent DM s-components.

For instance, consider descriptor system (1) with parameters given in (5). The corresponding DM decomposition of the bipartite representation G_{A-sF} is depicted in Fig. 3. G_1, G_2, and G_3 are the consistent DM components of G_{A-sF}, and $G_3 \preceq G_2 \preceq G_1$. That is, a maximal consistent DM s-component of this graph is G_1, because there is no edge to enter G_1 from other DM components.

The following graph-theoretic characterization of structural controllability for descriptor system (1) is found in [19].

Proposition 2: Descriptor system (1) is structurally controllable *if and only if*



(a) G_A (b) G_{A-sF} (c) $G_{[A|B]}$ (d) G

Fig. 2: Bipartite graph representation of descriptor system (1) with (5) and B_2 in (6). The bold edges represent s-arcs.

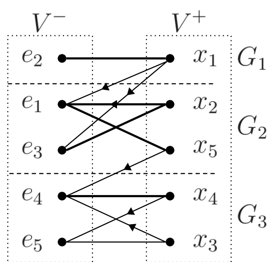


Fig. 3: DM decomposition of G_{A-sF} of descriptor system (1) with (5). The bold edges represent s-arcs.

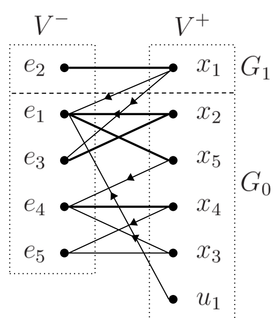


Fig. 4: DM decomposition of G of descriptor system (1) with (5) and B_1 in (6). The bold edges represent s-arcs.

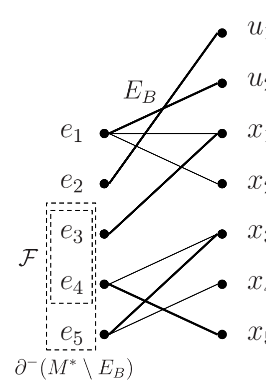


Fig. 5: The optimal solution B in (10) to Problem (4) with $\mathcal{F} = \{e_3, e_4\}$. The bold edges represent the maximum matching M^* of $G_{[A|B]}$. The optimality of B will be shown in Section IV-D.

1) $\nu(G_{A-sF}) = n$,
2) $\nu(G_{[A|B]}) = n$,
3) No consistent DM components of G contain s-arcs,

where $\nu(\mathcal{G})$ is the size of a maximum matching of a bipartite graph $\mathcal{G}$.

Condition 1) represents the solvability of the system described in Eq. (2) in Section III, and the latter two conditions correspond to (3) in Proposition 1. From assumption (2), condition 1) in Proposition 2 is always satisfied. Assumption (2) also implies that there are no inconsistent DM components G_0 and G_∞ of G_{A-sF}. This is because if we choose a maximum matching M with a size of n in G_{A-sF}, then V_0 and V_∞ in Steps 3–4 in Algorithm 1 are both empty.

The following lemma describes the characterization of the consistent DM components of G_{A-sF} and G.

Lemma 1: Consider descriptor system (1) which satisfies solvability condition (2). Then, the following conditions are equivalent:

1) A node $v \in X \cup V^-$ of a consistent DM component G_i in G_{A-sF} belongs to an inconsitent DM component G_0 in G.
2) There is a directed path on $\tilde{G}_M$ from an input node $u \in U$ to $v \in X \cup V^-$ in Step 2 of Algorithm 1 for G.

Proof : From assumption (2), condition 1) in Proposition 2 holds. This means that we can find the same maximum matching M with size of n in Step 1 of Algorithm 1 for G_{A-sF} and G. Then, in Step 2 of Algorithm 1, $\tilde{G}_M$ of G_{A-sF} is a directed subgraph of $\tilde{G}_M$ of G, as illustrated in Fig. 1. Also, $V^+ = X \cup U$, $\partial^+ M = X$, $V^- = \partial^- M$. Thus, $V^+ \setminus \partial^+ M$ and $V^- \setminus \partial^- M$ in Steps 3 and 4 of Algorithm 1 coincide with U and $\emptyset$, respectively. That is, the node sets V_0 and V_∞ of the inconsistent DM components in G are as follows:

$$V_0 = U \cup \{v \in X \cup V^- \mid \exists u \in U \; u \to_{\tilde{G}_M} v\}, \; V_\infty = \emptyset.$$

This implies that 1) and 2) are equivalent. □

For instance, consider descriptor system (1) with (5) and B_1 in (6). The DM decomposition of the bipartite representation G is depicted in Fig. 4. G_0 is the inconsistent DM component, G_1 is the consistent DM component, and $G_0 \preceq G_1$. That is, G_1 is the maximal consistent DM s-component. Then, the consistent DM component G_1 in G_{A-sF} (Fig 3) is also consistent in G

(Fig. 4), since there is no directed path from u_1 to e_2 or x_1 in G. Note that in general, there is a directed path from $u \in U$ to $x \in X$, because edges in a matching are undirected.

Lemma 1 gives another characterization of the structural controllability condition for system (1) which satisfies solvability condition (2). This property will be used in the proof of Theorem 1 in the next subsection.

Corollary 1: Consider descriptor system (1) which satisfies solvability condition (2). Then, descriptor system (1) is structurally controllable *if and only if*

2') $\nu(G_{[A|B]}) = n$,

3') For each maximal consistent DM s-component G_i in G_{A-sF}, which is a subgraph of G, there is a directed path on $\tilde{G}_M$ from U to G_i in Step 2 of Algorithm 1 for G.

Proof : From assumption (2), condition 1) in Proposition 2 holds. Condition 2') is equivalent to condition 2) in Proposition 2. Also, condition 3) in Proposition 2 holds *if and only if* all nodes $v \in X \cup V^-$ of consistent DM s-components G_i in G_{A-sF} belong to an inconsistent DM s-component G_0 in G, since consistent DM s-components contain s-arcs. Using Lemma 1, this is equivalent to the following:

3*) For each consistent DM s-component G_i in G_{A-sF}, which is a subgraph of G, there is a directed path on $\tilde{G}_M$ from U to G_i.

This condition is also equivalent to 3') in Corollary 1, since the non-maximal consistent DM components in G_{A-sF} have a directed path from the maximal consistent DM component in G_{A-sF} from the definition of the partial order $\preceq$. □

Consider descriptor system (1) with A and F given in (5) and B_1 in (6) again. In G_{A-sF}, there is no directed path from u_1 to G_1 (Fig. 4). This means condition 3') in Corollary 1 does not hold. Thus, descriptor system (1) with (5) and B_1 in (6) is not structurally controllable.

A. Existence of solution to Problem (4)

Unlike the traditional MCP0, it is not obvious whether or not an optimal solution exists for MCP0 with forbidden equations. By using Proposition 2 and Corollary 1, we obtain the following theorem which characterizes the existence of solutions to Problem (4). This theorem is employed to construct the optimal solution to Problem (4) in Section IV-D.

Theorem 1: Problem (4) has a solution if and only if both of the following conditions hold simultaneously:

a). For each node set V_i of maximal consistent DM s-components G_i in G_{A-sF}, there exists $e \in V_i \cap V^-$ such that $e \notin \mathcal{F}$.

b). The graph-theoretic problem

$$\begin{cases} \underset{M}{\text{maximize}} & |M| \\ \text{subject to} & M \text{ is matching of } G_A, \\ & \mathcal{F} \subseteq \partial^- M \end{cases} \tag{8}$$

is feasible.

Proof : We assume that $B \in \mathcal{G}^{n \times m}$ is a solution to Problem (4), and prove that conditions a) and b) hold. Note that this B defines E_B in (7). Suppose that a maximal consistent DM s-component G_i of G_{A-sF}, which is a subgraph of G, is not connected to U, where U denotes the set of input nodes. This means that there is no directed path on $\tilde{G}_M$ for G from U to G_i in Step 2 of Algorithm 1. From Corollary 1, this implies that system (1) with the given B is not structurally controllable, which contradicts the assumption. Thus, in G_{A-sF}, for the all node set V_i of maximal consistent DM s-components G_i, there exists an equation node $e \in V^- \cap V_i$ which is connected to an input $u \in U$. This means that $e \notin \mathcal{F}$ and condition a) holds. Moreover, condition 2) in Proposition 2 implies that the maximum matching M^* of $G_{[A|B]} = (V^+, V^- \cup U; E \cup E_B)$ and $|M^*| = n$ exists. Then, $M^* \setminus E_B$ is a matching in G_A and $\partial^-(M^* \setminus E_B)$ contains $\mathcal{F}$ since $\partial^- E_B \subseteq V^- \setminus \mathcal{F}$ (Fig. 5). Thus, $M^* \setminus E_B$ is a feasible solution to Problem (8), and condition b) holds.

Next, we assume that conditions a) and b) hold, and prove that Problem (4) has a solution. Consider the input nodes U with a size of $n - |M^*|$, where M^* is the optimal solution to Problem (8). Then we can connect each equation node $e \in V^- \setminus \partial^- M^* \subseteq V^- \setminus \mathcal{F}$ to an input node. This means that the edge set E_B in (7) with size $|E_B| = n - |M^*|$ was constructed and forms a part of a matching of $G_{[A|B]}$. That is, $M^* \cup E_B$ is a maximum matching of $G_{[A|B]}$, and thus condition 2') in Corollary 1 holds. Also, for each maximal consistent DM s-component V_i of G_{A-sF}, we can connect an equation node $e \in V_i \setminus \mathcal{F}$ to an input from condition a). That is, all maximal consistent DM s-components of G_{A-sF} have directed paths from the input node set U. This means that condition 3') in Corollary 1 holds. Thus, Problem (4) has a solution. □

Theorem 1 implies that the existence of an optimal solution to MCP0 with forbidden equations for system (1) can be verified using pure graph theory.

For instance, consider descriptor system (1) with (5) depicted in Figs. 2 and 3.

1) Consider the case $\mathcal{F} = \{e_1, e_3\}$. Then, MCP0 with $\mathcal{F}$ is not feasible since condition b) in Theorem 1 does not hold. In fact, the maximal consistent DM s-component G_1 in G_{A-sF} has a node e_2 which does not belong to $\mathcal{F}$. This means that condition a) in Theorem 1 holds. However, since a set of nodes that connect to e_1 or e_3 is $\{x_1\}$ in G_A in Fig. 2-(a), there is no matching M of G_A which satisfies $\mathcal{F} \subseteq \partial^- M$. Thus, Problem (8) is not feasible.
2) Consider the case $\mathcal{F} = \{e_2\}$. Then, MCP0 with $\mathcal{F}$ is not feasible since condition a) does not hold. In fact, we can choose a feasible solution for Problem (8) as $M_{\mathcal{F}} := \{(e_2, x_1)\}$ which satisfies $\partial^- M_{\mathcal{F}} = \{e_2\} \supseteq \mathcal{F}$. Thus, condition b) holds. However, a node of a maximal consistent DM s-component G_1 in G_{A-sF} is only e_2 and $e_2 \in \mathcal{F}$.

B. Optimal value for Problem (4)

Using the solution to Problem (8), we can compute the optimal value for Problem (4).

Theorem 2: Suppose that an optimal solution to Problem (4) exists. If the optimal value of Problem (8) is m^*, then the minimum number of inputs that satisfy the constraints of Problem (4) is

$$n_D = \max\{n - m^*, 1\}. \tag{9}$$

Proof : We assume that B is an optimal solution with n_D column size to Problem (4), and prove that $n_D \geq n - m^*$, where $m^* \leq n$ by the definition. To this end, we assume $n_D < n - m^*$. From condition 2) in Proposition 2, $\nu(G_{[A|B]}) = n$ holds. Let M be a maximum matching of $G_{[A|B]}$. Then, $M' := M \setminus E_B$ is a feasible solution to Problem (8) and $|M'| \geq n - n_D$. Thus, we obtain $|M'| > m^*$. This contradicts the maximality of m^*.

If $n_D = 0$, the corresponding system of the form (1) does not satisfy condition I) of Problem (4). Thus, $n_D \geq \max\{n - m^*, 1\}$.

To prove that the equality holds, that is, (9) holds, we assume that $M^*_{\mathcal{F}}$ is an optimal solution to Problem (8) such that $|M^*_{\mathcal{F}}| = m^*$, and construct a system that satisfies the constraints of Problem (4) as follows.

1) Connect each input node to a node of $V^- \setminus \partial^- M^*_{\mathcal{F}}$. Then, $\nu(G_{[A|B]}) = n$ is satisfied and *i.e.*, condition 2') in Corollary 1 holds.
2) Connect the input nodes to all maximal consistent DM s-components of G_{A-sF}. Then condition 3') in Corollary 1 is satisfied without increasing the number of input nodes.

A system constructed by 1) and 2) satisfies condition I) of Problem (4). Moreover, a system constructed by 1) satisfies condition II), which is equivalent to that $\partial^- E_B \subseteq V^- \setminus \mathcal{F}$. This is because $\mathcal{F} \subseteq \partial^- M^*_{\mathcal{F}}$. Note that in 2), we can choose nodes so that condition II) is satisfied, because condition a) in Theorem 1 implies that all maximal consistent DM s-component has at least one node that is not in $\mathcal{F}$. Therefore, this system of the form (1) satisfies the constraints of Problem (4) and n_D is given by (9). □

If the set $\mathcal{F}$ of forbidden equations is empty, Eq. (9) can be written as $n_D = \max\{n - \nu(G_A), 1\}$. This result is consistent with the result of the standard MCP0 for system (1) [16].

The argument in the proof of Theorem 2 will be used to develop an algorithm for Problem (4) in Section IV-D.

C. Algorithm for Problem (8)

Algorithm 2 describes an algorithm for solving Problem (8) based on an alternating path algorithm [25]. Steps 1–4 check the feasibility of solving Problem (8). This is because if there is a feasible solution M to Problem (8), by restricting the matching nodes in $\partial^- M$ to $\mathcal{F} \cap \partial^- M$, a matching $M_{\mathcal{F}}$ of $G_{\mathcal{F}} = (V^+, \mathcal{F}; E_A)$ is obtained that satisfies $|\mathcal{F}| = |M_{\mathcal{F}}|$ (Fig. 6). The converse is shown to hold by the following alternating path algorithm in Steps 5–9, as shown in the proof of Theorem 3.

The alternating path algorithm that is used in Steps 5–9 is explained as follows: Let M^* be a maximum matching of G_A which does not satisfy the condition $\mathcal{F} \subseteq \partial^- M^*$ in (8). That is, M^* is not a feasible solution to Problem (8). Note that M^*

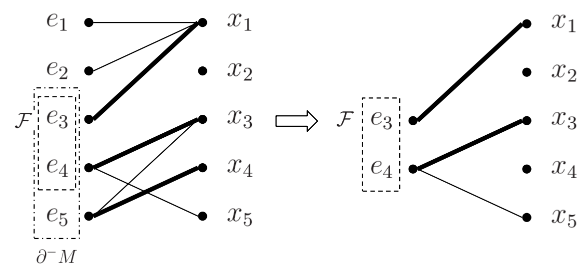


Fig. 6: G_A and $G_{\mathcal{F}} = (V^+, \mathcal{F}; E_A)$. The bold edges on the left are a feasible solution M to Problem (8) while the right is a matching $M_{\mathcal{F}}$ of $G_{\mathcal{F}}$ which satisfies $|\mathcal{F}| = |M_{\mathcal{F}}|$.

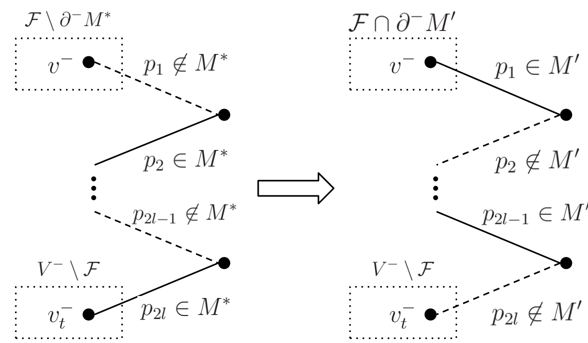


Fig. 7: Alternating path algorithm.

can be computed by the Hopcraft-Karp algorithm [25]. Then, $v^- \in \mathcal{F} \setminus \partial^- M^*$ exists. Suppose a path

$$\begin{aligned} P = p_1 \notin M^* \to p_2 \in M^* \to \cdots \\ \to p_{2l-1} \notin M^* \to p_{2l} \in M^*, \quad l = 1, 2, \ldots, \end{aligned}$$

exists, which starts with v^- and ends with $v_t^- \in V^- \setminus \mathcal{F}$. Then, a new maximum matching M' is defined by the symmetric difference $(M^* \setminus P) \cup (P \setminus M^*)$ between M^* and P. Moreover, the number of $\mathcal{F} \setminus \partial^- M'$ is just one less than the number of $\mathcal{F} \setminus \partial^- M^*$ since $\partial^- M' = (\partial^- M^* \setminus \{v_t^-\}) \cup \{v^-\}$. This procedure is illustrated in Fig. 7. By repeating this process, we can finally have a maximum matching $M^*_{\mathcal{F}}$ of G_A which satisfies $\mathcal{F} \subseteq \partial^- M^*_{\mathcal{F}}$. Note that from the construction of the alternating path, no paths that start at different vertices $v^- \in \mathcal{F} \setminus \partial^- M^*$ share edges with each other. This means that all desired alternating paths can be found by visiting each edge only once. Thus, we can execute Steps 6–9 by the breadth-first search algorithm with time complexity of $O(|E| + |V|)$ [25].

Theorem 3: If Problem (8) is feasible, Algorithm 2 outputs the optimal solution of Problem (8).

Proof : Let M^* be a maximum matching of G_A. From the feasibility assumption of Problem (8), there exists a maximum matching $M^*_{\mathcal{F}}$ of the bipartite graph $(V^+, V^- \cap \mathcal{F}; E_A)$ by Steps 1–4 in Algorithm 2. Because of the construction of Algorithm 2, it is sufficient to show that there is an alternating path P for each $v^- \in \mathcal{F} \setminus \partial^- M^*$. In particular, we show that such P alternates between edges of M^* and $M^*_{\mathcal{F}} \setminus M^*$. From the definition of $M^*_{\mathcal{F}}$, an edge $p \in M^*_{\mathcal{F}}$ exists that connects to v^-. Also, there exists $p' \in M^*$ that $\partial^+ p' = \partial^+ p$. This is because otherwise $M^* \cup \{p\}$ is a matching, whose size is larger

Algorithm 2 Algorithm for solving Problem (8)

1: Find the maximum matching $M^*_{\mathcal{F}}$ of the graph $G_{\mathcal{F}} := (V^+, \mathcal{F}; E_A)$.
2: **if** $|\mathcal{F}| > |M^*_{\mathcal{F}}|$ **then**
3: Problem (8) is infeasible.
4: **end if**
5: Find the maximum matching M^* of G_A.
6: **for** $v^- \in \mathcal{F} \setminus \partial^- M^*$ **do**
7: Find an alternating path $P = \{p_1, \cdots, p_{2l}\}$ which starts with v^- and ends with $v_t^- \in V^- \setminus \mathcal{F}$
8: $M^* := M^* \cup \{p_1, p_3, \cdots, p_{2l-1}\} \setminus \{p_2, p_4, \cdots, p_{2l}\}$
9: **end for**
10: Output M^*.

Algorithm 3 Algorithm for solving Problem (4)

1: **if** condition a) in Theorem 1 is not satisfied **then**
2: Problem (4) is infeasible.
3: **end if**
4: Run Algorithm 2 and let M^* be its output that is an optimal solution to Problem (8).
5: Connect inputs $U := \{u_1, \ldots, u_{n_D}\}$ to each equation node of $V^- \setminus \partial^- M^*$, where n_D is (9) in Theorem 2.
6: Connect inputs U to an arbitrary node of each maximal consistent DM s-components of G_{A-sF}, which does not belong to $\mathcal{F}$.

than M^* and this contradicts the maximality of M^*. Thus, there is a path with alternating the edges of M^* and $M^*_{\mathcal{F}} \setminus M^*$ with a length of at least 2. To show that all alternating paths end with a node v_t^- of V^-, not of V^+, suppose that an alternating path P exists that ends with $\tilde{v}^+ \in V^+$. We also assume that P alternates between the edges of M^* and $M^*_{\mathcal{F}} \setminus M^*$. In this case, we have $|M^*_{\mathcal{F}} \cap P| = |M^* \cap P| + 1$, since P must start with $p_1 \in M^*_{\mathcal{F}}$ and end with $p_{2l-1} \in M^*_{\mathcal{F}}$ (Fig. 8). Thus, by replacing the $M^* \cap P$ edges with $M^*_{\mathcal{F}} \cap P$ edges, we can obtain a matching larger than M^* in size. This contradicts the maximality of M^*. Thus, $p_{2l} \in M^*$ exists and $P \cup \{p_{2l}\}$ is an alternating path. That is, p_{2l} connects $v_t^- \in V^-$. If $v_t^- \in V^- \cap \mathcal{F}$, there is an edge $p' \in M^*_{\mathcal{F}} \setminus M^*$ that satisfies $v_t^- = \partial^- p'$, and $P \cup \{p_{2l}\} \cup \{p'\}$ is a new alternating path between the edges of M^* and $M^*_{\mathcal{F}} \setminus M^*$. Thus, by inductively applying the above argument, we finally have an alternating path P which ends with a node of $V^- \setminus \mathcal{F}$. □

To explain how Algorithm 2 works, consider Problem (8) with G_A of descriptor system (5) and forbidden equations $\mathcal{F} = \{e_3, e_4\}$. We can choose $M^*_{\mathcal{F}}$ in Step 1 of Algorithm 2 as $M^*_{\mathcal{F}} = \{(e_3, x_1), (e_4, x_3)\}$. In Step 5, we have the maximum matching $M^* = \{(e_3, x_1), (e_4, x_3), (e_5, x_4)\}$. Moreover, Step 6–9 illustrated in Fig. 9 produces the new maximum matching $M' = \{(e_3, x_1), (e_4, x_3), (e_5, x_4)\}$ that satisfies $\mathcal{F} \subseteq \partial^- M' = \{e_3, e_4, e_5\}$ in (8). In fact, we can

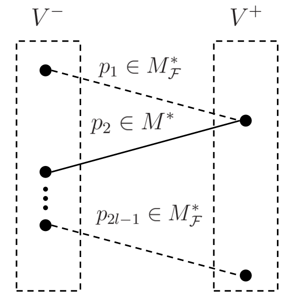


Fig. 8: Alternating path $P = \{p_1, p_2 \ldots, p_{2l-1}\}$. $p_1, p_3, \ldots, p_{2l-1}$ are in $M^*_{\mathcal{F}}$, and $p_2, p_4, \ldots, p_{2l-2}$ are in M^*.

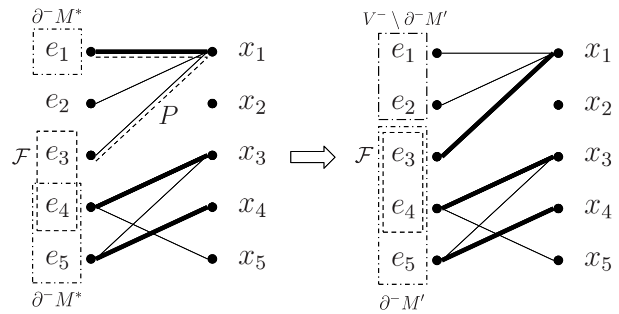


Fig. 9: Example of Step 6–9 in Algorithm 2. The bold edges represent M^* (left) and M' (right).

find an alternating path $P = \{(e_1, x_1), (e_3, x_1)\}$ and obtain the new matching M'.

D. *Algorithm for Problem (4)*

Algorithm 3 shows an algorithm for solving Problem (4) and has the following property.

Theorem 4: If Problem (4) has a feasible solution, then Algorithm 3 outputs the optimal solution to Problem (4). If Problem (4) is infeasible, Algorithm 3 determines the infeasibility. Furthermore, the time complexity is $O(|V| + |E|\sqrt{|V|})$.

Proof: Algorithm 3 is based on the argument in the proof of Theorem 2. Steps 1–4 determine whether or not Problem (4) is feasible, and Steps 1–3 are computed by Algorithm 1. Step 4 checks condition b) in Theorem 1 and output an optimal solution M^* of Problem (8). Steps 5 and 6 in Algorithm 3 output the optimal solution to Problem (4), as shown in the proof of Theorem 2.

Next, we show that the time complexity is $O(|V| + |E|\sqrt{|V|})$. Step 1–3 in Algorithm 3 can be checked by Algorithm 1, its time complexity is $O(|E|\sqrt{|V|})$. The time complexity of Step 4 is $O(|V| + |E|\sqrt{|V|})$. In fact, in Steps 1–5 in Algorithm 2 are computed by using the Hopcroft-Karp algorithm [25], the time complexity is $O(|E|\sqrt{|V|})$. Moreover, Steps 6–9 in Algorithm 2 can be computed by breadth-first search algorithm in $O(|V| + |E|)$, as mentioned already. Steps 5–6 in Algorithm 3 can be computed in $O(|V|)$ at most. Therefore, the time complexity of Algorithm 2 is $O(|V| + |E|\sqrt{|V|})$. □

Theorem 4 means that more general problems than those of [7], [16] can be solved with the same computational complexity. In fact, the problem addressed by [7] is a special case of Problem (4) with $\mathcal{F} = \emptyset$ and $F = I_n$, and it can be solved in $O(|V| + |E|\sqrt{|V|})$. Moreover, the algorithm proposed in [16] deals with Problem (4) with $\mathcal{F} = \emptyset$, and the time complexity is $O(|V| + |E|\sqrt{|V|})$.

To illustrate how Algorithm 3 works, consider descriptor system (1) with (5) and forbidden equations $\mathcal{F} = \{e_3, e_4\}$. In this case, Steps 1–4 determine that Problem (4) is feasible. In fact, the maximal consistent DM s-component G_1 in Fig. 3 has a node e_2 which does not belong to $\mathcal{F}$. Also, in Step 4, we obtain $M^* = \{(e_3, x_1), (e_4, x_3), (e_5, x_4)\}$ from the discussion of how Algorithm 2 works in IV-C. Thus, Problem (4) is feasible, since conditions a) and b) in Theorem 1 hold. Furthermore, Steps 5 and 6 produce

$$B = \begin{bmatrix} b_2 & 0 & 0 & 0 & 0 \\ 0 & b_1 & 0 & 0 & 0 \end{bmatrix}^\top, \tag{10}$$

that is an optimal solution to Problem (4). In fact, from Theorem 2, we have the minimum number of inputs $n_D = n - |M^*| = 2$. Also, $V^- \setminus \partial^- M' = \{e_1, e_2\}$. Thus, connecting u_1 to e_1, and u_2 to e_2 , we have B as (10).

V. Conclusion

In this study, we introduced the forbidden equations to MCP0 for structural descriptor systems. We gave a necessary and sufficient condition for the existence of solutions to the problem and provided the solution to MCP0. The algorithm for solving the problem can be computed in polynomial time as for the standard MCP0. That is, our proposed algorithm is well positioned for applications to large-scale descriptor systems.

In this paper, we focused on MCP0 for a structural descriptor system, since, for a structural descriptor system, MCP1 is NP-hard in general as shown in [16]. Thus, finding a solvable condition for MCP1 with forbidden equations in polynomial time would be a future project.

Furthermore, structural controllability considered in this paper requires that the system parameters be algebraically independent. This means that all non-zero system parameters are free, which may be a strong assumption for practical situations. To avoid this assumption, strong structural controllability has been proposed in [29], and studied from a graph-theoretic perspective in [30]. A strong structural controllability problem version in this paper is one of the future works.

Acknowledgment

This work was supported by the Japan Society for the Promotion of Science KAKENHI under Grant 20K14760 and 23K03899.

References

[1] M. Mesbahi and M. Egerstedt, *Graph theoretic methods in multiagent networks*. Princeton University Press, 2010.

[2] A. Rahmani, M. Ji, M. Mesbahi, and M. Egerstedt, “Controllability of multi-agent systems from a graph-theoretic perspective,” *SIAM Journal on Control and Optimization*, vol. 48, no. 1, pp. 162–186, 2009.

[3] S. Gu, F. Pasqualetti, M. Cieslak, Q. K. Telesford, B. Y. Alfred, A. E. Kahn, J. D. Medaglia, J. M. Vettel, M. B. Miller, S. T. Grafton *et al.*, “Controllability of structural brain networks,” *Nature Communications*, vol. 6, no. 1, pp. 1–10, 2015.

[4] S. F. Muldoon, F. Pasqualetti, S. Gu, M. Cieslak, S. T. Grafton, J. M. Vettel, and D. S. Bassett, “Stimulation-based control of dynamic brain networks,” *PLoS Computational Biology*, vol. 12, no. 9, p. e1005076, 2016.

[5] G. A. Pagani and M. Aiello, “The power grid as a complex network: a survey,” *Physica A: Statistical Mechanics and its Applications*, vol. 392, no. 11, pp. 2688–2700, 2013.

[6] D.-S. Yang, Y.-H. Sun, B.-W. Zhou, X.-T. Gao, and H.-G. Zhang, “Critical nodes identification of complex power systems based on electric cactus structure,” *IEEE Systems Journal*, vol. 14, no. 3, pp. 4477–4488, 2020.

[7] Y.-Y. Liu, J.-J. Slotine, and A.-L. Barabási, “Controllability of complex networks,” *Nature*, vol. 473, no. 7346, pp. 167–173, 2011.

[8] F. Pasqualetti, S. Zampieri, and F. Bullo, “Controllability metrics, limitations and algorithms for complex networks,” *IEEE Transactions on Control of Network Systems*, vol. 1, no. 1, pp. 40–52, 2014.

[9] T. H. Summers, F. L. Cortesi, and J. Lygeros, “On submodularity and controllability in complex dynamical networks,” *IEEE Transactions on Control of Network Systems*, vol. 3, no. 1, pp. 91–101, 2015.

[10] A. Clark, B. Alomair, L. Bushnell, and R. Poovendran, “Submodularity in input node selection for networked linear systems: Efficient algorithms for performance and controllability,” *IEEE Control Systems Magazine*, vol. 37, no. 6, pp. 52–74, 2017.

[11] L. Romao, K. Margellos, and A. Papachristodoulou, “Distributed actuator selection: Achieving optimality via a primal-dual algorithm,” *IEEE Control Systems Letters*, vol. 2, no. 4, pp. 779–784, 2018.

[12] K. Sato and A. Takeda, “Controllability maximization of large-scale systems using projected gradient method,” *IEEE Control Systems Letters*, vol. 4, no. 4, pp. 821–826, 2020.

[13] A. Olshevsky, “Minimal controllability problems,” *IEEE Transactions on Control of Network Systems*, vol. 1, no. 3, pp. 249–258, 2014.

[14] S. Pequito, S. Kar, and A. P. Aguiar, “A framework for structural input/output and control configuration selection in large-scale systems,” *IEEE Transactions on Automatic Control*, vol. 61, no. 2, pp. 303–318, 2015.

[15] A. Clark, B. Alomair, L. Bushnell, and R. Poovendran, “Input selection for performance and controllability of structured linear descriptor systems,” *SIAM Journal on Control and Optimization*, vol. 55, no. 1, pp. 457–485, 2017.

[16] S. Terasaki and K. Sato, “Minimal controllability problems on linear structural descriptor systems,” *IEEE Transactions on Automatic Control*, vol. 67, no. 5, pp. 2522–2528, 2022.

[17] C.-T. Lin, “Structural controllability,” *IEEE Transactions on Automatic Control*, vol. 19, no. 3, pp. 201–208, 1974.

[18] L. Dai, *Singular control systems*. Springer, 1989.

[19] K. Murota, *Matrices and matroids for systems analysis*. Springer Science & Business Media, 2009.

[20] G.-R. Duan, *Analysis and design of descriptor linear systems*. Springer Science & Business Media, 2010.

[21] Y.-Y. Liu and A.-L. Barabási, “Control principles of complex systems,” *Reviews of Modern Physics*, vol. 88, p. 035006, 2016.

[22] G. Ramos, A. P. Aguiar, and S. Pequito, “An overview of structural systems theory,” *Automatica*, vol. 140, p. 110229, 2022.

[23] K. Murota, “Structural controllability of a system in descriptor form expressed in terms of bipartite graphs (in japanese),” *Transactions of the Society of Instrument and Control Engineers*, vol. 20, no. 3, pp. 272–274, 1984.

[24] K. J. Reinschke and G. Wiedemann, “Digraph characterization of structural controllability for linear descriptor systems,” *Linear Algebra and its Applications*, vol. 266, pp. 199–217, 1997.

[25] B. Korte, J. Vygen, B. Korte, and J. Vygen, *Combinatorial optimization*. Springer, 2012, vol. 2.

[26] E. Yip and R. Sincovec, “Solvability, controllability, and observability of continuous descriptor systems,” *IEEE Transactions on Automatic Control*, vol. 26, no. 3, pp. 702–707, 1981.

[27] T. Berger and T. Reis, “Controllability of linear differential-algebraic systems—a survey,” in *Surveys in differential-algebraic equations I*. Springer, 2013, pp. 1–61.

[28] T. Yamada and D. Luenberger, “Generic controllability theorems for descriptor systems,” *IEEE Transactions on Automatic Control*, vol. 30, no. 2, pp. 144–152, 1985.

[29] H. Mayeda and T. Yamada, “Strong structural controllability,” *SIAM Journal on Control and Optimization*, vol. 17, no. 1, pp. 123–138, 1979.

[30] J. Jia, H. J. Van Waarde, H. L. Trentelman, and M. K. Camlibel, “A unifying framework for strong structural controllability,” *IEEE Transactions on Automatic Control*, vol. 66, no. 1, pp. 391–398, 2021.