A Flow-Network Based Polynomial-Time Approximation Algorithm for the Minimum Constrained Input Structural Controllability Problem

Shana Moothedath, Prasanna Chaporkar and Madhu N. Belur

Abstract—This paper deals with minimum cost constrained input selection (minCCIS) for state space structured systems. Our goal is to optimally select an input set from the given inputs such that the system is structurally controllable when the set of states that each input can influence is pre-specified and each input has a cost associated with it. This problem is known to be NP-hard. Firstly, we give a flow-network based novel necessary and sufficient graph theoretic condition for checking structural controllability. Using this condition we propose a polynomial reduction of the problem to a known NP-hard problem: the minimum-cost fixed-flow problem (MCFF). Subsequently, we show that approximation schemes of MCFF directly apply to the minCCIS problem. Using the special structure of the flownetwork constructed for the structured system, we formulate a polynomial time approximation algorithm based on minimum weight bipartite matching problem and a greedy scheme for solving the MCFF problem on the system flow-network. The proposed algorithm gives a so-called Δ -approximate solution to MCFF, where Δ denotes the maximum in-degree of input vertices in the flow-network of the structured system.

Index Terms—Structural controllability, Minimum input structural controllability, Maximum flow problem, Minimum-cost fixed-flow problem, Approximation algorithms.

1. INTRODUCTION

We consider a control system $\dot{x} = Ax + Bu$, where A is the state matrix and B is the input matrix. We assume that the exact entries of A and B are not known, rather only the location of the zero entries is known. Further, each input has a cost associated with it. Our aim in this paper is to choose a subset of inputs that keeps the system controllable while minimizing the cost.

A. Problem Formulation

Let \overline{A} , \overline{B} be matrices of dimensions $n \times n$, $n \times m$ respectively whose entries are either \star or 0. We say that \overline{A} and \overline{B} *structurally represent* state and input matrices of a control system $\dot{x} = Ax + Bu$ where A and B satisfy:

$$A_{ij} = 0 \text{ whenever } A_{ij} = 0, \text{ and}$$

$$B_{ij} = 0 \text{ whenever } \bar{B}_{ij} = 0.$$
(1)

We refer to matrices A and B that satisfy (1) as a *numerical realization* of \overline{A} and \overline{B} respectively and $(\overline{A}, \overline{B})$ as a *structured system*. Thus $(\overline{A}, \overline{B})$ represents a class of control systems corresponding to all possible numerical realizations. The key idea in structural controllability is to determine controllability of the class of control systems represented by $(\overline{A}, \overline{B})$ [1]. We have the following definition.

Definition 1.1. The structured system $(\overline{A}, \overline{B})$ is said to be structurally controllable if there exists at least one numerical realization (A, B) such that (A, B) is controllable.

Structural controllability can be verified in polynomial time [2]. In this paper, we first propose an alternate flownetwork based condition to establish structural controllability. Subsequently, for structurally controllable systems, we develop algorithms based on this flow-network to find optimal solutions to the optimization problem considered below. Let (\bar{A}, \bar{B}) be structurally controllable. Consider $\mathcal{W} \subseteq \{1, \dots, m\}$ and let $\bar{B}_{\mathcal{W}}$ be the restriction of \overline{B} to columns only in \mathcal{W} . Furthermore, let $\mathcal{K} := \{\mathcal{W} : (\bar{A}, \bar{B}_{\mathcal{W}}) \text{ is structurally controllable}\}$. The set \mathcal{K} is non-empty, since for $\mathcal{W} = \{1, \dots, m\}, (\bar{A}, \bar{B}_{\mathcal{W}}) = (\bar{A}, \bar{B})$ is structurally controllable. Given a structurally controllable structured system (\bar{A}, \bar{B}) and non-negative cost vector p_{μ} , where entry $p_u(j)$, j = 1, 2, ..., m, indicates the cost of actuating the jth input, the minimum cost constrained input selection (minCCIS) problem consists of finding a minimum cost input set such that the system is structurally controllable. Specifically, we wish to solve the following optimization: for any $\mathcal{I} \in \mathcal{K}$, define $p(\mathcal{I}) = \sum_{i \in \mathcal{I}} p_u(j)$.

Problem 1.2. Given a structurally controllable structured system $(\overline{A}, \overline{B})$ and $p_u(j), j = 1, 2, ..., m$, find

$$\mathcal{I}^{\star} \in \arg\min_{\mathcal{T} \in \mathcal{K}} p(\mathcal{I}),$$

where $\mathcal{K} := \{\mathcal{W} : (\bar{A}, \bar{B}_{\mathcal{W}}) \text{ is structurally controllable}\}.$

Let $p^* := p(\mathcal{I}^*)$. Thus, p^* denotes the minimum cost for constrained input selection that ensures structural controllability. If costs are non-zero and uniform, then Problem 1.2 is referred as *minimum constrained input selection* (minCIS) problem. In this paper, we formulate minCCIS as an instance of minimum-cost fixed-flow problem (MCFF), where the objective is to minimize the cost associated with the flow.

Note that in recent time structural controllability has attained significant interest on account of its applicability in diverse areas that include biological systems, electronic circuits, transportation, World Wide Web, social communication, power grids and robotics. For example, in a gene regulatory network the aim is to control the dynamics of cellular processes [3]. We refer to [3], [4], [5] and references therein to note the applicability of these concepts to several real networks.

B. Related Contribution

Structural controllability is a well studied area since its introduction in [2]. The key motivation then for studying system properties using their structural pattern instead of their exact numerical realization was that in many practical cases

The authors are in the Department of Electrical Engineering, Indian Institute of Technology Bombay, India. Email: {shana, chaporkar, belur}@ee.iitb.ac.in. This work was supported in part by SERB (DST) and BRNS, India.

the exact values of the parameters in the system description are not accurately known and controllability if lost, is often over a "thin set". Recently this area has regained interest because of its wide range of applications including social networks, biological networks, power grids and robotics. For instance, the problem of identifying the minimum number of inputs required to achieve structural controllability is considered and maximum matching condition is given in [6]. The sparsest input design problem when the input matrix is unconstrained is addressed for the single input case in [7] and for the multiinput case in [8], [9], [10]. For a detailed reading on various research done in this area, see [11] and references therein.

This paper deals with minimum cost input selection for a given structured system when the input matrix and the cost of each input is specified. The NP-hardness of this problem when the costs are non-zero and uniform is proved in [12]. However, for a subclass of systems where the state bipartite graph (see Section 2 for more details) has a perfect matching and the inputs are 'dedicated' (i.e., a diagonal input matrix) with uniform costs, the minCCIS problem is in P and this case is considered in [12]. Similarly, if the state digraph (see Section 2 for more details) is 'irreducible',¹ then the minCCIS problem is no longer NP-hard and this case is considered in [13]. However, minimum constrained input selection problems are not addressed in their full generality. Reducing these problems to the standard NP-hard problems with good approximation schemes was posed as an open problem in [12]. In this paper, we reduce the minCCIS problem to the minimumcost fixed-flow (MCFF) problem and present a polynomial time approximation algorithm for approximating the MCFF problem on the system flow-network. Note that in this work we do not impose any assumption on the structured system and the problem is considered in its full generality.

C. Summary of Contribution and Organization of the Paper

This paper develops an algorithm for finding a minimum (in the sense of cost) input set for structural controllability when the sparsity pattern of the input matrix B is specified.

Our key contributions are as follows:

• We formulate a new graph theoretic necessary and sufficient condition for checking structural controllability using flow-networks (see Theorem 3.3).

We reduce the minimum cost constrained input selection (minCCIS) problem to an MCFF problem in polynomial time.
We prove that an optimal solution to the MCFF problem corresponds to an optimal solution to the minCCIS problem (see Theorem 4.4).

• We prove that approximation schemes for MCFF apply to the minCCIS problem (see Theorem 4.5).

• We provide a polynomial time Δ -approximation algorithm² to solve the minCCIS problem using the MCFF problem (see Theorem 4.8).

The organization of this paper is as follows: in Section 2, we explain structural controllability using concepts from graph theory. In Section 3, we discuss a relation between structural

- ¹A digraph is said to be *irreducible* if there exists a directed path between every pair of nodes.
- $^{2}A \Delta$ -approximation algorithm is an algorithm whose solution value is at most Δ times that of the actual optimum value.



controllability and maximum flow problem. Using this, a new graph theoretic condition for checking structural controllability is also given. In Section 4, formulation of minCCIS as a flow problem is given. Using this formulation we obtain a Δ -approximate solution to the minCCIS problem. Finally, in Section 5 we give the concluding remarks.

Before discussing the minimum cost flow formulation of Problem 1.2, we explain structural controllability using concepts of graph theory in the next section.

2. REVIEW OF ESSENTIAL GRAPH THEORETIC RESULTS FOR STRUCTURAL CONTROLLABILITY

In this section we briefly describe some existing graph theoretic concepts associated with structural controllability for the sake of completeness (see [11] for details).

The key idea behind considering graph for studying structural controllability is because we can represent the influences of states and inputs on each state through a directed graph. We first construct the state digraph $\mathcal{D}(\bar{A})$ with vertex set $V_X = \{x_1, \ldots, x_n\}$ and edge set E_X , where $(x_i, x_i) \in E_X$ if $\bar{A}_{ij} \neq 0$. Presence of an edge (x_j, x_i) in $\mathcal{D}(\bar{A})$ indicates that state x_i of the system *influences* state x_i . To capture the effect of inputs, we construct the system digraph $\mathcal{D}(\bar{A},\bar{B})$ with vertex set $V_X \cup V_U$ and edge set $E_X \cup E_U$. An edge $(u_i, x_i) \in E_U$ if $\overline{B}_{i,i} \neq 0$ and we say that input u_i influences state x_i . Construction of $\mathcal{D}(\bar{A})$ and $\mathcal{D}(\bar{A},\bar{B})$ is illustrated for a structured system shown in Figure 1 in Figures 2 and 3b respectively. In a structured system, a state x_i is said to be *inaccessible* if it is not reachable³ from any input vertex. To check if all states are accessible, we first generate a directed acyclic graph (DAG) associated with $\mathcal{D}(\bar{A})$ by condensing each strongly connected component⁴ (SCC) of the graph to a supernode. Thus in this DAG, vertex set comprises of all SCCs. A directed edge exists between two nodes of DAG if there exists a directed edge connecting vertices in the respective SCCs in the original digraph. Using this DAG, we have a following definition characterizing SCCs of $\mathcal{D}(\bar{A})$.

Definition 2.1. An SCC is said to be <u>non-top linked</u> if it has no incoming edges to its vertices from the vertices of another SCC.

While accessibility of all states is necessary for structural controllability, it is not sufficient. In addition to accessibility, we must also ensure that $\mathcal{D}(\bar{A},\bar{B})$ has *no dilation*⁵. Lin proved the sufficiency of these conditions through the following result.

³A vertex v_k is said to be *reachable* from a vertex v_1 if there exists a sequence of directed edges with all vertices distinct from v_1 to v_k .

⁴A digraph is said to be strongly connected if for each ordered pair of vertices (v_1, v_k) there exists an elementary path from v_1 to v_k . A maximal strongly connected subgraph of a digraph, called a *strongly connected component* (SCC), is a subgraph that is strongly connected and is not properly contained in any other subgraph that is strongly connected.

⁵The digraph $\mathcal{D}(\bar{A},\bar{B})$ is said to have dilation if there exists a set of nodes $S \subset V_X$ whose neighbourhood node set denoted by T(S) (where node $x_i \in T(S)$, if there exists a directed edge from x_i to a node in S) has fewer nodes than S.

Proposition 2.2. [2, pp.207] The structured system (\bar{A}, \bar{B}) is structurally controllable if and only if the associated digraph $\mathcal{D}(\bar{A}, \bar{B})$ has no inaccessible states and has no dilations.

Our main result requires Proposition 2.3 which needs the following background. Given a bipartite graph G_B with vertex set $V \cup \widetilde{V}$ satisfying $V \cap \widetilde{V} = \phi$ and edge set E satisfying $E \subseteq V \times V$, a matching M is a collection of edges $M \subseteq E$ such that no two edges in M share the same vertex. That is, for any (i, j) and $(u, v) \in M$, we have $i \neq u$ and $j \neq v$, where $i, u \in V$ and $j, v \in V$. A perfect matching is a matching M, where |M| = $\min(|V|, |V|)$. Given G_B and a weight function w from the set E to the set of non-negative real numbers \mathbb{R}_+ , a minimum weight perfect matching is a perfect matching M such that $\sum_{e \in M} w(e) \leq \sum_{e \in M'} w(e)$, where M' is any perfect matching of G_B [14]. Corresponding to the state digraph $\mathcal{D}(\bar{A})$, we associate the bipartite graph $\mathcal{B}(\bar{A})$ with vertex set $V_X \cup V_{X'}$ and edge set \mathcal{E}_X , where $V_X = \{x_1, x_2, \dots, x_n\}, V_{X'} = \{x'_1, x'_2, \dots, x'_n\}$ and $(x_i, x'_i) \in \mathcal{E}_X \Leftrightarrow (x_i, x_j) \in E_X$. Similarly, corresponding to the system digraph $\mathcal{D}(\bar{A},\bar{B})$, we associate the bipartite graph $\mathcal{B}(\bar{A},\bar{B})$, with vertex set $(V_X \cup V_U) \cup V_{X'}$ and edge set $\mathcal{E}_X \cup \mathcal{E}_U$. Here, $V_U = \{u_1, u_2, \dots, u_m\}$ and $(u_i, x'_i) \in \mathcal{E}_U \Leftrightarrow (u_i, x_j) \in E_U$. Note that $\mathcal{D}(\bar{A}), \mathcal{D}(\bar{A}, \bar{B})$ are digraphs, but $\mathcal{B}(\bar{A}), \mathcal{B}(\bar{A}, \bar{B})$ are undirected graphs. Now we give the following result.

Proposition 2.3. [10, Theorem 2] Digraph $\mathcal{D}(\bar{A},\bar{B})$ has no dilations if and only if $\mathcal{B}(\bar{A},\bar{B})$ has a perfect matching.

Using the two graph theoretic conditions explained in this section, checking structural controllability of a system has polynomial complexity [14]. However, these conditions do not give ample insight about solving minCCIS problem. In the next section, we give an alternate graph theoretical condition for checking structural controllability using flownetworks. This condition will be subsequently used to provide an approximation algorithm for the minCCIS problem.

3. FLOW-NETWORK REFORMULATION OF STRUCTURAL CONTROLLABILITY

In this section, we establish a relation between the structural controllability and the maximum flow problem [15]. In the context of structured systems, flow-networks are used previously. For instance, a minimum cost network flow problem is used in [16] for finding an elementary i/o matching and a primal-dual algorithm is used in [17] for finding structure at infinity. Given a flow-network \mathcal{F} with vertex set V, directed edge set E, source-sink pair s, t and non-negative capacities b(e) for every $e \in E$, we define a flow vector $f: E \to \mathbb{R}_+$.

Definition 3.1. In a flow-network \mathcal{F} with vertex and edge sets V and E respectively, a source-sink pair s,t and non-negative edge capacities b(e), a flow vector f is said to be <u>feasible</u> if

(i) for every
$$e \in E$$
, $f(e) \leq b(e)$, and

(ii) for every
$$v \in V \setminus \{s,t\}$$
, $\sum_{e=(u,v)\in E} f(e) = \sum_{e'=(v,w)\in E} f(e')$.

The requirements (i) and (ii) in Definition 3.1 are called capacity constraint and flow conservation constraint respectively. Capacity constraint ensures that the flow through each edge is less than the edge capacity. The flow conservation constraint ensures that at every node, except possibly the source and the sink nodes, the flow leaving a node equals the flow entering Algorithm 3.1 Pseudo-code for constructing the flow-network of a structured system (\bar{A}, \bar{B})

Input: Structured matrices $\bar{A} \in \{0, \star\}^{n \times n}$ and $\bar{B} \in \{0, \star\}^{n \times m}$ **Output:** Flow-network $\mathcal{F}(\bar{A}, \bar{B})$, source, sink nodes: *s*,*t* and capacity vector *b*

- 1: Find non-top linked SCCs of $\mathcal{D}(\bar{A}), \mathcal{N} := \{\mathcal{N}_i\}_{i=1}^q$
- 2: Construct flow-network $\mathcal{F}(\bar{A},\bar{B})$ with vertex set V_F and edge set E_F as follows:

3:
$$V_F \leftarrow \left\{ \{s,t\} \cup \{\mathcal{N}_i\}_{i=1}^q \cup \{x'_k, x_k\}_{k=1}^n \cup \{u'_j, u_j\}_{j=1}^m \right\}$$

4: $e \in E_F \leftarrow \begin{cases} (s,\mathcal{N}_i), \text{ for } i \in \{1,2,\ldots,q\}, \\ (s,x'_k), \text{ for } k \in \{1,2,\ldots,n\}, \\ (\mathcal{N}_i,u'_j), \bar{B}_{r,j} = \star \text{ and } x_r \in \mathcal{N}_i, \\ (x'_k, x_r), \bar{A}_{kr} \neq 0, \\ (x'_k, u_j), \bar{B}_{kj} \neq 0, \\ (u'_j, u'_j), \text{ for } j \in \{1,2,\ldots,m\}, \\ (u'_j, t), \text{ for } j \in \{1,2,\ldots,m\}, \\ (x_k, t), \text{ for } k \in \{1,2,\ldots,n\}. \end{cases}$
5: $b(e) \leftarrow \begin{cases} n+1, \text{ for } e = (u'_j, t), j \in \{1,2,\ldots,m\}, \\ 1, \text{ otherwise.} \end{cases}$

the node. We define the flow from the source to the sink under a feasible flow vector f as

$$\varphi_f = \sum_{e=(s,v)\in E} f(e).$$
(2)

The objective of a maximum flow problem is to find a feasible flow vector f^* such that $\varphi_{f^*} \ge \varphi_f$ for any feasible flow vector f. It is a well studied problem and there exist many algorithms that find the maximum flow f^* in time polynomial in the number of nodes and edges of the flow-network. For example, the algorithm in [18] computes f^* in O(|V||E|). The case of zero flow is ruled out from the rest of the paper.

In order to establish a relation between these two problems, i.e., maximum flow and structural controllability, we first construct the flow-network $\mathcal{F}(\bar{A},\bar{B})$ corresponding to the given structured system (\bar{A}, \bar{B}) . The pseudo-code for constructing the flow-network $\mathcal{F}(\bar{A},\bar{B})$ is presented in Algorithm 3.1. Given (\bar{A},\bar{B}) , we first find the digraph $\mathcal{D}(\bar{A})$, the bipartite graph $\mathcal{B}(\bar{A},\bar{B})$ and the non-top linked SCCs in $\mathcal{D}(\bar{A}), \mathcal{N} = \{\mathcal{N}_i\}_{i=1}^q$ (see Step 1). In order to avoid new symbols, with some abuse of notation we denote the condensed version of non-top linked SCCs using the same notation, $\{\mathcal{N}_i\}_{i=1}^q$. Then we define the vertex set V_F (see Step 3), edge set E_F (see Step 4), sourcesink pair s, t and capacity vector b (see Step 5) as shown in the algorithm⁶. The flow-network $\mathcal{F}(\bar{A},\bar{B})$ of a system (\bar{A},\bar{B}) given in Figure 3a is shown in Figure 3b. Note that in $\mathcal{F}(\bar{A},\bar{B})$, Block-1 corresponds to the non-top linked SCCs in $\mathcal{D}(\bar{A})$ and Block-2 is the directed version of $\mathcal{B}(\bar{A},\bar{B})$. The flows entering Block-1 and Block-2 are defined as $\sum_{e \in \{(s,\mathscr{N}_i)\}_{i=1}^q} f(e)$ and $\sum_{e \in \{(s,x'_k)\}_{k=1}^n} f(e)$ respectively. These flows are later used in the sequel to provide a necessary and sufficient condition for

⁶Note that even though V_F and E_F depend on (\bar{A}, \bar{B}) , we are not making the dependence explicit in our notations for brevity. This is because (V_F, E_F) can be obtained unambiguously given the context.



Figure 3: Illustrative example showing system digraph $\mathcal{D}(\bar{A},\bar{B})$ and flow-network $\mathcal{F}(\bar{A},\bar{B})$ for a structured system (\bar{A},\bar{B}) .

structural controllability of (\bar{A}, \bar{B}) . A flow network with all integer capacities satisfies the *integrality theorem*, which is as follows.

Proposition 3.2. [19] If all capacities in a flow-network are integers, then there exists an integer maximum flow solution.

The following result relates structural controllability and maximum flow problems.

Theorem 3.3. Consider a structured system (\bar{A}, \bar{B}) with n states and q non-top linked SCCs in $\mathcal{D}(\bar{A})$. Then, (\bar{A}, \bar{B}) is structurally controllable if and only if the maximum flow $\varphi_{f^{\star}}$ in the flow-network $\mathcal{F}(\bar{A}, \bar{B})$ is equal to q + n.

Proof. Recall the integrality theorem given in Proposition 3.2. Since $b(e) \in \mathbb{Z}_+$ without loss of generality, we assume that the optimal flow vector f^* is an integer valued function from the edge set E_F .

Only-if part: We assume that (\bar{A}, \bar{B}) is structurally controllable and show that the maximum flow is equal to q+n, i.e., $\varphi_{f^*} = q+n$. For ease of understanding, we first outline the key steps: **Step (i)** We use structural controllability to deduce a perfect matching in $\mathcal{B}(\bar{A}, \bar{B})$ and accessibility of all x_i 's in $\mathcal{D}(\bar{A}, \bar{B})$. **Step (ii)** We construct a flow vector f with flow $\varphi_f = q+n$.

Step (iii) We then show f satisfies the capacity constraint.

Step (iv) We finally prove that the flow is conserved at all nodes except source s and sink t.

Step (i): Assume (\bar{A}, \bar{B}) is structurally controllable. Then by Propositions 2.2 and 2.3, all the states are accessible and there exists a perfect matching in $\mathcal{B}(\bar{A}, \bar{B})$. All states being accessible implies that all non-top linked SCCs are connected to some input vertex. Denote by $u(\mathcal{N}_i)$ an input that connects to some state in a non-top linked SCC \mathcal{N}_i . There can be many inputs connecting to a vertex in \mathcal{N}_i , we choose any one of them as $u(\mathcal{N}_i)$. Furthermore, since $\mathcal{B}(\bar{A}, \bar{B})$ has a perfect matching, say M, for every vertex x'_k there exists a unique $y_k \in V_X \cup V_U$ such that $(x'_k, y_k) \in M$. Now, we construct a feasible flow vector f for the flow-network $\mathcal{F}(\bar{A}, \bar{B})$ such that $\varphi_f = q + n$. This would prove the required as $\varphi_{f^*} \ge \varphi_f$. However, since the maximum outflow through source is q + n, $\varphi_{f^*} = \varphi_f$.

Step (ii): Construct a flow vector f in $\mathcal{F}(\bar{A}, \bar{B})$ as follows: a. f((s,v)) = 1 for every $v \in \{\mathscr{N}_1, \ldots, \mathscr{N}_q\} \cup \{x'_1, \ldots, x'_n\}$, b. $f((\mathscr{N}_i, u'(\mathscr{N}_i))) = 1$ for every $i \in \{1, \ldots, q\}$, c. $f((x'_k, y_k)) = 1$, for every $k \in \{1, \ldots, n\}$, d. if $y_i \in V_U$, then $f((y_i, y'_i)) = 1$, else $f(y_i, t) = 1$, and e. $f((u'_j, t)) = |\{i : u_j = u(\mathcal{N}_i)\}| + f((u_j, u'_j))$, where |D| denotes the cardinality of set D.

From Step (ii-a) and equation (2), it follows that $\varphi_f = q + n$. **Step (iii):** Thus, it suffices to show that f is feasible. First, we show that f satisfies the capacity constraint. Note from Steps (ii-a) to (ii-d), each edge except that emanating from nodes u'_j 's has unit flow. Thus, for these edges f(e) = b(e) =1. Now we show that the capacity constraint is satisfied for the edges (u'_j, t) as flow through all other edges is at most one. Edges (u'_j, t) considered in construction Step (ii-e) are shown to be in E_F . By construction of $\mathcal{F}(\bar{A}, \bar{B})$, each of these edges has capacity n+1. Since $q \leq n$ and since u'_j can have unit capacity incoming edges only from u_j and $\mathcal{N}_1, \ldots, \mathcal{N}_q$, the total flow coming in u'_j is bounded above by n+1. This concludes that f satisfies the capacity constraint.

Step (iv): To see that *f* satisfies the flow conservation constraint, note that the flow being pushed in Step (ii-a) is pushed out in Steps (ii-b) and (ii-c), subsequently this flow is further pushed to sink *t* in Steps (ii-d) and (ii-e). Thus, *f* is feasible. **If part:** Here we assume $\varphi_{f^*} = q + n$ and show that the system (\bar{A}, \bar{B}) is structurally controllable by proving that both the accessibility and no-dilation conditions are satisfied.

Assume $\varphi_{f^*} = q + n$ in the flow-network $\mathcal{F}(\bar{A}, \bar{B})$. Since there are exactly q+n edges, each with capacity one, emanating from the source vertex *s*, each of these edges should carry exactly one unit flow. Since f^* is a feasible flow vector, it satisfies flow conservation at each node in $V_F \setminus \{s,t\}$. Specifically, the flow conservation is satisfied at nodes $\mathcal{N}_1, \ldots, \mathcal{N}_q$ and x'_1, \ldots, x'_n . Thus, for every \mathcal{N}_i , there exist u'_j such that $(\mathcal{N}_i, u'_j) \in E_F$. Thus, all non-top linked SCCs are connected to atleast one input. This ensures that all states are accessible.

Furthermore, on account of flow conservation at nodes x'_1, \ldots, x'_n and flow integrality, there exists $y_k \in V_X \cup V_U$ such that $f^*((x'_k, y_k)) = 1$. Since the capacity of outgoing edges from each node in $V_X \cup V_U$ is one, it follows that $y_{k_1} \neq y_{k_2}$ unless $k_1 = k_2$. Now, note that the set $\{(x'_k, y_k) : k = 1, \ldots, n\}$ is a matching in $\mathcal{B}(\bar{A}, \bar{B})$. Since there exists a perfect matching in $\mathcal{B}(\bar{A}, \bar{B})$, there are no dilations in $\mathcal{D}(\bar{A}, \bar{B})$. This proves the required using Propositions 2.2 and 2.3.

The following result about certain subsets of inputs also yielding structural controllability follows from Theorem 3.3 by observing that the maximum flow through $\mathcal{F}(\bar{A}, \bar{B}_{W_r}) = q + n$.

Corollary 3.4. Consider $\mathcal{F}(\bar{A},\bar{B})$ and any feasible flow vector f such that $\varphi_f = q + n$ and define $\mathcal{W}_f = \{j : f(u'_j,t) > 0\}$. Then, $(\bar{A},\bar{B}_{\mathcal{W}_f})$ is structurally controllable.

Remark 3.5. The above result allows for obtaining a subset of all possible inputs that are enough to retain controllability of the structured system from the obtained flow vector. Conversely, the structural controllability of the system with a given subset of inputs, say W, can be checked using Theorem 3.3 for any flow-network $\mathcal{F}(\bar{A}, \bar{B}_W)$.

Constructing the flow-network $\mathcal{F}(\bar{A},\bar{B})$ corresponding to a structured system (\bar{A},\bar{B}) has $O(n^2)$ complexity⁷ and checking

⁷Finding SCCs has $O(n^2)$ complexity and other constructions in Algorithm 3.1 are of linear complexity [14].

structural controllability using the maximum flow formulation has $O(n^3)$ complexity [18], where *n* denotes the number of states in the system. However, using the two conditions from the literature (i.e., accessibility and no-dilation), structural controllability can be checked with complexity $O(n^{2.5})$. Even though complexity is slightly higher in our case, the flownetwork constructed above is useful in two ways: (a) for checking if a given system is structurally controllable and (b) to optimize the input set for solving the minCCIS problem. In order to use $\mathcal{F}(\bar{A}, \bar{B})$ to optimize the number of inputs for the minCCIS problem, we augment the maximum flow problem with some additional features, called the minimum-cost fixedflow problem. This is elaborated in Section 4.

4. APPROXIMATING THE MINCCIS PROBLEM

In this section, we augment the flow-network with a cost for edge usage in order to solve the minCCIS problem. Specifically, we show that solving an MCFF problem (see for example [20]) on a flow-network we design is equivalent to solving the minCCIS problem.

A. Minimum Controllability Problem as MCFF

We first describe the MCFF problem for completeness and subsequently demonstrate its utility for solving the minCCIS problem. A MCFF problem takes as input a directed flownetwork $\mathcal{F}(V, E)$, with vertex set V, edge set E, source-sink pair s,t, non-negative capacities b(e), non-negative costs c(e)for $e \in E$ and flow requirement φ_{\min} . Then, the solution to MCFF(φ_{\min}) is a feasible flow vector f_M^* such that $\varphi_{f_M^*} \ge \varphi_{\min}$ and $\sum_{e \in E: f_M^*(e) > 0} c(e) \le \sum_{e \in E: f(e) > 0} c(e)$, for any feasible flow vector f. Thus, MCFF(φ_{\min}) solves the following constrained optimization:

Problem 4.1. *Minimize:* $\sum_{e \in E: f(e) > 0} c(e)$ Subject to: (1) f is a feasible flow vector, and (2) $\varphi_f \ge \varphi_{\min}$.

Any feasible solution to Problem 4.1 is referred to as a feasible solution to MCFF(φ_{\min}). Note that MCFF(φ_{\min}) has a feasible solution if and only if $\varphi_{f^*} \ge \varphi_{\min}$.

To establish a relation between the MCFF and the minCCIS problems, we formulate minCCIS as an instance of MCFF such that an optimal solution f_M^* to MCFF corresponds to an optimal solution to minCCIS. Given a structured system (\bar{A},\bar{B}) and a cost vector p_u , such that each entry $p_u(j)$, for j = 1, 2, ..., m, corresponds to the cost associated with each input, we consider flow-network $\mathcal{F}(\bar{A},\bar{B})$ augmented with cost vector c (referred to as $\mathcal{F}(\bar{A},\bar{B},c)$ in the sequel) as follows:

$$c(e) \leftarrow \begin{cases} p_u(j), & \text{for } e = (u'_j, t), j \in \{1, 2, \dots, m\}, \\ 0, & \text{otherwise.} \end{cases}$$
(3)

On this flow-network, we solve MCFF(q+n). We have the following preliminary result.

Lemma 4.2. A structured system (\bar{A}, \bar{B}) is structurally controllable if and only if MCFF(q+n) has a feasible solution on $\mathcal{F}(\bar{A}, \bar{B}, c)$.

Proof. Only-if part: Here we will prove that if $(\overline{A}, \overline{B})$ is structurally controllable, then MCFF(q+n) has a feasible

solution on $\mathcal{F}(\bar{A}, \bar{B}, c)$. By Theorem 3.3 we know that if (\bar{A}, \bar{B}) is structurally controllable, then the maximum flow through $\mathcal{F}(\bar{A}, \bar{B})$ equals q+n. Thus there exists a feasible flow vector f of $\mathcal{F}(\bar{A}, \bar{B}, c)$ such that $\varphi_f \ge q+n$. Thus f is indeed a feasible solution to MCFF(q+n). This completes the only-if part.

If part: Here we will prove that if there exists a feasible solution to MCFF(q+n) on $\mathcal{F}(\bar{A},\bar{B},c)$, then (\bar{A},\bar{B}) is structurally controllable. A feasible solution to MCFF(q+n) is a feasible flow vector f such that $\varphi_f \ge q+n$. Since $\varphi_f \ge q+n$, the maximum flow vector f^* in $\mathcal{F}(\bar{A},\bar{B},c)$ gives $\varphi_f^* \ge q+n$. Thus by Theorem 3.3, the structured system (\bar{A},\bar{B}) is structurally controllable. This completes the proof.

Henceforth, we consider a structurally controllable system (\bar{A}, \bar{B}) with *n* states and *q* number of non-top linked SCCs in $\mathcal{D}(\bar{A})$. Let *f* be any feasible solution to MCFF(*q*+*n*). Define,

$$\mathcal{I}_f := \{ j : f(u'_j, t) > 0 \}, \text{ and}$$
 (4)

$$c_f := \sum_{e:f(e)>0} c(e).$$
(5)

Also define $c^* = c_{f_M^*}$ as the optimal cost for MCFF(q+n) on $\mathcal{F}(\bar{A}, \bar{B}, c)$. Using (4) and (5), we now describe how a solution to minCCIS can be obtained from a feasible solution f of MCFF(q+n). For a given f, if $j \in \mathcal{I}_f$, then we propose to use input u_j . Following result holds.

Lemma 4.3. If f is a feasible solution to MCFF(q+n)on $\mathcal{F}(\bar{A},\bar{B},c)$, then $(\bar{A},\bar{B}_{\mathcal{I}_f})$ is structurally controllable and $p(\mathcal{I}_f) = c_f$.

Proof. Since f is a feasible solution to MCFF(q + n) on $\mathcal{F}(\bar{A}, \bar{B}, c)$, f is also a feasible solution to MCFF(q+n) on the flow-network $\mathcal{F}(\bar{A}, \bar{B}_{\mathcal{I}_f})$. Now by Lemma 4.2 the structured system $(\bar{A}, \bar{B}_{\mathcal{I}_f})$ is structurally controllable. Finally, $p(\mathcal{I}_f) = c_f$ follows from equations (4), (5) and the cost definition given by equation (3).

Now we prove the equivalence between minCCIS and MCFF(q+n) through the following theorem.

Theorem 4.4. Consider a structured system (\bar{A}, \bar{B}) with n states and q non-top linked SCCs in $\mathcal{D}(\bar{A})$. The flow-network $\mathcal{F}(\bar{A}, \bar{B}, c)$ can be constructed in $O(n^2)$ computations. Further, for an optimal flow vector f_M^* of MCFF(q+n) on $\mathcal{F}(\bar{A}, \bar{B}, c)$, $\mathcal{I}_{f_M^*} = \{j : f_M^*((u'_j, t) > 0)\}$ is an optimal solution to minCCIS. Moreover, the optimal cost of minCCIS equals the optimum cost of MCFF(q+n), i.e., $p^* = c^*$.

Proof. Constructing the flow-network $\mathcal{F}(\bar{A}, \bar{B})$ has complexity $O(n^2)$. In addition to this we define a cost vector c and flow requirement φ_{\min} to construct $\mathcal{F}(\bar{A}, \bar{B}, c)$. Since these are each of linear complexity, we conclude that complexity involved in constructing the flow-network $\mathcal{F}(\bar{A}, \bar{B}, c)$ is $O(n^2)$.

We now prove that $p^* = c^*$. Let f_M^* be an optimal flow vector for MCFF(q+n) on $\mathcal{F}(\bar{A},\bar{B},c)$. Note that by definition $c_{f_M^*} = c^*$, and by Lemma 4.3 $c^* = p(\mathcal{I}_{f_M^*})$. Now we show that $\mathcal{I}_{f_M^*}$ is an optimal solution to minCCIS. First, we argue that $\mathcal{I}_{f_M^*}$ is a feasible solution to minCCIS, i.e. $\mathcal{I}_{f_M^*} \in \mathcal{K}$. It follows from Lemma 4.3 that the system $(\bar{A}, \bar{B}_{\mathcal{I}_{f_M^*}})$ is structurally controllable. Thus, $\mathcal{I}_{f_M^*} \in \mathcal{K}$. Suppose $\mathcal{I}_{f_M^*}$ is not an optimal solution to minCCIS. Then there exists $\mathcal{I} \in \mathcal{K}$ such that $p(\mathcal{I}) < p(\mathcal{I}_{f_{M}^{\star}})$. Consider the flow-network $\mathcal{F}(\bar{A}, \bar{B}_{\mathcal{I}})$. By Theorem 3.3, there exists a feasible flow vector f in $\mathcal{F}(\bar{A}, \bar{B}_{\mathcal{I}})$ such that $\varphi_f \ge q+n$. Since $\mathcal{F}(\bar{A}, \bar{B}_{\mathcal{I}})$ is a sub-graph of $\mathcal{F}(\bar{A}, \bar{B})$, f is also a feasible flow vector in $\mathcal{F}(\bar{A}, \bar{B})$ with $\varphi_f \ge q+n$. We note that $\mathcal{I}_f \subseteq \mathcal{I}$. Thus, from Lemma 4.3, $c_f \le p(\mathcal{I}) < p(\mathcal{I}_{f_M^{\star}}) = c^{\star}$. This contradicts optimality of f_M^{\star} . Thus, $\mathcal{I}_{f_M^{\star}}$ is an optimal solution to minCCIS. Finally, $p^{\star} = c^{\star}$ follows from Lemma 4.3 and optimality of $\mathcal{I}_{f_M^{\star}}$ for minCCIS. This completes the proof.

Thus given an instance of minCCIS, we construct $\mathcal{F}(\bar{A}, \bar{B}, c)$ and reduce⁸ it to an MCFF as discussed. After solving MCFF(q+n), we get an optimal flow f_M^{\star} . From f_M^{\star} , we get back the corresponding optimal solution to minCCIS, $\mathcal{I}^{\star} = \mathcal{I}_{f_M^{\star}} = \{j : f_M^{\star}(u_j', t) > 0\}, \text{ i.e., the minimum cost incur$ ring set of inputs selected under f_M^{\star} . Unfortunately, MCFF is also a known NP-hard problem. However, it is a well studied problem as it relates to many fields including job-shop scheduling, transportation networks and computer networks [21]. For MCFF, approximation algorithms in addition to many good heuristics exist (see [22] and references therein). One can potentially use these existing algorithms to obtain an approximate solution to minCCIS. However to do this an approximate solution to the MCFF problem must yield an approximate solution to the minCCIS problem. We establish this next.

Theorem 4.5. Let $\varepsilon \ge 1$ and let f be any feasible solution to MCFF(q+n) on $\mathcal{F}(\bar{A},\bar{B},c)$. Suppose c^* and p^* are the optimum costs of MCFF(q+n) and minCCIS respectively. Then $c_f \le \varepsilon c^*$ implies $p(\mathcal{I}_f) \le \varepsilon p^*$.

Proof. The result follows from Lemma 4.3 and Theorem 4.4. However, for completeness we explain the proof. Since f is a feasible solution to MCFF(q+n) on $\mathcal{F}(\bar{A},\bar{B},c)$, by Lemma 4.3, $p(\mathcal{I}_f) = c_f$. Also, by Theorem 4.4, $p^* = c^*$. Since $c_f \leq \varepsilon c^*$, we get $p(\mathcal{I}_f) \leq \varepsilon c^* = \varepsilon p^*$.

Note that a feasible solution f that satisfies the condition in Theorem 4.5 is called an ε -optimal solution or an ε approximate solution. In the next section, we obtain an approximation algorithm for MCFF(q+n) on $\mathcal{F}(\bar{A},\bar{B},c)$.

B. An Approximation Algorithm for MCFF on $\mathcal{F}(\bar{A}, \bar{B}, c)$

MCFF over general graphs are shown to be hard to approximate [20], [23]. The best known algorithm gives a φ_{\min} -approximate solution to MCFF(φ_{\min}) [20]. Note that in our case $\varphi_{\min} = q + n$ and thus the approximation factor is linear in n, which is not desirable. Next we propose a polynomial complexity Δ -approximate solution to MCFF(q + n) on $\mathcal{F}(\bar{A}, \bar{B}, c)$, where Δ is the maximum in-degree of nodes u'_j 's in $\mathcal{F}(\bar{A}, \bar{B}, c)$. Note that $1 < \Delta \leq q + 1$.

Problem 4.6. Consider a flow-network $\mathcal{F}(\bar{A}, \bar{B}, c)$ and define the following Linear Program (LP): Minimize: $\sum_{c} c(e) f(e)$

Minimize: $\sum_{e \in E_F} c(e)f(e)$ Subject to: (1) f is a feasible flow vector, and (2) $\varphi_f \ge q + n$.

⁸A *polynomial reduction* of Problem A into Problem B helps in obtaining a solution to Problem A from a solution to Problem B in number of computations that is at most polynomial in the size of the input data. Problem 4.6 is a well studied flow problem, the *minimum* cost flow problem (MCFP) [15]. The key difference between the MCFF problem and the MCFP is that in the former, the cost incurred does not depend on the flow through an edge, rather the cost depends only on whether the edge is used; however in the latter, the cost increases linearly with the flow through the edge. MCFP can be solved in polynomial time [24]. Let the value of the objective function in Problem 4.6 for a feasible flow vector be C_f and let C_{OPT} denote the minimum value of the objective function of Problem 4.6. Also, let f_{LP}^{\star} be the corresponding optimal flow vector, i.e. $C_{\text{OPT}} \leq C_f$ for any feasible solution f of the LP. The following preliminary result holds as a direct consequence of [15, Theorem 9.8].

Lemma 4.7. Let f_{LP}^{\star} be an optimal flow vector for Problem 4.6. Then, for every $e \in E_F$, $f_{LP}^{\star}(e) \in \mathbb{Z}$.

In Theorem 4.8, we obtain a relation between the cost of the inputs selected under f_{LP}^* and the optimal cost of MCFF(q+n).

Theorem 4.8. The flow vector f_{LP}^{\star} is a Δ -approximate solution of MCFF(q+n), i.e. $c_{f_{LP}^{\star}} \leq \Delta c^{\star}$, where Δ is the maximum indegree for nodes in $\{u_1^{\prime}, \ldots, u_m^{\prime}\}$.

Proof. Let C_{OPT} be the optimal value of the LP and given c^* is the optimal cost of MCFF(q + n). Note that b(e) = 1 for every $e \in E_F \setminus \{(u'_j, t)\}_{j=1}^m$. Thus, the total flow carried by any $e \in E_F$ is at most Δ under any feasible flow vector f. Hence, we have the following:

$$C_{\text{OPT}} = \sum_{e} f_{\text{LP}}^{\star}(e)c(e) \leqslant \sum_{e} f_{M}^{\star}(e)c(e), \tag{6}$$

$$=\sum_{e:f_M^{\star}(e)>0}f_M^{\star}(e)c(e) \leqslant \Delta \sum_{e:f_M^{\star}(e)>0}c(e) = \Delta c^{\star}.$$
 (7)

The inequality in equation (6) follows as f_M^* is a feasible solution to the LP. The inequality in equation (7) follows as $f(e) \leq \Delta$ for every edge *e* and every feasible flow vector *f*. Thus, $C_{\text{OPT}} \leq \Delta c^*$. The cost of the inputs selected under f_{LP}^*

$$c_{f_{\text{LP}}^{\star}} = \sum_{e: f_{\text{LP}}^{\star}(e) > 0} c(e) \leqslant \sum_{e: f_{\text{LP}}^{\star}(e) > 0} f_{\text{LP}}^{\star}(e) c(e) = C_{\text{OPT}}.$$
 (8)

The inequality in equation (8) follows as if $f_{LP}^{\star}(e) > 0$, then $f_{LP}^{\star}(e) \ge 1$ by Lemma 4.7. Thus, $C_{OPT} \le \Delta c^{\star}$ and $c_{f_{LP}^{\star}} \le C_{OPT}$. Thus by combining both the inequalities we get, $c_{f_{LP}^{\star}} \le C_{OPT} \le \Delta c^{\star}$. This completes the proof.

Remark 4.9. Note that the number of non-top linked SCCs is atmost n. Thus, in the worst case $\Delta = O(n)$. This corresponds to states being decoupled. However, in practical systems the states interact and as a result the number of non-top linked SCCs is usually much smaller than n [4]. In such cases, the above algorithm gives a tighter approximation.

There exist various polynomial time algorithms for solving Problem 4.6 [15], [24]. The algorithm in [24] has complexity $O(\ell^4 \log \ell)$ on a generic flow-network with ℓ nodes. However, because of the special structure of the flow-network $\mathcal{F}(\bar{A}, \bar{B}, c)$, Problem 4.6 can be solved using a simpler algorithm that incorporates a minimum weight perfect matching algorithm and a greedy scheme. We describe the pseudo-code of this twostage procedure in Algorithm 4.1. In Stage 1 of Algorithm 4.1, we run a minimum weight perfect matching on $\mathcal{B}(\bar{A}, \bar{B})$ with Algorithm 4.1 Pseudo-code for solving Problem 4.6 using minimum weight maximum matching and a greedy scheme

Input: Structured system (\bar{A}, \bar{B}) , input cost vector p_u and flow-network $\mathcal{F}(\bar{A}, \bar{B}, c)$ **Output:** Flow vector f_A constructed using Algorithm 4.2 with inputs M_A and S_A .

Algorithm 4.2 Pseudo-code for constructing flow vector *f* from perfect matching *M* and set of edges $S = \{(\mathcal{N}_i, u'(\mathcal{N}_i))\}_{i=1}^q$

Input: Perfect matching *M* and edge set *S* Output: Flow vector *f* 1: $\mathcal{X}_M \leftarrow \{x_r : (x_r, x'_k) \in M\}$ 2: $\mathcal{U}_M \leftarrow \{u_j : (u_j, x'_k) \in M\}$ 3: $\mathcal{U}_S \leftarrow \{u_j : (\mathcal{N}_i, u'_j) \in S\}$ 4: Define the flow vector *f* as 1, for $e \in \{(s, x'_k), (x'_k, x_r), (x_r, t)\}$, for k = 1, ..., n and $x_r \in \mathcal{X}_M$, 1, for $e \in \{(s, x'_k), (x'_k, u_j), (u_j, u'_j)\}$, for k = 1, ..., n and $u_j \in \mathcal{U}_M$, 1, for $e \in \{(s, \mathcal{N}_i), (\mathcal{N}_i, u'_j)\}$, for i = 1, ..., n and $u_j \in \mathcal{U}_S$, $\sum_{k=1}^n \sum_{j=1}^m \mathbb{I}_{\{(x'_k, u_j) \in M\}} + \sum_{i=1}^q \sum_{j=1}^m \mathbb{I}_{\{(\mathcal{N}_i, u'_j) \in S\}}$, for $e = (u'_j, t)$, for j = 1, ..., m, 0, otherwise.

weights defined as shown in Step 2. Let M_A be a matching obtained as solution of Stage 1 (see Step 3). In Stage 2, we perform a greedy selection to connect all the non-top linked SCCs to some input. To achieve this for all \mathcal{N}_i 's, $i \in \{1, \ldots, q\}$, we greedily assign the least cost input from the set of all inputs that has an edge to some state in \mathcal{N}_i (see Step 5). Let the least cost input corresponding to \mathcal{N}_i be $u(\mathcal{N}_i)$. We define $S_A = \{(\mathcal{N}_i, u'(\mathcal{N}_i))\}_{i=1}^q$ (see Step 6). Finally, we use Algorithm 4.2 to construct a flow vector f_A based on M_A and S_A . We prove the optimality of the constructed flow vector f_A after stating the following supporting lemmas.

Lemma 4.10. Given any valid inputs M and S of Algorithm 4.2, let f denote the output flow vector. Then f is a feasible solution to the LP given in Problem 4.6. Moreover, the value

$$C_{f} = \sum_{k=1}^{n} \sum_{j=1}^{m} p_{u}(j) \mathbb{I}_{\{(x'_{k}, u_{j}) \in M\}} + \sum_{i=1}^{q} \sum_{j=1}^{m} p_{u}(j) \mathbb{I}_{\{(\mathcal{N}_{i}, u'_{j}) \in \mathcal{S}\}}, \quad (9)$$

where $\mathbb{I}_{\mathscr{A}}$ is the indicator function on subset $\mathscr{A} \subset E_F$.

Proof. From the construction of f as per Algorithm 4.2, it follows that f satisfies both flow conservation and capacity constraint on $\mathcal{F}(\bar{A},\bar{B},c)$. Thus, f is feasible in $\mathcal{F}(\bar{A},\bar{B},c)$. Moreover, note that f(e) = 1 for every e that emanates from the source vertex s. Hence $\varphi_f = q + n$. This shows that f is a feasible solution to the LP given in Problem 4.6.

Now, note that since costs are non-zero only for the edges between u'_j and t. Also, $c((u'_j,t)) = p_u(j)$. Thus, $C_f = \sum_{j=1}^{m} p_u(j) f((u'_j,t))$. The flow value $f((u'_j,t))$ equals the sum of the flows coming from edges (u_j, u'_j) and (\mathcal{N}_i, u'_j) . Note that the second term in equation (9) corresponds to the total cost contributed by the flow from edges (\mathcal{N}_i, u'_j) . Now, if the flow on (u_j, u'_j) is greater than zero, then it has to come from some edge $(x'_k, u_j) \in M$. Thus, the first term in equation (9) corresponds to the total cost contributed by the flow from edges (u_j, u'_j) . This proves the required.

Lemma 4.11. The sets M_A and S_A given by Algorithm 4.1 satisfies the following:

(1) For any perfect matching M of $\mathcal{B}(\bar{A},\bar{B})$,

$$\sum_{k=1}^{n} \sum_{j=1}^{m} p_{u}(j) \mathbb{I}_{\{(x'_{k}, u_{j}) \in M_{A}\}} \leqslant \sum_{k=1}^{n} \sum_{j=1}^{m} p_{u}(j) \mathbb{I}_{\{(x'_{k}, u_{j}) \in M\}}.$$
(2) For any set $S = \{(\mathscr{N}_{i}, y'_{i}) : y_{i} \in V_{U}, (\mathscr{N}_{i}, y'_{i}) \in E_{F}\}_{i=1}^{q},$

$$\sum_{i=1}^{7} \sum_{j=1}^{m} p_{u}(j) \mathbb{I}_{\{(\mathcal{N}_{i}, u'_{j}) \in \mathcal{S}_{A}\}} \leq \sum_{i=1}^{7} \sum_{j=1}^{m} p_{u}(j) \mathbb{I}_{\{(\mathcal{N}_{i}, u'_{j}) \in \mathcal{S}\}}.$$

Proof. The result is an immediate consequence of the way in which the Algorithm 4.1 constructs M_A and S_A . Here, (1) follows since M_A is a minimum weight perfect matching in $\mathcal{B}(\bar{A},\bar{B})$ and (2) follows since set S_A is obtained by greedily selecting a minimum cost input corresponding to each non-top linked SCC \mathcal{N}_i .

Using the above lemmas we prove the following result.

Theorem 4.12. Let C_{f_A} be the value of the objective function in Problem 4.6 for a flow vector f_A constructed using Algorithm 4.2 and let C_{OPT} be the optimal value of the LP. Then $C_{f_A} = C_{\text{OPT}}$.

Proof. First, observe from Lemma 4.10 that f_A is a feasible solution to the LP described in Problem 4.6. Thus, $C_{f_A} \ge C_{\text{OPT}}$. Now, we get the result if we can show $C_{\text{OPT}} \ge C_{f_A}$. Let f_{LP}^* be the optimal flow vector. Define, the following sets:

$$M^{\star} = \{ (x'_j, y_j) : f_{LP}^{\star}((x'_j, y_j)) > 0, y_j \in V_X \cup V_U \}, \\ S^{\star} = \{ (\mathcal{N}_i, u'_j) : f_{LP}^{\star}(\mathcal{N}_i, u'_j) > 0 \text{ for some } i \in \{1, \dots, q\} \}.$$

Note that M^* is a perfect matching in $\mathcal{B}(\bar{A}, \bar{B})$. Also, \mathcal{S}^* has an outgoing edge from every non-top linked SCC to some input u'_j . Note that f^*_{LP} can be thought as a flow vector constructed from (M^*, \mathcal{S}^*) using Algorithm 4.2. Now the result follows from (9) and Lemma 4.11.

Note that though we showed approximation using greedy scheme, any heuristics can also be used. The following result quantifies complexity of Algorithm 4.1. **Lemma 4.13.** Algorithm 4.1, which computes an approximate solution to the minCCIS problem, has complexity $O(n^3)$, where *n* denotes the number of states in the structured system (\bar{A}, \bar{B}) .

Proof. Stage 1 of Algorithm 4.1 where we solve a minimum weight perfect matching has complexity $O(n^3)$ [14]. Stage 2 of Algorithm 4.1 where a greedy scheme is employed to connect all non-top linked SCCs has $O(n^2)$ complexity, since q = O(n) and m = O(n). From the two stages we get M_A and S_A . Construction of flow vector f_A using M_A and S_A given in Algorithm 4.2 is of linear complexity. Thus, Algorithm 4.1 has complexity $O(n^3)$.

C. Special Cases

In this subsection, we discuss two special cases of Problem 1.2. Consider the case when $\mathcal{B}(\bar{A})$ has a perfect matching. Note that Problem 1.2 is NP-hard over this special class of structured systems also [12]. We give the following approximation result for this class of systems.

Corollary 4.14. Suppose there exists a perfect matching in $\mathcal{B}(\bar{A})$. Then, an optimal solution to Problem 4.6, where flow f_{LP}^{\star} is to be computed for an MCFP, is a $(\Delta - 1)$ -approximate solution to MCFF(q+n), i.e. $c_{f_{LP}^{\star}} \leq (\Delta - 1)c^{\star}$.

Now consider the case when $\mathcal{D}(\bar{A})$ is *irreducible*. Then, Problem 1.2 is in P and the following result holds.

Corollary 4.15. Suppose $\mathcal{D}(\bar{A})$ is irreducible. Then an optimal solution to Problem 4.6, where flow f_{LP}^{\star} is to be computed for an MCFP, gives an optimal solution to MCFF(q + n). Also, $c_{f_{LP}^{\star}} = c^{\star}$.

Remark 4.16. By duality between controllability and observability in linear time invariant systems, all the analysis and results discussed in this paper are directly applicable to the minimum cost constrained output selection problem.

5. CONCLUSION

This paper addressed minimum cost constrained input selection (minCCIS) problem for structural controllability. We studied minCCIS on a general structured system, which is a known NP-hard problem [12]. In this paper, we provided a new graph theoretic necessary and sufficient condition based on flow-networks for checking structural controllability (Theorem 3.3). The link between structural controllability and flowbased-networks is one of our central contributions in this paper. Then we obtained a polynomial reduction of minCCIS to an NP-hard variant of the maximum flow problem: the minimum-cost fixed-flow problem (MCFF). We showed that an optimal solution to MCFF corresponds to an optimal solution to minCCIS (Theorem 4.4). We also showed that approximation schemes available for solving MCFF can be used to solve minCCIS (Theorem 4.5). Using the special structure of the flow-network constructed from the structured system (\bar{A}, \bar{B}) , we proposed an approximation algorithm to solve minCCIS. In our main result (Theorem 4.8) we gave a polynomial time algorithm that obtains a Δ -approximate solution to minCCIS. Needless to elaborate the same result holds for structural *observability* of (\bar{C},\bar{A}) and the minimum sensor selection problem. Further, the results are applicable to discrete time systems too.

REFERENCES

- K. J. Reinschke, Multivariable control: A graph theoretic approach. Springer-Verlag, 1988.
- [2] C.-T. Lin, "Structural controllability," *IEEE Transactions on Automatic Control*, vol. 19, no. 3, pp. 201–208, 1974.
- [3] Y.-Y. Liu, J.-J. Slotine, and A.-L. Barabási, "Controllability of complex networks," *Nature*, vol. 473, no. 7346, pp. 167–173, 2011.
- [4] S. Pequito, V. M. Preciado, A.-L. Barabási, and G. J. Pappas, "Tradeoffs between driving nodes and time-to-control in complex networks," *Scientific Reports*, vol. 7, pp. 1–7, 2017.
- [5] G. Ramos, S. Pequito, A. P. Aguiar, and S. Kar, "Analysis and design of electric power grids with p-robustness guarantees using a structural hybrid system approach," in *Proceedings of the IEEE European Control Conference (ECC)*, Linz, Austria, 2015, pp. 3542–3547.
- [6] C. Commault, J.-M. Dion, and J. W. van der Woude, "Characterization of generic properties of linear structured systems for efficient computations," *Kybernetika*, vol. 38, no. 5, pp. 503–520, 2002.
- [7] C. Commault and J.-M. Dion, "The single-input minimal controllability problem for structured systems," *Systems & Control Letters*, vol. 80, pp. 50–55, 2015.
- [8] S. Pequito, S. Kar, and A. P. Aguiar, "A structured systems approach for optimal actuator-sensor placement in linear time-invariant systems," in *Proceedings of the IEEE American Control Conference (ACC)*, Washington DC, USA, 2013, pp. 6108–6113.
- [9] 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, 2016.
- [10] A. Olshevsky, "Minimum input selection for structural controllability," in Proceedings of the IEEE American Control Conference (ACC), Chicago, USA, 2015, pp. 2218–2223.
- [11] Y.-Y. Liu and A.-L. Barabási, "Control principles of complex systems," *Reviews of Modern Physics*, vol. 88, no. 3, p. 035006, 2016.
 [12] S. Pequito, S. Kar, and A. P. Aguiar, "On the complexity of the
- [12] S. Pequito, S. Kar, and A. P. Aguiar, "On the complexity of the constrained input selection problem for structural linear systems," *Automatica*, vol. 62, pp. 193–199, 2015.
- [13] S. Pequito, S. Kar, and G. J. Pappas, "Minimum cost constrained inputoutput and control configuration co-design problem: A structural systems approach," in *Proceedings of the IEEE American Control Conference* (ACC), Chicago, USA, 2015, pp. 4099–4105.
- [14] R. Diestel, Graph Theory. Springer: New York, 2000.
- [15] R. K. Ahuja, T. L. Magnanti, and J. B. Orlin, *Network Flows: Theory, Algorithms, and Applications.* Prentice Hall: New Jersey, 1993.
- [16] T. Yamada, "A network flow algorithm to find an elementary i/o matching," *Networks*, vol. 18, no. 2, pp. 105–109, 1988.
- [17] V. Hovelaque, C. Commault, and J. Dion, "Analysis of linear structured systems using a primal-dual algorithm," *Systems & control letters*, vol. 27, no. 2, pp. 73–85, 1996.
- [18] J. B. Orlin, "Max flows in O(nm) time, or better," in *Proceedings of the 45-th Annual ACM Symposium on Theory of Computing*, Palo Alto, USA, 2013, pp. 765–774.
- [19] L. R. Ford Jr and D. R. Fulkerson, "A simple algorithm for finding maximal network flows and an application to the hitchcock problem," DTIC Document, Tech. Rep., 1955.
- [20] S. O. Krumke, H. Noltemeier, S. Schwarz, H.-C. Wirth, and R. Ravi, "Flow improvement and network flows with fixed costs," in *Operations Research Proceedings*. Springer: Berlin, Heidelberg, 1999, pp. 158–167.
- [21] T. L. Magnanti and R. T. Wong, "Network design and transportation planning: Models and algorithms," *Transportation Science*, vol. 18, no. 1, pp. 1–55, 1984.
- [22] D. B. Khang and O. Fujiwara, "Approximate solutions of capacitated fixed-charge minimum cost network flow problems," *Networks*, vol. 21, no. 6, pp. 689–704, 1991.
- [23] G. Even, G. Kortsarz, and W. Slany, "On network design problems: fixed cost flows and the covering steiner problem," ACM Transactions on Algorithms, vol. 1, no. 1, pp. 74–101, 2005.
- [24] J. B. Orlin, "A faster strongly polynomial minimum cost flow algorithm," Operations Research, vol. 41, no. 2, pp. 338–350, 1993.