Energy-Transfer Edge Centrality and Its Role in Enhancing Network Controllability

Prasad Vilas Chanekar* Erfan Nozari Jorge Cortés

Abstract—The ability to modify the structure of network systems offers great opportunities to enhance their operation, improve their efficiency, and increase their resilience against failures and attacks. This paper focuses on the edge modification problem, i.e., improving network controllability by adding and/or re-weighting interconnections while keeping the actuation structure fixed. We consider a network system following linear dynamics and propose a novel edge centrality measure that captures the extent to which an edge facilitates energy exchange across the network through its defining nodes. We analyze the effectiveness of the proposed measure by characterizing its relationship with the gradient (with respect to edge weights) of trace, log determinant, and inverse of the trace inverse of the Gramian. We show that the optimal solution of the edge modification problem lies on the boundary of the feasible search space when the objective is the trace of the Gramian or the network has a diagonal controllability Gramian and the objective is either log determinant or the inverse of the trace inverse of the Gramian. Finally, using the proposed edge centrality measure we design two network modification algorithms that restrict the search space to a smaller subset of all possible edges and numerically demonstrate their efficacy.

Index Terms—Complex networks, network controllability, edge centrality, network structure modification

1 Introduction

Recent years have witnessed a surge of interest from the scientific community in improving the controllability of complex dynamical networks, with a wide range of applications in infrastructure, robotic, biological, and social networks. This body of work has been largely focused on optimizing the location of the actuators and sensors while assuming a given network interconnection (whether fixed or time-varying). Complementary to this effort, the modification of the network structure raises exciting and unexplored opportunities for improving functionality and resilience. While network modification is certainly possible for man-made systems, it is becoming increasingly feasible for biological ones, such as gene regulatory and neuronal networks [2], [3]. These observations motive the focus on this paper on the study of the optimal edge modification problem for network systems described by linear dynamics.

Literature review: The controllability notion for a dynamical system addresses the question of whether its state can be steered to any desired goal. Controllability of linear control systems can be established in a number of equivalent ways [4], [5] which, however, do not quantify the effort required to steer the system [6]. In the context of network systems, a recent body of work [7]–[9] has introduced metrics based on the controllability Gramian that quantify the minimal energy necessary to move between any pair of states. This has opened the way to explore a number of exciting problems, including the relationship between network structure and ease of controllability [8], the role of centrality [10]–[12], and the optimal placement of actuators and sensors [9], [13]–[15] to maximize network controllability. In the latter problem, it is worth noting that if the existing edge structure of the network is not suitable for control, this may pose significant constraints on the best achievable degree of controllability via actuator scheduling. This has motivated recent works that study the dual problem of maximizing network controllability via edge modification. The work [16] gives a procedure to compute the minimum number of edges required to make a network structurally controllable and proposes a polynomial-time algorithm to determine their locations, but does not take into account the role of control energy or the weights of newly-added edges. The work [17] considers the edge addition problem in consensus networks and develops an algorithm for enhancing network performance measured by the trace of the observability Gramian. The work [18] introduces a class of systemic performance measures which are spectral functions of the Laplacian eigenvalues of the coupling graph and proposes a greedy procedure for edge addition limited to linear consensus networks. The work [19] develops a procedure to compute optimal perturbations of existing edges given various energy-based metrics which also allows for the addition of new edges if their location is given a priori. The work [20] establishes explicit relationships between network structure, weight distribution, and its degree of controllability when the Gramian is diagonal, and proposes an edge modification algorithm to generate stable and controllable networks with pre-specified diagonal Gramians. The edge modification problem has also been tackled making use of notions of network centrality. Originated in the network science literature, centrality is a measure to quantify the influence and importance of nodes and edges in a network. Many of these measures are topological in nature (relying on the network interconnection structure), generally oblivious to individual node dynamics and its effect on network behavior. While topological-based centrality measures include metrics for both nodes and edges [21]–[23] and have also been used for edge modification [24], [25], dynamics-based centrality measures encompass fewer metrics, which are limited to node centrality [11], [12], [26]. The recent work [27] proposes an edge centrality measure with respect to the H_2-norm for networks with continuous-time consensus dynamics having

A preliminary version appeared at the IEEE Conference on Decision and Control as [1]. This work was supported by ARO Award W911NF-18-1-0213 (PVC and JC) and NSF Award CMMI-1826065 (EN and JC).
*Corresponding author.
PVC and JC are with the Department of Aerospace and Mechanical Engineering, University of California, San Diego, {pchanekar, cortes}@ucsd.edu. *EN is with the Department of Mechanical Engineering, University of California, Riverside,* erfan.nozari@ucr.edu

time delays and structured uncertainties.

In this work, we propose an energy-based measure of edge centrality and build on it to design computational efficient edge modification algorithms to improve network controllability. This is in contrast with the state of the art where the edge modification problem is addressed for special cases (consensus networks or networks with diagonal controllability Gramian), or only on existing edges (i.e., without the possibility of adding new edges), or using topological notions of edge centrality that do not take the effect of network dynamics into account.

Statement of contributions: The contributions of this work pertain to the optimal edge modification problem for network systems described by linear dynamics. We introduce a novel energy-transfer edge centrality (ETEC) measure based on the Gramian that captures the energy transfer across the network through the nodes defining the edge. This measure is computationally inexpensive, making it suitable for use in large networks. We show the relationship of the proposed measure with the gradients of the trace, $\log\det$ and inverse of the trace inverse of the controllability Gramian by deriving lower and upper bounds that involve ETEC. On the basis of this relationship, existing or potentially new edges can be ranked in descending order of their ETEC as a proxy for their effect on enhancing network controllability. We further analytically establish the super-modularity nature of the trace of the Gramian with respect to edge selection and weighting. This, together with its non-decreasing nature with respect to edge weights, implies that the standard performance guarantees of greedy algorithms for optimization are not applicable. We also show that the optimal solution of the edge modification problem with the trace of the Gramian as objective lies on the boundary of its feasible search space. Additionally, we arrive at the same conclusion for systems with diagonal Gramian and $\log\det$ or inverse of the trace inverse of the controllability Gramian as objective. We rely on this result and the ordered list of edges to reduce the search space to those edges with the largest ETEC centrality measures and propose two computationally efficient network modification procedures, termed restricted-set optimization and restricted-set exhaustive greedy algorithms. We provide examples to numerically demonstrate the utility of the proposed algorithms in enhancing network controllability as measured by the trace and $\log\det$ of the Gramian.

Notation: We let $\mathbb{R}$ and $\mathbb{R}_{>0}$ denote the set of reals and positive reals, respectively. The j^{th} canonical unit vector is denoted by $e_j \in \mathbb{R}^n, j \in \{1,\dots,n\}$. For $x \in \mathbb{R}$, $\lfloor x \rfloor$ denotes the greatest integer less than or equal to x. For a matrix $X \in \mathbb{R}^{m\times m}$, we let its transpose, trace, logarithm-determinant, eigenvalue with smallest magnitude and its column-wise vectorized form be denoted by $X^\top$, $\mathrm{tr}(X)$, $\log\det(X)$, $\lambda_{\min}(X)$ and $\mathrm{vec}(X)$, respectively. The notation $X \geq 0$ $(X > 0)$ signifies that all the entries of the matrix X are non-negative (positive). The notation $X \leq 0$ $(X < 0)$ signifies that all the entries of the matrix X are non-positive (negative). We denote the Frobenius inner product of $X,Y \in \mathbb{R}^{n\times n}$ by $\langle X,Y\rangle_F = \mathrm{tr}\left(X^\top Y\right)$ and the Frobenius norm by $\|X\|_F = \sqrt{\langle X,X\rangle_F}$. Note that $\langle X,Y\rangle_F = \|X\|_F\|Y\|_F\cos\phi$, where $\phi = \angle(\mathrm{vec}(X),\mathrm{vec}(Y))$ is the angle between the vectors $\mathrm{vec}(X)$ and $\mathrm{vec}(Y)$. For a set $\mathscr{X}$, we denote its cardinality by $\mathrm{card}(\mathscr{X})$. We let $\mathbf{1}_k$ represent the vector of all 1's of dimension k.

2 Problem Description

We consider a directed network of n nodes represented by the triplet $\mathscr{G}_A = (\mathscr{V},\mathscr{E}_A,w_A)$, where $\mathscr{V} = \{1,2,\dots,n\}$ is the node set, $\mathscr{E}_A = \{(i,j) \mid i \in \mathscr{V}, j \in \mathscr{V}, i \neq j\}$ is the edge set, and $w_A : \mathscr{E}_A \mapsto \mathbb{R}_{>0}$ is the weight function. The pair (i,j) denotes an edge directed from node i to node j. We consider linear time-invariant dynamics,

$$x(t+1) = Ax(t) + Bu(t),\ t = \{0,\dots,T-1\}, \tag{1}$$

where $T > 0$ is a finite time horizon and $x \in \mathbb{R}^n$ and $u \in \mathbb{R}^m$ denote state and input vectors, respectively. Here, $A = (a_{ij}) \in \mathbb{R}^{n\times n}_{\geq 0}$ is the non-negative weighted adjacency matrix defined by $a_{ij} = w_A(j,i) > 0$ if $(j,i) \in \mathscr{E}_A$ and $a_{ij} = 0$ otherwise, and $B = (b_1 \cdots b_i \cdots b_m) \in \mathbb{R}^{n\times m}_{\geq 0}$ is the input matrix. In the case when the actuation is through individual nodes, then each $b_i \in \{0,1\}^n$ is a vector with 0's everywhere except at one entry, signifying the presence of an actuator at the corresponding node. Our treatment, however, is valid for arbitrary input matrices.

The dynamical system (1) is controllable in T steps if any arbitrary initial state $x(0) = x_0$ can be steered to any arbitrary final state $x(T) = x_T$ using a finite control input sequence $\{u(0),u(1),\dots,u(T-1)\}$. Equivalently, (A,B) is controllable in T steps if and only if the controllability Gramian

$$\mathscr{W}_A = \sum_{t=0}^{T-1} A^t BB^\top {A^t}^\top, \tag{2}$$

is non-singular [5]. This controllability test, however, does not distinguish between systems that are easier or harder to control. To do so, the literature [8], [9] has explored various performance metrics based on the spectral properties of $\mathscr{W}_A$, including $\mathrm{tr}(\mathscr{W}_A)$, $(\mathrm{tr}(\mathscr{W}_A^{-1}))^{-1}$, $\det(\mathscr{W}_A)$ (or its logarithm), and $\lambda_{\min}(\mathscr{W}_A)$. Here, we employ $\mathrm{tr}(\mathscr{W}_A)$, $\log\det(\mathscr{W}_A)$ and $(\mathrm{tr}(\mathscr{W}_A^{-1}))^{-1}$ as our measures of network controllability. The metric $\mathrm{tr}(\mathscr{W}_A)$ is related to the average controllability in all directions in the state space. The metric $\det(\mathscr{W}_A)$ is a volumetric measure of the set of states that can be reached with up to one unit of input energy. The metric $(\mathrm{tr}(\mathscr{W}_A^{-1}))^{-1}$ is related to the average energy needed to move the system around in the state space. Maximizing $\mathrm{tr}(\mathscr{W}_A)$ generally gives more relevance to large eigenvalues of $\mathscr{W}_A$ over smaller ones, which might make a few particular directions hard to reach. This is part compensated by the fact that, for large networks, controllability is not needed in all the state-space directions in general, cf. [12]. For large networks, $\mathrm{tr}(\mathscr{W}_A)$ is analytically tractable and computationally feasible to compute, and also intrinsically less conservative than other metrics that rely on the smallest eigenvalue of $\mathscr{W}_A$ (which also makes them hard to compute due to limited machine precision). We refer the reader to [12, Appendix B] for a comparison and detailed discussion of the (dis)advantages of these measures.

Our problem of interest is *improving controllability by modifying the network structure*. In doing so and for simplicity, we keep the input structure fixed. Edge modification may involve perturbing existing edges or adding new ones. We represent the modification or the sub-graph to be added by $\mathscr{G}_{\delta A} = (\mathscr{V},\mathscr{E}_{\delta A},w_{\delta A})$ with weighted adjacency matrix $\delta A = (\delta a_{ij}) \in \mathbb{R}^{n\times n}_{\geq 0}$. The resulting network is $\mathscr{G}_A \cup \mathscr{G}_{\delta A} = (\mathscr{V},\mathscr{E}_A \cup \mathscr{E}_{\delta A}, w_A + w_{\delta A})$. Both the number of modified edges and the total added weight may be constrained by bounds $N_{\max}$ and $w_{\max} \in \mathbb{R}_{>0}$, respectively. Formally, we seek to solve

$$\max_{\delta A} \qquad f(A+\delta A) \tag{3}$$

$$
\begin{aligned}
\text{s.t.} \quad & \operatorname{card}(\mathscr{E}_{\delta A}) \leq N_{\max}, \\
& \sum w_{\delta A}(\mathscr{E}_{\delta A}) \leq w_{\max},
\end{aligned}
$$

where $f(A+\delta A) = \operatorname{tr}(\mathscr{W}_{A+\delta A})$, $f(A+\delta A) = \log\det(\mathscr{W}_{A+\delta A})$, or $f(A+\delta A) = (\operatorname{tr}(\mathscr{W}_{A+\delta A}^{-1}))^{-1}$, and $\mathscr{W}_{A+\delta A}$ is the controllability Gramian of the modified network,

$$
\mathscr{W}_{A+\delta A} = \sum_{t=0}^{T-1} (A+\delta A)^t BB^\top (A+\delta A)^{t\top}.
$$

Note that the formulation (3) does not explicitly consider the stability of the resulting network. One could consider stability explicitly in the design problem by, for instance, resorting to techniques [28] for converting spectral stability constraints into algebraic ones. As our ensuing discussion shows, the study of the controllability problem presents significant technical challenges when considered on its own, and we defer for future work its consideration in combination with the stability problem.

The problem (3) is a non-convex mixed-integer nonlinear program (MINLP) which quickly becomes computationally intractable as the network size grows. The selection of $N_{\max}$ edges and their respective optimal weights requires reasoning over $2n^2$ variables, where n^2 are binary (corresponding to the edge selection) and n^2 are continuous (corresponding to the edge weight). This quadratic growth of the number of variables with the network size and the presence of integer variables prohibits the direct solution of (3) for large networks.

Our proposal instead consists of first restricting the possible edge choices to a smaller search space $\mathscr{E}_S$ with cardinality N_S such that $N_{\max} < N_S \leq n^2$. This does not change the nonconvex nature of the optimization problem but reduces the number of optimization variables to $2N_S$. The optimization problem then takes the form

$$
\begin{aligned}
\max_{\eta, w} \quad & f(A+\delta A) \\
\text{s.t.} \quad & \delta A = \sum_{k=1}^{N_S} \eta_k w_k \Delta_k, \qquad (4)\\
& \eta = \begin{pmatrix} \eta_1 & \eta_2 & \dots & \eta_k \dots & \eta_{N_S} \end{pmatrix}^\top \in \{0,1\}^{N_S}, \\
& w = \begin{pmatrix} w_1 & w_2 & \dots & w_k & \dots & w_{N_S} \end{pmatrix}^\top, \\
& 0 \leq w_k \leq w_{ub}, \quad \sum_{k=1}^{N_S} \eta_k \leq N_{\max}, \quad \sum_{k=1}^{N_S} w_k \leq w_{\max}.
\end{aligned}
$$

Here, η and w are the binary edge selection vector and the edge weight vector, respectively. Also, for each $1 \leq k \leq N_S$, if (j,i) is the corresponding element in $\mathscr{E}_S$, $\Delta_k \in \mathbb{R}^{n\times n}$ is such that $(\Delta_k)_{ij} = 1$ while all its other entries are zero. Note that if $N_S = n^2$, then (4) corresponds to (3).

In general, to simplify the computation of a solution to (4), we can take $N_S \ll n^2$. The key question then is how to select the set of candidate edges $\mathscr{E}_S$ over which the optimal search is performed. The underlying idea is that a good selection would yield an optimal solution of (4) that is a near-optimum of (3). An alternative viewpoint is that, by removing edges from consideration that do not significantly affect the objective function in the original optimization problem, the resulting optimal value is not negatively affected. Our strategy to tackle this selection it is to characterize what makes edges important in enhancing controllability and build on this understanding to perform the selection of candidate edges over which to restrict the search.

3 Energy-Transfer Edge Centrality

We introduce here a novel notion of edge centrality that seeks to quantify the influence of an edge in shaping network behavior. This notion builds on measures of the energy exchange between an individual node and the network, for which we employ the controllability and observability Gramians. We later demonstrate how the proposed notion of edge centrality can be invoked to address the network modification problem stated in Section 2.

3.1 Node-Network Interactions

As the input matrix B is fixed, so is the location of input (energy) injection from external sources. The distribution and flow of energy is therefore only dependent on the network edge structure. Due to the original structure of the network, the energy supplied may get accumulated at some nodes. At the same time, some nodes which are efficient at distributing energy may not receive adequate energy from input nodes. By modifying the network structure, one can thus redistribute the energy flow that takes place in the system. Intuitively, the modified edges should connect (or strengthen existing connections from) nodes which accumulate energy to the nodes which are in a good position to distribute it. Following the exposition of [11], we next quantify these properties by the complementary notions of the *influence of the network on a node* and the *influence of a node on the network*, respectively.

Node-to-Network Influence: To quantify the influence of a particular node on the network, we consider that particular node as the only input node to the network and compute the trace of the resulting controllability Gramian [5]. For node j and time horizon $t \in \{1,\dots,T\}$, this Gramian takes the form

$$
\mathscr{W}_j^{(t)} = \sum_{k=0}^{t-1} A^k e_j e_j^\top A^{k\top}. \qquad (5)
$$

We compute the influence of node j on the network at time t, denoted by $p_j^{(t)}$, as

$$
p_j^{(t)} = \operatorname{tr}(\mathscr{W}_j^{(t)}) = \sum_{u=1}^{n} \sum_{k=0}^{t-1} e_u^\top A^k e_j e_j^\top A^{k\top} e_u. \qquad (6)
$$

Clearly $p_j^{(t)} \geq 1$. If $p_j^{(t)} = 1$ for all $t > 1$, then that node j is a sink and thus accumulates energy, cf. [11].

Network-to-Node Influence: In parallel to the above definition, we quantify the influence of the network on a particular node by considering that particular node as the only output node of the network and computing the trace of the resulting *observability* Gramian [5]. For node i and time horizon $t \in \{1,\dots,T\}$, this Gramian takes the form

$$
\mathscr{M}_i^{(t)} = \sum_{k=0}^{t-1} A^{k\top} e_i e_i^\top A^k, \qquad (7)
$$

and we compute the influence of the network on node i at time t, denoted by $q_i^{(t)}$, as

$$
q_i^{(t)} = \operatorname{tr}(\mathscr{M}_i^{(t)}) = \sum_{v=1}^{n} \sum_{k=0}^{t-1} e_v^\top A^{k\top} e_i e_i^\top A^k e_v. \qquad (8)
$$

Similarly, $q_i^{(t)} \geq 1$ and if $q_i^{(t)} = 1$ for all $t > 1$, node i is a source [11] and distributes energy (if receiving external input).

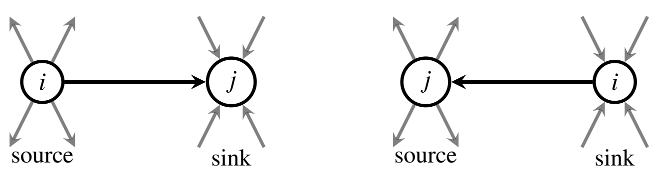
3.2 Edge Centrality Measure

Here, we combine the node-network notions above to define a novel edge centrality measure based on energy exchange. Intuitively, the influence of an edge in a network is related to the nodes it connects and the extent to which it facilities the energy distribution in the network. If an edge connects an energy-rich node (one with high q_i) to a node with high potential for facilitating energy distribution (one with high p_j), then the edge has more influence on the energy distribution in the network. Thus, we propose the energy-transfer edge centrality (ETEC)

$$c_{ij} = \sum_{t=1}^{T-1} c_{ij}^{(t)}, \tag{9a}$$

$$c_{ij}^{(t)} = q_i^{(t)} p_j^{(t)}, \tag{9b}$$

as a measure of the centrality of the edge directed from node i to node j. For simplicity, we also refer to $c_{ij}^{(t)}$ as *the edge centrality at time* t. Clearly $c_{ij}^{(t)} \geq 1$, so the minimum attainable centrality at time t is 1, which has a simple intuitive interpretation. If $c_{ij}^{(t)} = 1$ for all $t > 1$, then the corresponding edge is directed from a source node i to a sink node j, and thus constitutes a minimally-influential interconnection, cf. Figure 1(a). A highly-influential interconnection, in contrast, is one with $c_{ij}^{(t)} \gg 1$ for all $t > 1$, e.g., one connecting a sink node i to a source node j, cf. Figure 1(b). In general, there may be limited or no source/sink nodes in a network. In such situations, the concept of edge centrality can then help us to identify and compare the importance of edges in regards to the input energy distribution across the network.



(a) Edge with low centrality **(b)** Edge with high centrality

Fig. 1: Examples of edges with (a) low ($c_{ij} = 1$) and (b) high ($c_{ij} \gg 1$) centrality.

Remarkably, ETEC has a strong relationship with the objective function of the optimization problem (3). Figure 2 shows the histogram of the correlation coefficients [29] between the elements of the gradient of the trace of the Gramian (with respect to all the edge weights $\{a_{ji}\}_{i,j}$) and $\{c_{ij}\}_{i,j}$ for 10^3 Erdős-Rényi (ER) random networks with $n = 25$ nodes, $m = 8$ input nodes, 0.2 edge probability, and $T = n$. This is calculated in MATLAB using the function '*corr*', which returns the pairwise correlation coefficient between the gradient with respect to an edge weight and its corresponding edge centrality as well as the p-value for testing the hypothesis of no correlation against the alternative hypothesis of a nonzero correlation. The plot shows a remarkable average correlation coefficient of $R \simeq 0.9$, with p-value $< 10^{-6}$ in all cases, revealing an extreme statistical significance [30]. These results motivate our ensuing investigation of the relationship between both notions. Incidentally, through numerical experiments, we have also found that statistically the ETEC measure is related to the notion of non-normality of network quantified using, cf. [31], the distance from commutativity and Henrici's departure from normality, and also with the network diameter.

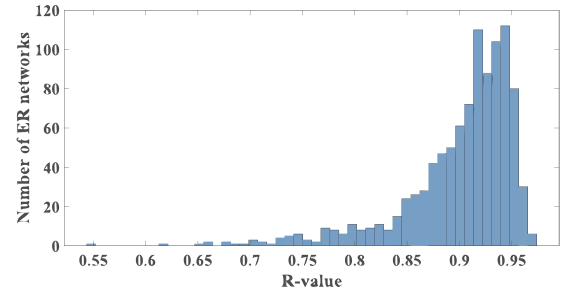


Fig. 2: Histogram of the correlation coefficient (R-value) between the values of edge centrality $\{c_{ij}\}_{i,j}$ and $\{\frac{\partial \operatorname{tr}(\mathscr{W}_A)}{\partial a_{ji}}\}_{i,j}$ for 10^3 random networks. The plot suggests that the former can be used as a proxy for the latter.

3.3 Relationship with Functions of the Gramian

In this section we establish a relationship between the proposed notion of ETEC and the gradient direction of the objective functions in (3). Consider the modification of the weight of an edge directed from node i to node j. The corresponding adjacency matrix δA only has one non-zero element, $\delta A = e_j e_i^\top$. Let,

$$\begin{aligned} C_j^{(t)} &= (A^{t-1}e_j \quad A^{t-2}e_j \quad \cdots \quad Ae_j \quad e_j), \\ O_i^{(t)} &= (e_i \quad A^\top e_i \quad \cdots \quad A^{t-2^\top} e_i \quad A^{t-1^\top} e_i)^\top. \end{aligned} \tag{10}$$

The next result provides an expression for the derivative of the Gramian with respect to edge weights in terms of the matrices $C_j^{(t)}$ and $O_i^{(t)}$.

Theorem 3.1. *(Derivative of Gramian with respect to edge weight). For the network dynamics* (1) *and any* $i, j \in \mathscr{V}$*,*

$$\frac{\partial}{\partial a_{ji}} \mathscr{W}_A = \sum_{t=1}^{T-1} \left(C_j^{(t)} O_i^{(t)} BB^\top A^{t^\top} + A^t BB^\top \left(C_j^{(t)} O_i^{(t)} \right)^\top \right),$$

where the controllability Gramian $\mathscr{W}_A$ *is defined in* (2).

Proof. From (2), consider the derivative of a general term in the expression of $\mathscr{W}_A$ with respect to the scalar a_{ji}. Using $\frac{\partial A^t}{\partial a_{ji}} = \sum_{k=0}^{t-1} A^k \frac{\partial A}{\partial a_{ji}} A^{t-1-k}$ we get,

$$\begin{aligned} \frac{\partial}{\partial a_{ji}} A^t BB^\top A^{t^\top} = \sum_{k=0}^{t-1} \Big(& A^k \frac{\partial A}{\partial a_{ji}} A^{t-1-k} BB^\top A^{t^\top} \\ & + A^t BB^\top \left(A^k \frac{\partial A}{\partial a_{ji}} A^{t-1-k} \right)^\top \Big). \end{aligned} \tag{11}$$

Now $\frac{\partial A}{\partial a_{ji}}$ is a matrix with 1 as the $(j,i)^{th}$ element and rest all 0s. So, using $\frac{\partial A}{\partial a_{ji}} = e_j {e_i}^\top$ and

$$\sum_{k=0}^{t-1} A^k \frac{\partial A}{\partial a_{ji}} A^{t-k-1} = \sum_{k=0}^{t-1} A^k e_j e_i^\top A^{t-k-1} = C_j^{(t)} O_i^{(t)},$$

where $C_j^{(t)}$ and $O_i^{(t)}$ are defined in (10) yields

$$\frac{\partial}{\partial a_{ji}} A^t BB^\top A^{t^\top} = \left(C_j^{(t)} O_i^{(t)} BB^\top A^{t^\top} + A^t BB^\top \left(C_j^{(t)} O_i^{(t)} \right)^\top \right).$$

The result follows by summing up $\frac{\partial}{\partial a_{ji}} A^t BB^\top A^{t^\top}$ above from $t = 1$ to $t = T - 1$. $\square$

We next rely on Theorem 3.1 to derive expressions for the gradient of the trace, $\log\det$, and inverse of the trace inverse of the Gramian.

Corollary 3.2. *(Gradient of the trace, $\log det$ and inverse of the trace inverse of the Gramian with respect to edge weight). For the network dynamics* (1)*, any $i,j \in \mathscr{V}$,*

1) $\frac{\partial}{\partial a_{ji}} \operatorname{tr}(\mathscr{W}_A) = 2\sum_{t=1}^{T-1} \operatorname{tr}\left(BB^\top A^{t\top} C_j^{(t)} O_i^{(t)}\right)$,
2) $\frac{\partial}{\partial a_{ji}} \log det(\mathscr{W}_A) = 2\sum_{t=1}^{T-1} \operatorname{tr}\left(BB^\top A^{t\top} \mathscr{W}_A^{-1} C_j^{(t)} O_i^{(t)}\right)$,
3) $\frac{\partial}{\partial a_{ji}} (\operatorname{tr}(\mathscr{W}_A^{-1}))^{-1} = \frac{2}{(\operatorname{tr}(\mathscr{W}_A^{-1}))^2} \sum_{t=1}^{T-1} \operatorname{tr}\left(BB^\top A^{t\top} \mathscr{W}_A^{-2} C_j^{(t)} O_i^{(t)}\right)$.

Proof. The proof of 1) follows from taking the trace of the expression $\frac{\partial}{\partial a_{ji}} \mathscr{W}_A$ in Theorem 3.1 and using the facts $\operatorname{tr}(\frac{\partial U}{\partial a_{ji}}) = \frac{\partial}{\partial a_{ji}} \operatorname{tr}(U)$; $\operatorname{tr}(U_1U_2U_3) = \operatorname{tr}(U_2U_3U_1)$; and $\operatorname{tr}(U_1U_2^\top) = \operatorname{tr}(U_1^\top U_2)$.

To establish 2), we rely on the following equality, cf. [32, Appendix A],

$$\frac{\partial}{\partial a_{ji}} \log\det(\mathscr{W}_A) = \operatorname{tr}\left(\mathscr{W}_A^{-1} \frac{\partial \mathscr{W}_A}{\partial a_{ji}}\right).$$

The result now follows from combining Theorem 3.1, $\operatorname{tr}(\sum_i U_i) = \sum_i \operatorname{tr}(U_i)$, $\operatorname{tr}(U_1^\top U_2) = \operatorname{tr}(U_1U_2^\top)$, and $\operatorname{tr}(U_1U_2U_3) = \operatorname{tr}(U_3U_1U_2)$.

To establish 3), we note that $\frac{\partial}{\partial a_{ji}} (\operatorname{tr}(\mathscr{W}_A^{-1}))^{-1} = \frac{-1}{(\operatorname{tr}(\mathscr{W}_A^{-1}))^2} \frac{\partial}{\partial a_{ji}} \operatorname{tr}(\mathscr{W}_A^{-1})$. The result follows by combining

$$\frac{\partial}{\partial a_{ji}} \operatorname{tr}(\mathscr{W}_A^{-1}) = -\operatorname{tr}\left(\mathscr{W}_A^{-1} \frac{\partial \mathscr{W}_A}{\partial a_{ji}} \mathscr{W}_A^{-1}\right),$$

cf. [32, Appendix A] with the properties of the trace invoked to establish 2). ☐

Corollary 3.2 establishes an important relationship between the gradient of the trace of the controllability Gramian of the original network with input matrix B, the controllability matrix of the network with node j as the only input node, and the observability matrix of the network with node i as the only output node. It is worth noticing that in the case where $A = -L$, where L is a Laplacian matrix, then Part 1 in Corollary 3.2 corresponds to the centrality measure proposed in [27] for the no-delay case. Corollary 3.2 serves as a basis for establishing a relationship between the proposed ETEC measure and the gradient of the trace of Gramian, as we show next.

For any edge (i,j) and $t \in \{1,\dots,T-1\}$, let

$$g_{ij}^{(t)} = \operatorname{tr}\left(BB^\top A^{t\top} C_j^{(t)} O_i^{(t)}\right) \tag{12}$$

so that, according to Corollary 3.2, the gradient can be compactly expressed as $\frac{\partial}{\partial a_{ji}} \operatorname{tr}(\mathscr{W}_A) = 2\sum_{t=1}^{T-1} g_{ij}^{(t)}$. The matrix $C_j^{(t)} O_i^{(t)}$ in (12) is of size $n \times n$, and its $(u,v)^{th}$ element is given by,

$$\left[C_j^{(t)} O_i^{(t)}\right]_{uv} = e_u^\top \left(\sum_{k=0}^{t-1} A^{t-1-k} e_j e_i^\top A^k\right) e_v. \tag{13}$$

The following results present two independent sets of bounds relating $g_{ij}^{(t)}$ with the ETEC $c_{ij}^{(t)}$ at time $t = 1,\dots,T-1$.

Theorem 3.3. *(ETEC-based bounds on the gradient of trace of Gramian (I)). For any $t \in \{1,\dots,T-1\}$,*

$$0 \le g_{ij}^{(t)} \le \|A^t BB^\top\|_F \sqrt{c_{ij}^{(t)}}. \tag{14}$$

Proof. From the definition (12) of $g_{ij}^{(t)}$ and A being a non-negative weighted adjacency matrix, it is clear that $g_{ij}^{(t)} \ge 0$. Using the definition of the trace of the product of two matrices,

$$\begin{aligned} \operatorname{tr}(BB^\top A^{t\top} C_j^{(t)} O_i^{(t)}) &= \sum_{u=1}^{n}\sum_{v=1}^{n} \left[A^t BB^\top\right]_{uv} \left[C_j^{(t)} O_i^{(t)}\right]_{uv}, \\ &\le \|A^t BB^\top\|_F \|C_j^{(t)} O_i^{(t)}\|_F. \end{aligned} \tag{15}$$

Using the definition of Frobenius norm [33],

$$\begin{aligned} \|C_j^{(t)} O_i^{(t)}\|_F^2 &= \sum_{u=1}^{n}\sum_{v=1}^{n} \left[C_j^{(t)} O_i^{(t)}\right]_{uv}^2, \\ &= \sum_{u=1}^{n}\sum_{v=1}^{n} \left[e_u^\top \left(\sum_{k=0}^{t-1} A^{t-1-k} e_j e_i^\top A^k\right) e_v\right]^2. \end{aligned} \tag{16}$$

Consider

$$\begin{aligned} &\left[e_u^\top \left(\sum_{k=0}^{t-1} A^{t-1-k} e_j e_i^\top A^k\right) e_v\right]^2 \\ &\qquad = \left[\sum_{k=0}^{t-1} \left(e_u^\top A^{t-1-k} e_j\right)\left(e_i^\top A^k e_v\right)\right]^2 \\ &\qquad \le \sum_{k=0}^{t-1} \left(e_u^\top A^{t-1-k} e_j\right)^2 \sum_{k=0}^{t-1} \left(e_i^\top A^k e_v\right)^2. \end{aligned}$$

Summing from $u = 1$ to $u = n$ and from $v = 1$ to $v = n$ in this expression, and using

$$\sum_{u=1}^{n}\sum_{k=0}^{t-1} \left(e_u^\top A^k e_j\right)^2 = \sum_{u=1}^{n}\sum_{k=0}^{t-1} e_u^\top A^k e_j e_j^\top A^{k\top} e_u = p_j^{(t)},$$

$$\sum_{v=1}^{n}\sum_{k=0}^{t-1} \left(e_i^\top A^k e_v\right)^2 = \sum_{v=1}^{n}\sum_{k=0}^{t-1} e_v^\top A^{k\top} e_i e_i^\top A^k e_v = q_i^{(t)},$$

gives $\|C_j^{(t)} O_i^{(t)}\|_F \le \sqrt{p_j^{(t)} q_i^{(t)}} = \sqrt{c_{ij}^{(t)}}$, which together with (15) implies the result. ☐

Theorem 3.3 provides analytical evidence of the high correlation between the gradient of $\operatorname{tr}(\mathscr{W}_A)$ and the centrality measured c_{ij} observed in Figure 2. We build on this relationship later in Section 5 in our algorithm design for optimal edge selection.

In general, one cannot discard that the upper bound obtained in Theorem 3.3 provides a loose approximation to $g_{ij}^{(t)}$ because the lower bound does not scale with ETEC. Interestingly, the edge centrality measure c_{ij} is not only related to the gradient of $\operatorname{tr}(\mathscr{W}_A)$ but also to the gradient of $\log\det(\mathscr{W}_A)$ and $(\operatorname{tr}(\mathscr{W}_A^{-1}))^{-1}$, as we show next. For any edge (i,j) and $t \in \{1,\dots,T-1\}$, let

$$h_{ij}^{(t)} = \operatorname{tr}\left(BB^\top A^{t\top} P C_j^{(t)} O_i^{(t)}\right) \tag{17}$$

so that, according to Corollary 3.2, the gradient can be compactly expressed as $\frac{\partial f}{\partial a_{ji}} = 2\sum_{t=1}^{T-1} h_{ij}^{(t)}$ with $P = I$ for $f = \operatorname{tr}(\mathscr{W}_A)$, $P = \mathscr{W}_A^{-1}$ for $f = \log\det(\mathscr{W}_A)$ and $P = \frac{\mathscr{W}_A^{-2}}{(\operatorname{tr}(\mathscr{W}_A^{-1}))^2}$ for $f = (\operatorname{tr}(\mathscr{W}_A^{-1}))^{-1}$. Note that for $P = I$, $h_{ij}^{(t)} = g_{ij}^{(t)}$. Let $\Omega_{ij}^{(t)}$ be the projection of $PA^t BB^\top$ on $C_j^{(t)} O_i^{(t)}$ denoted as,

$$\Omega_{ij}^{(t)} = \|PA^t BB^\top\|_F \cos\omega_{ij}^{(t)}.$$

with the projection angle $\omega_{ij}^{(t)} = \angle(\mathrm{vec}(PA^tBB^\top), \mathrm{vec}(C_j^{(t)}O_i^{(t)}))$. In addition, let ${s_j^{(t)}}^2$ be the variance of the eigenvalues of ${C_j^{(t)}}^\top C_j^{(t)}$ given by,

$$s_j^{(t)} = \sqrt{\frac{\mathrm{tr}\big(({C_j^{(t)}}^\top C_j^{(t)})^2\big)}{r_j^{(t)}} - \left(\frac{p_j^{(t)}}{r_j^{(t)}}\right)^2}, \quad r_j^{(t)} = \mathrm{rank}({C_j^{(t)}}^\top C_j^{(t)}). \tag{18}$$

We are now ready to state our next result.

Theorem 3.4. *(ETEC-based bounds on the gradient square of trace,* $\log det$ *and inverse of the trace inverse of Gramian). For* $P = I$ *for* $f = \mathrm{tr}(\mathscr{W}_A)$, $P = \mathscr{W}_A^{-1}$ *for* $f = \log det(\mathscr{W}_A)$, *and* $P = \frac{\mathscr{W}_A^{-2}}{(\mathrm{tr}(\mathscr{W}_A^{-1}))^2}$ *for* $f = (\mathrm{tr}(\mathscr{W}_A^{-1}))^{-1}$, *and any* $t \in \{1, \ldots, T-1\}$,

$$\underline{\mathfrak{h}}_{ij}^{(t)} \le {h_{ij}^{(t)}}^2 \le \overline{\mathfrak{h}}_{ij}^{(t)}, \tag{19}$$

where,

$$\underline{\mathfrak{h}}_{ij}^{(t)} = {\Omega_{ij}^{(t)}}^2 \left(\frac{c_{ij}^{(t)}}{r_j^{(t)}} - q_i^{(t)} s_j^{(t)} \sqrt{r_j^{(t)} - 1}\right),$$

$$\overline{\mathfrak{h}}_{ij}^{(t)} = {\Omega_{ij}^{(t)}}^2 \left(\frac{c_{ij}^{(t)}}{r_j^{(t)}} + q_i^{(t)} s_j^{(t)} \sqrt{r_j^{(t)} - 1}\right).$$

Proof. Using the definitions of $h_{ij}^{(t)}$, $\omega_{ij}^{(t)}$ and $\Omega_{ij}^{(t)}$, we observe that

$$\begin{aligned} h_{ij}^{(t)} &= \|PA^tBB^\top\|_F \|C_j^{(t)}O_i^{(t)}\|_F \cos\omega_{ij}^{(t)} \\ &= \Omega_{ij}^{(t)} \|C_j^{(t)}O_i^{(t)}\|_F. \end{aligned}$$

Using the definition of Frobenius norm [33],

$$\|C_j^{(t)}O_i^{(t)}\|_F = \sqrt{\mathrm{tr}\big({C_j^{(t)}}^\top C_j^{(t)} O_i^{(t)} {O_i^{(t)}}^\top\big)}. \tag{20}$$

From [34, Lemma 1], the trace of the product in the latter term can be lower and upper bounded as

$$\begin{aligned} \lambda_{\min}\big({C_j^{(t)}}^\top C_j^{(t)}\big)\mathrm{tr}\big(O_i^{(t)}{O_i^{(t)}}^\top\big) &\le \mathrm{tr}\big({C_j^{(t)}}^\top C_j^{(t)} O_i^{(t)} {O_i^{(t)}}^\top\big) \\ &\le \lambda_{\max}\big({C_j^{(t)}}^\top C_j^{(t)}\big)\mathrm{tr}\big(O_i^{(t)}{O_i^{(t)}}^\top\big), \end{aligned} \tag{21}$$

where $\lambda_{\min}(\cdot)$ and $\lambda_{\max}(\cdot)$ denote the smallest and largest eigenvalues of a symmetric matrix, respectively. Using

$$\mathrm{tr}\big({C_j^{(t)}}^\top C_j^{(t)}\big) = \sum_{u=1}^{n}\sum_{k=0}^{t-1} e_u^\top A^k e_j e_j^\top {A^k}^\top e_u = p_j^{(t)},$$

and [35, Theorem 2.1] and (18), we further have the bounds

$$\frac{p_j^{(t)}}{r_j^{(t)}} - s_j^{(t)}\sqrt{r_j^{(t)} - 1} \le \lambda_{\min}\big({C_j^{(t)}}^\top C_j^{(t)}\big),$$

$$\lambda_{\max}\big({C_j^{(t)}}^\top C_j^{(t)}\big) \le \frac{p_j^{(t)}}{r_j^{(t)}} + s_j^{(t)}\sqrt{r_j^{(t)} - 1}. \tag{22}$$

Now using (21), (22), and the fact that

$$\mathrm{tr}\big(O_i^{(t)}{O_i^{(t)}}^\top\big) = \sum_{v=1}^{n}\sum_{k=0}^{t-1} e_v^\top {A^k}^\top e_i e_i^\top A^k e_v = q_i^{(t)},$$

we get

$$(\Omega_{ij}^{(t)})^2 \left(\frac{c_{ij}^{(t)}}{r_j^{(t)}} - q_i^{(t)} s_j^{(t)} \sqrt{r_j^{(t)} - 1}\right) \tag{23}$$

$$\begin{aligned} &\le \big(\mathrm{tr}\big(BB^\top {A^t}^\top P C_j^{(t)} O_i^{(t)}\big)\big)^2 \\ &\le (\Omega_{ij}^{(t)})^2 \left(\frac{c_{ij}^{(t)}}{r_j^{(t)}} + q_i^{(t)} s_j^{(t)} \sqrt{r_j^{(t)} - 1}\right), \end{aligned}$$

from which the result follows. □

Theorem 3.4 shows that each ${h_{ij}^{(t)}}^2$ belongs to an interval whose lower and upper bounds are functions of $c_{ij}^{(t)}$ and their average,

$${\Omega_{ij}^{(t)}}^2 \frac{c_{ij}^{(t)}}{r_j^{(t)}}, \tag{24}$$

is again directly proportional to ETEC. However, an important difference is given by the fact that the elements of the matrices $\mathscr{W}_A$ and $BB^\top {A^t}^\top C_j^{(t)} O_i^{(t)}$ are non-negative while the elements of the matrix $\mathscr{W}_A^{-1}$ may be negative. Hence, unlike $g_{ij}^{(t)}$ in (12), which is always non-negative, $h_{ij}^{(t)}$ may be negative when considering the $\log\det$ or the inverse of the trace inverse of the Gramian and, in such cases, a large $c_{ij}^{(t)}$ may not always correspond to a large $h_{ij}^{(t)}$.

We note that the lower bound in Theorem 3.4 may become negative as t increases, in which case it would be possible for $c_{ij}^{(t)}$ to be large while $h_{ij}^{(t)}$ is not, even in the case of the trace of the Gramian. Understanding whether this possibility can actually occur is an open problem, but we note that Theorem 3.4 provides an analytical basis to understand the tight correlation, cf. Figure 2, between ETEC and the gradient of the trace of Gramian observed in simulation.

Also, note that for the case of the trace of the Gramian, the upper bound in Theorem 3.4 is better than the upper bound in Theorem 3.3. In fact, we have

$$\frac{\overline{\mathfrak{h}}_{ij}^{(t)}}{\|A^tBB^\top\|_F^2 c_{ij}^{(t)}} = \cos^2\omega_{ij}^{(t)} \left(\frac{1}{r_j^{(t)}} + \frac{s_j^{(t)}}{p_j^{(t)}}\sqrt{r_j^{(t)} - 1}\right).$$

As $p_j^{(t)}$ and ${s_j^{(t)}}^2$ are the sum and variance of the eigenvalues of ${C_j^{(t)}}^\top C_j^{(t)}$, respectively, we have $\frac{s_j^{(t)}}{p_j^{(t)}} \le \frac{1}{r_j^{(t)}}\sqrt{r_j^{(t)} - 1}$, and therefore,

$$\frac{\overline{\mathfrak{h}}_{ij}^{(t)}}{\|A^tBB^\top\|_F^2 c_{ij}^{(t)}} \le \cos^2\omega_{ij}^{(t)} \le 1.$$

3.4 Efficient Computation of Edge Centrality

In this section, we show that computing the proposed edge centrality measure c_{ij} offers a significant reduction in computational complexity with respect to the computation of the gradient of the trace or $\log\det$ of Gramian in the edge modification problem (3). In particular, we show how the redundancy in the calculations of the network-node influences can be exploited to enhance the computational efficiency of calculating ETEC, and compare this to the computation of the gradients $\frac{\partial}{\partial a_{ji}}\mathrm{tr}(\mathscr{W}_A)$, $\frac{\partial}{\partial a_{ji}}\log\det(\mathscr{W}_A)$ and $\frac{\partial}{\partial a_{ji}}(\mathrm{tr}(\mathscr{W}_A^{-1}))^{-1}$. Let $\widehat{p}_j^{(t)}$ be the vector of the diagonal elements of $\mathscr{W}_j^{(t)}$ in (5) and

$$\mathscr{H}^{(t)} = \begin{pmatrix} \widehat{p}_1^{(t)} & \ldots & \widehat{p}_j^{(t)} & \ldots & \widehat{p}_n^{(t)} \end{pmatrix} \in \mathbb{R}^{n \times n}. \tag{25}$$

The following result shows how to employ this matrix to compute the network-node influences in the definition of ETEC.

Proposition 3.5. *(Network-node influences are column and row sums of $\mathscr{H}^{(t)}$). Let $\mathscr{H}^{(t)}$ be defined as in (25). Then,*

$$p_j^{(t)} = \sum_{i=1}^{n} \mathscr{H}_{ij}^{(t)}, \tag{26a}$$

$$q_i^{(t)} = \sum_{j=1}^{n} \mathscr{H}_{ij}^{(t)}. \tag{26b}$$

Proof. $\mathscr{H}_{ij}^{(t)}$ is the i^{th} diagonal element of $\mathscr{W}_j^{(t)}$, i.e.,

$$\mathscr{H}_{ij}^{(t)} = \sum_{k=0}^{t-1} e_i^\top A^k e_j e_j^\top A^{k\top} e_i,$$

and thus (26a) follows from the definition of $p_j^{(t)}$ in (6). Similarly, since $\mathscr{H}_{ij}^{(t)}$ is a scalar,

$$\mathscr{H}_{ij}^{(t)} = \sum_{k=0}^{t-1} e_i^\top A^k e_j e_j^\top A^{k\top} e_i = \sum_{k=0}^{t-1} e_j^\top A^{k\top} e_i e_i^\top A^k e_j,$$

and so (26b) follows by noting that

$$\sum_{j=1}^{n} \mathscr{H}_{ij}^{(t)} = \sum_{j=1}^{n} \sum_{k=0}^{t-1} e_j^\top A^{k\top} e_i e_i^\top A^k e_j = q_i^{(t)},$$

using (8). □

Based on this result, we can simplify the computation of the centrality measures $\{c_{ij}\}_{i,j}$ as follows. Define the vectors of network-node influences,

$$p^{(t)} = \begin{pmatrix} p_1^{(t)} & p_2^{(t)} & \ldots & p_j^{(t)} & \ldots & p_{n-1}^{(t)} & p_n^{(t)} \end{pmatrix}^\top,$$
$$q^{(t)} = \begin{pmatrix} q_1^{(t)} & q_2^{(t)} & \ldots & q_i^{(t)} & \ldots & q_{n-1}^{(t)} & q_n^{(t)} \end{pmatrix}^\top.$$

Using Proposition 3.5, both $p^{(t)}$ and $q^{(t)}$ can be computed from $\mathscr{H}^{(t)}$. We define $\Theta^{(t)} = p^{(t)} q^{(t)\top} \in \mathbb{R}^{n\times n}$ and

$$\Theta = \sum_{t=1}^{T-1} \Theta^{(t)} \in \mathbb{R}^{n\times n}. \tag{27}$$

Then, from (9) and (27), it is straightforward to see that

$$c_{ij}^{(t)} = q_i^{(t)} p_j^{(t)} = \Theta_{ji}^{(t)} \quad \text{and} \quad c_{ij} = \sum_{t=1}^{T-1} q_i^{(t)} p_j^{(t)} = \Theta_{ji}.$$

This discussion, together with the definition of $\mathscr{H}^{(t)}$, shows that the knowledge of $\{\widehat{p}_j^{(t)}\}_{j=1}^n$ is sufficient to compute $\Theta^{(t)}$. We are now ready to substantiate our claim of computational efficiency of ETEC versus the gradients of the objective functions:

Computation of gradients of objective functions: the computation of $\frac{\partial}{\partial a_{ji}} \operatorname{tr}(\mathscr{W}_A)$, $\frac{\partial}{\partial a_{ji}} \log\det(\mathscr{W}_A)$ and $\frac{\partial}{\partial a_{ji}} (\operatorname{tr}(\mathscr{W}_A^{-1}))^{-1}$ requires $g_{ij}^{(t)}$ and $h_{ij}^{(t)}$ respectively. The latter can be efficiently computed using their definitions (12) and (17), respectively, as discussed next. At each time t and for each edge (i,j), we have A, $A^2, \ldots, A^{t-1}$, $e_j, Ae_j, \ldots A^{t-2}e_j$, $e_i{}^\top, e_i{}^\top A, \ldots, e_i{}^\top A^{t-2}$ and $\mathscr{W}_A^{-1}$ available from previous times $1,2,\ldots,t-1$. To compute $g_{ij}^{(t)}$ or $h_{ij}^{(t)}$, we need to additionally compute A^t, $A^{t-1}e_j$, and $e_i^\top A^{t-1}$. With all this information, we then form $C_j^{(t)}$, $O_i^{(t)}$, compute the products $BB^\top A^{t\top} C_j^{(t)} O_i^{(t)}$ and $BB^\top A^{t\top} P C_j^{(t)} O_i^{(t)}$, and take their trace. Note that this computation needs to be done for n^2 edges.

Computation of ETEC: To calculate $c_{ij}^{(t)}$, we first need to compute $\widehat{p}_j^{(t)}$ for each node j. This computation requires to form the Gramian $\sum_{k=0}^{t} A^k e_j e_j^\top A^{k\top}$. To do this, we only need the term $A^t e_j e_j^\top A^{t\top}$, as $\sum_{k=0}^{t-1} A^k e_j e_j^\top A^{k\top}$ is available from the previous times. Once $\widehat{p}_j^{(t)}$ are computed for each node, we form the matrix $\mathscr{H}^{(t)}$, sum its rows and columns to form vectors $p^{(t)}, q^{(t)}$ and then perform the outer product $p^{(t)} q^{(t)\top}$ to get $\Theta^{(t)}$. Note that this computation needs to be done for n nodes.

The above discussion shows that the computational effort in calculating $c_{ij}^{(t)}$ is far less than what is required to calculate $g_{ij}^{(t)}$ and $h_{ij}^{(t)}$, respectively.

4 Properties of the Metrics and Implications for Optimal Edge Modification

The notion of ETEC introduced in Section 3 helps identify those edges that have a significant effect on enhancing network controllability. In this section we characterize properties of the solution of the optimization problems (3) and (4) that will later guide our design of computationally efficient algorithms for network edge modification. We first analyze the case of trace of the Gramian followed by $\log\det$ and inverse of the trace inverse of the Gramian.

4.1 Trace of Gramian as Objective

Here we prove that the optimal solution of either (3) or (4), when the objective function is the trace of Gramian, lies on the boundary of its feasibility set. To establish this, we rely on the non-decreasing nature of the trace of Gramian and its gradient, shown next. Throughout this section, $f(A) = \operatorname{tr}(\mathscr{W}_A)$.

Lemma 4.1. *(Trace of Gramian is non-decreasing). Let A_1 and A_2 be non-negative weighted adjacency matrices. If $A_2 - A_1 \geq 0$ then $f(A_2) \geq f(A_1)$.*

Proof. Let $\delta A = A_2 - A_1$ and note that δA is a non-negative matrix. From (2),

$$\begin{aligned}\mathscr{W}_{A_2} &= \sum_{t=0}^{T-1} A_2^t BB^\top A_2^{t\top} \\ &= \sum_{t=0}^{T-1} (A_1 + \delta A)^t BB^\top (A_1 + \delta A)^{t\top} \\ &= \sum_{t=0}^{T-1} A_1^t BB^\top A_1^{t\top} + \Psi(A_1, \delta A),\end{aligned}$$

where $\Psi(A_1, \delta A)$ is an appropriate matrix function of A_1 and δA consisting of the remaining terms of the Gramian expansion at A_2. Therefore, $f(A_2) = f(A_1) + \operatorname{tr}(\Psi(A_1, \delta A))$. The non-negativeness of $A_1, \delta A$ and the input matrix B implies $f(A_2) \geq f(A_1)$. □

From Lemma 4.1 it follows that with the addition or enhancement of each new edge, the trace of the controllability Gramian can only increase. Next we show that the gradient of the trace of the controllability Gramian with respect to any edge is also a non-decreasing function.

Lemma 4.2. *(Gradient of trace of Gramian is non-decreasing). Let A_1 and A_2 be non-negative weighted adjacency matrices. If $A_2 - A_1 \geq 0$, then $\frac{\partial}{\partial a_{ji}} f(A_2) \geq \frac{\partial}{\partial a_{ji}} f(A_1)$ for any $(i,j) \in \mathscr{V} \times \mathscr{V}$.*

Proof. From (11) in the proof of Theorem 3.1, we deduce

$$\frac{\partial}{\partial a_{ji}} f(A) = 2 \sum_{t=1}^{T-1} \operatorname{tr}\Big(\sum_{p=0}^{t-1} \big(A^p \Delta_{ji} A^{t-p-1} B B^\top A^{t\top} \big) \Big),$$

where $\Delta_{ji} = \frac{\partial A}{\partial a_{ji}}$ is a matrix whose $(j,i)^{th}$ element is 1 and all other elements are 0. Evaluating the expression above for $\frac{\partial}{\partial a_{ji}} f(A)$ at $A = A_2$ and noting that $A_2 = A_1 + \delta A$, we can write

$$\frac{\partial}{\partial a_{ji}} f(A_2) = \frac{\partial}{\partial a_{ji}} f(A_1) + \operatorname{tr}(\Theta(A_1, \delta A)),$$

where $\Theta(A_1, \delta A)$ is an appropriate matrix function of A_1 and δA consisting of the remaining terms of the gradient expansion at A_2. Using the non-negativeness of $A_1, \delta A$ and B, we deduce that $\operatorname{tr}(\Theta(A_1, \delta A)) \geq 0$, which implies the result. □

Next, we turn our attention to characterize the optimal solution of (3) and (4). Note that we can break the edge addition process in two parts: (i) the selection of $N_{\max}$ edge locations and (ii) the computation of the weights to be added to the $N_{\max}$ selected edges. Let $\mathscr{E}_{opt} \subset \mathscr{E}_S$ be the set of optimal edge locations of cardinality $N_{\max}$ solving (i). Fixing $\mathscr{E}_{opt}$, the optimal weight computation in (ii) becomes

$$\max_{w} \quad f(A + \delta A) = \operatorname{tr}(\mathscr{W}_{A+\delta A}), \tag{28a}$$

$$\text{s.t.} \quad \delta A = \sum_{k=1}^{N_{\max}} w_k \Delta_k, \tag{28b}$$

$$w = \begin{pmatrix} w_1 & w_2 & \dots & w_{N_{\max}} \end{pmatrix}^\top, \tag{28c}$$

$$0 \leq w_k \leq w_{ub}, \quad k = 1, \dots, N_{\max}, \tag{28d}$$

$$\sum_{k=1}^{N_{\max}} w_k \leq w_{\max}. \tag{28e}$$

The next result characterizes the solutions of this problem.

Proposition 4.3. *(Characterization of optimal edge weight modification for trace). Given the edge addition set $\mathscr{E}_{opt}$ of cardinality $N_{\max}$, consider the optimization problem* (28)*. Then,*

(i) there always exists a global maximum w^* of (28) *satisfying*

$$\sum_{k=1}^{N_{\max}} w_k^* = w_{\max}, \tag{29}$$

when $w_{\max} \leq N_{\max} w_{ub}$; otherwise $w_k^* = w_{ub}$ for all k;

(ii) if, further, f is not constant as a function of w_i for any $i = 1, \dots, N_{\max}$, then any *global maximum of* (29) *satisfies* (29)*. In particular, no global maximum may lie in the interior of the feasibility set.*

Proof. (i) Let us first consider $w_{\max} \leq N_{\max} w_{ub}$. If a global optimum w^* of (28) does not already satisfy (29), then

$$\sum_{k=1}^{N_{\max}} w_k^* < w_{\max}. \tag{30}$$

However, it is trivial to find $\hat{w} \geq w^*$ satisfying (29). By Lemma 4.1 and using the fact that w^* is a global maximum, it follows that $\hat{w}$ is also a global maximum. For the case $w_{\max} > N_{\max} w_{ub}$ due to Lemma 4.1, it is trivial to see that each w_k^* takes the value w_{ub}.

(ii) We prove the result by contradiction. Assume w^* is a global maximum satisfying (30) and let $N_{w^*}^s > 0$ be the number of components of w^* that equal w_{ub}. Define

$$U_{w^*} = \{ w \geq w^* \mid w \leq w_{ub} \mathbf{1}_{N_{\max}}, \ \mathbf{1}_{N_{\max}}^T w \leq w_{\max} \}.$$

Note that U_{w^*} has a positive Lebesgue measure in $\mathbb{R}^{N_{\max} - N_{w^*}^s}$. Moreover, f is constant on U_{w^*} by Lemma 4.1 and the fact that w^* is a global maximum. Therefore f is a constant function over $\mathbb{R}^{N_{\max} - N_{w^*}^s}$, cf. [36, §4.1], which is a contradiction. Therefore, any global maximum satisfies (29). □

Proposition 4.3 means that the solutions of (4) have optimal edge weights lying on the boundary of the feasible weight set. A similar result also holds for the original optimization problem (3). Proposition 4.3, however, does not determine what happens on the boundary (29): in fact, the global optima may lie at a corner point of the boundary or in its relative interior. Both scenarios may happen, cf. Figure 3, although corner points seem to be predominant. Finally, the following result gives a sufficient (but not necessary) condition for the trace of the Gramian not to be constant as a function of any present or non-present edge weight in the network.

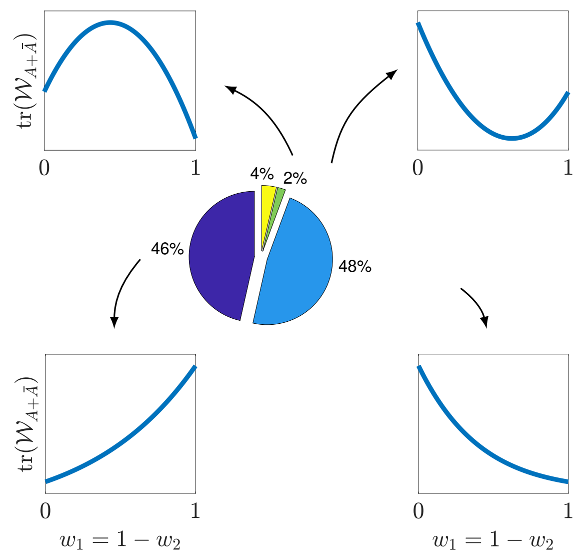


Fig. 3: Dependence patterns of $\operatorname{tr}(\mathscr{W}_{A+\delta A})$ on edge weights when the constraint (28e) holds with equality. We constructed 10^4 Erdős-Rényi random graphs with $n = 25$, $m = 5$, 20% edge probability, and $T = 2n$ and, for each graph, chose two non-self-loop edges at random and solved problem (28) with $N_{\max} = 2$, and $w_{ub} = w_{\max} = 1$ to specifically study the trade-off between w_1 and w_2. As shown in Proposition 4.3, we know that the optimal solution satisfies $w_1 + w_2 = 1$, so we classified the networks according to the pattern of $\operatorname{tr}(\mathscr{W}_{A+\delta A})$ on the edge of the feasibility set where $w_1 + w_2 = 1$ into four classes: monotonically increasing, monotonically decreasing, local minimum in the interior, and local maximum in the interior. In particular, we see that in 96% of the cases the globally optimal solution lies at a corner point of the feasibility set.

Lemma 4.4. *(Non-trivial dependence of trace of Gramian on edge weights). The function f in* (28) *is not constant as a function of any entry of δA if (A, B) is controllable and $T - 2$ is at least as large as the input depth of the network*[1].

1. Similar to the notion of the depth of a directed tree [37], the input depth of the network is defined as $\max_{j \in \mathscr{V}} \min_{i \in \mathscr{V}_B} \operatorname{dist}(i, j)$, where $\mathscr{V}_B$ is the set of input nodes of the network and $\operatorname{dist}(i, j)$ is the topological (unweighted) distance from node i to node j.

Proof. Expanding the definition of the trace of the controllability Gramian, we have

$$f(A+\delta A)=\sum_{t=0}^{T-1}\sum_{\ell=1}^{n}\sum_{i=1}^{n}\sum_{r=1}^{n}B_{i\ell}^2((A+\delta A)^t)_{ri}^2+\tilde{f}(A+\delta A)$$

where $\tilde{f}(A+\delta A)$ only contains sums and products of positive terms. Therefore, the derivative with respect to any edge weight of both the first term on the right-hand side and $\tilde{f}$ is always nonnegative. From this, we deduce that the result holds if the first term is not a constant function of any entry of δA (since the derivative of $\tilde{f}$ cannot cancel out the derivative of the first term). The term $B_{i\ell}$ equals the input weight from input $u_\ell(t)$ to node i while $((A+\delta A)^t)_{ri}$ equals the sum of the weight of all directed paths of length t from node i to node r (with the weight of a path being equal to the product of the weights of the edges along it). Note that all of the terms inside the sum also consist of sums and products of positive values and, hence, the result would hold if at least one $B_{i\ell}^2((A+\delta A)^t)_{ri}^2$ is not constant as a function of any entry of δA.

Consider an arbitrary entry (r,s) of δA, whether already present in A or not. Recall that this indicates an edge from node r to node s. Since the network is controllable, there is at least one path from one input node i (receiving some input signal $u_\ell(t)$) to r. The length of this path can be at most equal to the input depth of the network, so appending it with the edge (r,s) creates a path with length $t\le T-1$. Therefore, the term $B_{i\ell}^2((A+\delta A)^t)_{ri}^2$ for these values of t,ℓ,i,r is non zero and, thus, is not a constant function of $(\delta A)_{rs}$, completing the proof. □

Note that, if T is small, it is possible that for a controllable pair (A,B), the trace of the Gramian remains constant as a function of some entries of δA. We conclude this section by establishing an important property of the trace of the Gramian when viewed as a set function. Given a node set $\mathscr{V}$, for an arbitrary matrix A with non-negative edge weights, let $\mathscr{G}_A=(\mathscr{V},\mathscr{E}_A,w_A)$ denote the associated network triplet. Let $\mathscr{Z}_\mathscr{G}$ denote the set of all possible directed network triplets $\mathscr{G}_A=(\mathscr{V},\mathscr{E}_A,w_A)$. Given adjacency matrices A_1 and A_2 with corresponding directed network triplets $\mathscr{G}_{A_1}$ and $\mathscr{G}_{A_2}$, we note that $\mathscr{G}_{A_1}\cup\mathscr{G}_{A_2}$ is the network triplet associated to A_1+A_2. We also say, cf. [18],

$$\mathscr{G}_{A_1}\subset\mathscr{G}_{A_2}\text{ if and only if }A_2>A_1. \tag{31}$$

Given the input matrix B, we view the trace as a set function by defining $\widetilde{f}:\mathscr{Z}_\mathscr{G}\mapsto\mathbb{R}_{\ge 0}$ such that for $\mathscr{G}_A\in\mathscr{Z}_\mathscr{G}$, $\widetilde{f}(\mathscr{G}_A)=\operatorname{tr}(\mathscr{W}_A)$.

Proposition 4.5. *(Trace of Gramian is a super-modular set function of edge locations and weights). The set function $\widetilde{f}$ is super-modular under the definition of set inclusion given by* (31)*.*

Proof. From the definition of $\widetilde{f}$, we have $\widetilde{f}(\mathscr{G}_A)=\operatorname{tr}(\mathscr{W}_A)=f(A)$. For arbitrary $\gamma\in\mathbb{R}_{\ge 0}$ and $D\in\mathbb{R}^{n\times n}$, note that

$$\frac{df(A+\gamma D)}{d\gamma}=\operatorname{tr}\left(\nabla f(A+\gamma D)^\top D\right).$$

Now, consider directed network triplet sets $\mathscr{G}_{A_1}$ and $\mathscr{G}_{A_2}$ such that $\mathscr{G}_{A_1}\subset\mathscr{G}_{A_2}$. Consider the addition of $D\in\mathbb{R}^{n\times n}$ with non-negative entries, and associated directed network triplet set $\mathscr{G}_D$, to A_1 and A_2. Then,

$$\begin{aligned}\frac{d}{d\gamma}&\left(f(A_1+\gamma D)-f(A_2+\gamma D)\right)\\&=\operatorname{tr}\left(\left(\nabla f(A_1+\gamma D)-\nabla f(A_2+\gamma D)\right)^\top D\right).\end{aligned}$$

From Lemma 4.2, $\frac{\partial f(A)}{\partial a_{ji}}$ is non-decreasing for all $1\le i,j\le n$. Using this fact and (31), we deduce that $\nabla f(A_1+\gamma D)-\nabla f(A_2+\gamma D)\le 0$. Since the entries of D are non-negative, we also have $\operatorname{tr}\left(\left(\nabla f(A_1+\gamma D)-\nabla f(A_2+\gamma D)\right)^\top D\right)\le 0$, which implies

$$\frac{d}{d\gamma}\left(f(A_1+\gamma D)-f(A_2+\gamma D)\right)\le 0.$$

Integrating this expression gives

$$\int_0^1\frac{d}{d\gamma}\left(f(A_1+\gamma D)\right)d\gamma\le\int_0^1\frac{d}{d\gamma}\left(f(A_2+\gamma D)\right)d\gamma,$$

which in turn yields $f(A_1+D)-f(A_1)\le f(A_2+D)-f(A_2)$. In terms of $\widetilde{f}$, this inequality reads as

$$\widetilde{f}(\mathscr{G}_{A_1}\cup\mathscr{G}_D)-\widetilde{f}(\mathscr{G}_{A_1})\le\widetilde{f}(\mathscr{G}_{A_2}\cup\mathscr{G}_D)-\widetilde{f}(\mathscr{G}_{A_2}),$$

as claimed. □

Proposition 4.5 generalizes the result in [38, Theorem 1], which considers the case where edge weights are fixed (in which case (31) simplifies to the standard definition of set inclusion) and equal to 1.

Remark 4.6. *(Non-applicability of classical greedy algorithms for edge modification).* Classical greedy algorithms provide approximations of the optimal solution with guaranteed accuracy for set function optimization when the set function is super-modular and non-increasing in the set variable, see cf. [39]–[41]. From Proposition 4.5 and Lemma 4.1, however, the trace of Gramian is a super-modular and non-decreasing set function, and hence the guarantees on the performance of greedy algorithms are not applicable. The optimal edge addition problem falls in the category of super-modular function maximization (equivalently, sub-modular function minimization) problems, for which we refer the reader to [42] for a survey of available solution procedures. As the edge set cardinality is n^2, these solution procedures in general have an order of computational complexity n^k, where $k\ge 6$, which becomes computationally infeasible as the network size increases. □

4.2 Log-det or Inverse of Trace Inverse of Gramian as Objective

The task of characterizing the optimal solution of (3) or (4) when the objective function is $\log\det(\mathscr{W}_A)$ or $(\operatorname{tr}(\mathscr{W}_A^{-1}))^{-1}$ is challenging. Here, we restrict our attention to systems with diagonal controllability Gramian and prove that the optimal solution lies on the boundary of the feasibility set if modifications are made only on the existing edges (edge re-weighing). An example of systems with diagonal controllability Gramian are '*stem-bud*' networks, cf. [43]. We start by showing that the diagonal nature of the Gramian is preserved after weight modifications in existing edges.

Lemma 4.7. *(Diagonal Gramian systems under existing edge modification). Consider a network system* (1) *with diagonal controllability Gramian $\mathscr{W}_A$. Let δA be any modification with positive weights in its existing edges. Then, the resulting controllability Gramian $\mathscr{W}_{A+\delta A}$ is also diagonal.*

Proof. By assumption, $A,B,\delta A$ have non-negative elements and modifications are made only on existing edges. Therefore, the matrices $\mathscr{W}_{A+\delta A}$ and $\mathscr{W}_A$ have zero (resp. non-zero entries) at identical places, and the result follows. □

Note that, when modifications are made in non-existing edges, the Gramian may lose its diagonal character. Next, we prove that gradient of $\log\det(\mathscr{W}_A)$ and $(\mathrm{tr}(\mathscr{W}_A^{-1}))^{-1}$ for systems with diagonal controllability Gramian is non-negative.

Lemma 4.8. *(Log-det and inverse of trace inverse for diagonal Gramians are non-decreasing). Consider a network system* (1) *with diagonal controllability Gramian $\mathscr{W}_A$. If $f = \log\det(\mathscr{W}_A)$ or $f = (\mathrm{tr}(\mathscr{W}_A^{-1}))^{-1}$, then $\frac{\partial f}{\partial a_{ji}} \geq 0$, i.e., if $A_2 - A_1 \geq 0$ then $f(A_2) \geq f(A_1)$.*

Proof. Both $\mathscr{W}_A$ and $\mathscr{W}_A^{-1}$ are diagonal and positive definite, with $\frac{\partial \mathscr{W}_A}{\partial a_{ji}}$ having non-negative entries. The result follows from Corollary 3.2. □

From Lemmas 4.7 and 4.8, we infer that when we modify existing edges, the system retains the diagonal nature of the Gramian and $\log\det(\mathscr{W}_A)$ and $(\mathrm{tr}(\mathscr{W}_A^{-1}))^{-1}$ are non-decreasing. Consider the following optimization problem similar to (28) by fixing $\mathscr{E}_{opt}$ as a subset of existing edge locations,

$$\max_{w} \quad f(A+\delta A) = \log\det(\mathscr{W}_{A+\delta A}) \text{ or } (\mathrm{tr}(\mathscr{W}_{A+\delta A}))^{-1}, \tag{32a}$$

$$\text{s.t.} \quad \delta A = \sum_{k=1}^{N_{\max}} w_k \Delta_k, \tag{32b}$$

$$w = \begin{pmatrix} w_1 & w_2 & \dots & w_{N_{\max}} \end{pmatrix}^\top, \tag{32c}$$

$$0 \leq w_k \leq w_{ub}, \quad k = 1, \dots, N_{\max}, \tag{32d}$$

$$\sum_{k=1}^{N_{\max}} w_k \leq w_{\max}. \tag{32e}$$

We next characterize the optimal solution of the edge modification problem when the objective is $\log\det(\mathscr{W}_A)$ or $(\mathrm{tr}(\mathscr{W}_A^{-1}))^{-1}$.

Proposition 4.9. *(Characterization of optimal solution for* $\log$*det and inverse of trace inverse of Gramian). Consider a network system* (1) *with diagonal controllability Gramian. Given the edge addition set $\mathscr{E}_{opt}$ as a subset of the existing edge set, with cardinality $N_{\max}$, consider the optimization problem* (32)*. Then,*

(i) there always exists a global maximum w^* of (32) *satisfying*

$$\sum_{k=1}^{N_{\max}} w_k^* = w_{\max}, \tag{33}$$

when $w_{\max} \leq N_{\max} w_{ub}$; otherwise $w_k^* = w_{ub}$ for all k;

(ii) if, further, f is not constant as a function of w_i for any $i = 1, \dots, N_{\max}$, then any *global maximum of* (32) *satisfies* (33)*. In particular, no global maximum may lie in the interior of the feasibility set.*

The proof of Proposition 4.9 follows the same line of argumentation as that of Proposition 4.3, and we omit it for brevity. This result is not true if $\mathscr{W}_A$ is non-diagonal and/or modifications are made in non-existing edges. This is because a non-diagonal $\mathscr{W}_A^{-1}$ may have negative non-diagonal elements, due to which no guarantee can be made regarding the monotonicity of f. To prove super/sub-modularity of f, one needs to investigate the increasing/decreasing nature of the gradient, which is itself a challenging task and a direction of future research.

5 Edge Modification Algorithms

In this section, we build on the results of Sections 3 and 4 to propose two computationally efficient edge modification procedures that yield approximate solutions to the optimization problem (3).

Given the computational complexity of solving this problem, our designs are based on the idea of restricting the search space. In both cases, we employ the term '*Restricted-set*' to refer to the shrinking of the original feasible set to a smaller edge selection subset $\mathscr{E}_S$ as formulated in (4). We construct $\mathscr{E}_S$ using the ETEC measure introduced in Section 3.2. To do so, we compute the ETEC of every possible edge (including non-existing ones), sort the edges in decreasing order according to their ETEC, and select the top N_S edges. The value of N_S is a design parameter whose choice may be driven by a number of factors (e.g., considerations about computational capability, significant gaps in the sequence of ETEC values, etc.).

5.1 Restricted-Set Optimization (RSO)

Restricted-Set Optimization consists of solving the optimization problem (4) with the set $\mathscr{E}_S$ of N_S edges of largest ETEC measure. The resulting MINLP is then computationally tractable and we solve it by using publicly available software such as the OPTI TOOLBOX [44] for MATLAB which uses the BONMIN solver [45] for integer variables. Algorithm 1 describes the pseudocode of the resulting edge modification procedure.

Algorithm 1 Restricted-set optimization

Input: $A, B, n, T, N_{\max}, w_{\max}, N_S, w_{ub}$
Output: δA

1: Compute $\{\mathscr{W}_j^{(t)}\}_{t=1}^T$ as in (6) for each node $j \in \{1, \dots, n\}$
2: Form $\mathscr{H}^{(t)}$ as in (25)
3: Compute $p^{(t)}, q^{(t)}, \Theta^{(t)}$ and Θ as in (27)
4: Rank all the edges according to decreasing ETEC measure c_{ij}
5: Select the first N_S edges and form the set $\mathscr{E}_S$
6: Solve the optimization problem (4) to get η, w
7: Compute $\delta A = \sum_{k=1}^{N_S} \eta_k w_k \Delta_k$

We note that, albeit the optimization problem solved by RSO is certainly more tractable than the original optimization (3), it still is an MINLP problem and, as such, is subject to the limitations of nonconvex solvers (particularly, the choice of initial conditions).

5.2 Restricted-Set Exhaustive Greedy (RSEG)

This approach is based on the following observations: (i) from Proposition 4.3, the optimal solution has weights lying on the boundary of the feasible region; (ii) the optimal solution tends to have weights in a corner point of the boundary, cf Figure 3; and (iii) the supermodularity of the trace of Gramian, cf. Proposition 4.5, means that, when adding new or strengthening existing edges, larger weights are more beneficial. Therefore, in this second approach we choose the weights first (by letting $N_m = \lfloor \frac{w_{\max}}{w_{ub}} \rfloor$ edges to have weight w_{ub} and one edge to have weight $w_r = w_{\max} - N_m w_{ub}$) and then determine the edge locations, choosing them one at a time. For the latter, we sort the edges with respect to their ETEC, select the search set $\mathscr{E}_S$, exhaustively search for the best edge location in the set $\mathscr{E}_S$ (only considering weight assignments of the form $w_{ub}, w_{ub}, \dots, w_{ub}, w_r$), update the adjacency matrix A with the best edge location, and repeat the procedure N_m+1 times to get the optimal edge set $\mathscr{E}_{opt}$ (in case $N_m+1<N_{\max}$, we assign the first N_m edges weight w_{ub}, w_r to the next edge, and zero weight to the remaining $N_{\max}-N_m-1$ edges). We summarize the procedure in Algorithm 2.

Algorithm 2 Restricted-set exhaustive greedy

Input: $A, B, n, T, N_{\max}, w_{\max}, N_S, w_{ub}$
Output: δA

1: Compute N_m, w_r
2: Set $k=1$, $\delta A=0$, $\mathscr{E}_{opt}=\{\}$, and $\mathscr{E}$ equal to the set of all possible edges
3: Compute $\{\mathscr{W}_j^{(t)}\}_{t=1}^{T}$ as in (6) for each node $j\in\{1,\ldots,n\}$
4: Form $\mathscr{H}^{(t)}$ as in (25)
5: Compute $p^{(t)}, q^{(t)}, \Theta^{(t)}$ and Θ as in (27)
6: Rank all the edges according to decreasing ETEC measure c_{ij} for edges in $\mathscr{E}\setminus\mathscr{E}_{opt}$.
7: Select the first N_S edges and form the set $\mathscr{E}_S$
8: Assign $w=w_{ub}$ if $1\le k\le N_m$ else $w=w_r$
9: Compute $(i,j)^*=\underset{(i,j)\in\mathscr{E}_S}{\operatorname{argmax}} f(A+w\Delta_{ji})$ ($\Delta_{ji}\in\mathbb{R}^{n\times n}$ has the (j,i) entry equal to 1 and all other entries equal to zero)
10: Set $\delta A=\delta A+w\Delta_{ji}{}^*$, $A=A+w\Delta_{ji}{}^*$. Update $k=k+1$, $\mathscr{E}_{opt}=\mathscr{E}_{opt}\cup\{(i,j)^*\}$ and repeat Step 3 to Step 10 until $k=N_m+2$

5.3 Exhaustive Greedy (EG)

For comparison purposes with the two methods proposed above, we also implement a third approach termed *Exhaustive greedy*. In this method, the weight vector is the same as in the '*Restricted-set exhaustive greedy*' method. However, we then exhaustively search the best edge to be modified in the set of all possible edges $\mathscr{E}$. This step is similar to Step 9 in Algorithm 2 with $\mathscr{E}_S$ replaced by $\mathscr{E}$. After each edge modification, we decrease the size of the set $\mathscr{E}$ by 1 as the modified edge location is deleted from it. In the case of using *RSEG* or *EG*, one could ensure the stability of the resulting network by checking whether the magnitude of the maximum eigenvalue of the system matrix is smaller than 1 after each edge modification.

5.4 Numerical Examples

Here, we provide two examples to illustrate the efficacy of the proposed edge modification procedures in improving network controllability. We perform all the simulations using MATLAB on a desktop with Intel core-i7-8700, 3.20 GHz processor with 16 GB of RAM. It should be noted that, while modifying edges, we exclude the addition of self-loops.

The optimal edge modification problem (3) is a non-convex MINLP problem and has discrete edge location variables. This results in multiple local maxima corresponding to different combinations of variables. For such non-convex MINLPs, the number of local maxima increases with the number of optimization variables ($2N_S$) and the computed solution will eventually depend on the initial starting solution used for the optimization process [46]. Further, as the number of variables increases, the computational efficiency deteriorates rapidly, making the computation of even local maximum a challenging task.

5.4.1 10-Node Numerical Example

We consider a network with $n=10$ nodes and 14 edges without self-loops, and adjacency matrix A whose entries are all zero except for

$$\begin{aligned} &a_{12}=0.69, && a_{18}=0.36, && a_{1,10}=1.24,\\ &a_{23}=0.20, && a_{25}=0.02, && a_{26}=0.87,\\ &a_{27}=0.64, && a_{37}=0.37, && a_{46}=0.76,\\ &a_{51}=0.66, && a_{76}=0.99, && a_{84}=0.50,\\ &a_{91}=0.52, && a_{10,9}=0.74, \end{aligned}$$

and a diagonal $B\in\mathbb{R}^{n\times n}$ with diagonal entries

$$\operatorname{diag}(B)=\begin{pmatrix}0 & 0 & 0 & 1 & 1 & 1 & 0 & 1 & 0 & 0\end{pmatrix}.$$

We have constraint bounds on the optimal edge modification problem (4) as $N_{\max}=3$, $w_{\max}=0.6$, and $w_{ub}=0.25$, and the time horizon is $T=2n$. We take $N_S=5$, which represents just 5.6% of the complete edge selection space (of cardinality 90 due to the removal of self-loops). The top five edges in descending order of their ETEC are $1\to6$, $1\to10$, $1\to9$, $5\to6$, and $5\to10$.

We compare the performance of the three approaches discussed in Section 5 to find an approximate solution to (4). We perform the edge modification for the trace as well as the log det of Gramian as objective functions, and present the results in Tables 1 and 2.

TABLE 1: Performance comparison of edge modification algorithms with trace of the Gramian as objective function. Here, both RSO and RSEG only reason over 5.6% of the set of all possible edges, whereas EG reasons over all 100% of them.

Property		*RSO*	*RSEG*	*EG*
Edges	Edge 1	$1\to6$	$1\to9$	$1\to9$
	Edge 2	$1\to10$	$1\to10$	$1\to10$
	Edge 3	$1\to9$	$1\to6$	$1\to6$
Weights	w_1	0.25	0.25	0.25
	w_2	0.25	0.25	0.25
	w_3	0.10	0.10	0.10
Initial $\operatorname{tr}(\mathscr{W}_A)$		9.27	9.27	9.27
Final $\operatorname{tr}(\mathscr{W}_A)$		36.4	32.8	32.8
Percentage increase		293%	254%	254%
Initial $\log\det(\mathscr{W}_A)$		-12.37	-12.37	-12.37
Final $\log\det(\mathscr{W}_A)$		-9.2	-7.9	-7.9
Initial $\lambda_{\min}(\mathscr{W}_A)$		0.00030	0.00030	0.00030
Final $\lambda_{\min}(\mathscr{W}_A)$		0.00033	0.00040	0.00040

From Table 1, with trace of Gramian as the objective function, we can see that the performance of *RSEG* is the same as that of *EG*, and that *RSO* performs the best. Interestingly, all the three approaches result in the same edge selections but with different weight combinations. Furthermore, even though the *RSO* and *RSEG* approaches seek to maximize $\operatorname{tr}(\mathscr{W}_A)$, one can see that they also maintain (and even improve) worst-case controllability. We show the original and modified graphs using all approaches in Figure 4.

From Table 2, with log det of Gramian as the performance objective, we observe that all approaches produce similar performance improvements. We note that the *EG* approach produces an edge combination with the lowest increase in the trace. It is also worth noticing that for $N_S=16$ and $N_S=19$, *RSEG* obtains an

TABLE 2: Performance comparison of edge modification algorithms with log det of the Gramian as objective function. Here, both RSO and RSEG only reason over 5.6% of the set of all possible edges, whereas EG reasons over all 100% of them.

Property		*RSO*	*RSEG*	*EG*
Edges	Edge 1	$1 \to 9$	$1 \to 9$	$1 \to 9$
	Edge 2	$5 \to 10$	$9 \to 10$	$3 \to 5$
	Edge 3	$1 \to 6$	$5 \to 10$	$9 \to 10$
Weights	w_1	0.25	0.25	0.25
	w_2	0.25	0.25	0.25
	w_3	0.10	0.10	0.10
Initial $\log\det(\mathscr{W}_A)$		-12.37	-12.37	-12.37
Final $\log\det(\mathscr{W}_A)$		-4.76	-3.9	-3.25
Percentage increase		61.5%	68.5%	73.7%
Initial $\lambda_{\min}(\mathscr{W}_A)$		0.00030	0.00030	0.00030
Final $\lambda_{\min}(\mathscr{W}_A)$		0.00050	0.00060	0.035
Initial $\mathrm{tr}(\mathscr{W}_A)$		9.27	9.27	9.27
Final $\mathrm{tr}(\mathscr{W}_A)$		22.8	24.2	13.2

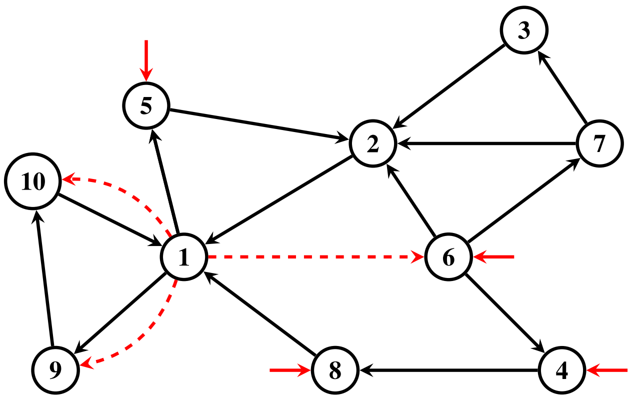


Fig. 4: The result of the proposed edge modification algorithms on the network of Example 1 with trace of the Gramian as objective (cf. Table 1). Nodes with red arrows (4, 5, 6, 8) represent the actuator locations, while solid and dashed edges represent original and modified edges, respectively. All three approaches result in the same edge selections but with different weight assignments.

TABLE 3: Performance comparison RSO and RSEG combined with different notions of edge centrality. Here, $N_S = 5$.

f	Initial	SE+*RSO*	SE+*RSEG*	TEC+*RSO*	TEC+*RSEG*	ETEC+*RSO*	ETEC+*RSEG*
tr	9.27	11.9	11.6	26.9	26	36.4	32.8
log det	-12.37	-8.45	-10.8	-10.1	-8.9	-4.7	-3.9

optimal objective function value of -3.07 and -2.88, respectively, which is better than *EG*.

We also compare the *RSO* and *RSEG* edge modification algorithms based on the ETEC measure against the following two benchmark procedures: (i) strengthening existing edges (SE) by using *RSO* and *RSEG* and (ii) using topological edge centrality (TEC), cf. [22] in combination with *RSO* and *RSEG*. Table 3 shows the comparison. One can observe that the best result is obtained with the ETEC-based implementation, which reinforces our claim that, along with the network topology, the consideration of the dynamical effects in the network (i.e., energy exchange effects) is important for enhancing controllability. In Tables 1 and 2, we have omitted the values of $(\mathrm{tr}(\mathscr{W}_A^{-1}))^{-1}$, which are approximately equal to $\lambda_{\min}(\mathscr{W}_A)$.

5.4.2 Random Erdős-Rényi Networks

In this example, we implement the proposed edge modification procedures on 10^3 random Erdős-Rényi networks [47] without self-loops. We consider networks having $n = 25$ nodes, 8 input nodes, and edge probability of 0.2. We modify a maximum number of $N_{\max} = 3$ edges with budget constraint $w_{\max} = 1$ on the added weight, and the maximum weight to be added to each edge is $w_{ub} = 0.4$. The time horizon is $T = 2n$. We then apply the algorithm *RSO* with $N_S = 15$ and the algorithm *RSEG* with N_S equal to 3% of $n^2 - n$ (i.e., 18). For each network, we compute the percentage of increase in controllability and show the histogram of the resulting values in Figure 5. The mean percentage increase in the performance is about 10^5%, showing the utility of the proposed edge centrality measure and edge modification algorithms.

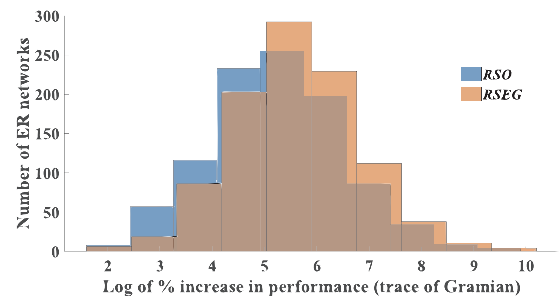


Fig. 5: The histogram of $\log_{10}$(% increase in $\mathrm{tr}(\mathscr{W}_A)$) for 10^3 Erdős-Rényi random networks following the proposed edge modification algorithm.

To illustrate the impact of the reduction of the search space in finding the best locations for edge addition, Figure 6 shows, for the same 10^3 Erdős-Rényi random networks, the histogram of the minimum size of the set $\mathscr{E}_S$, i.e., minimum N_S for the algorithm *RSEG*, that is required to match or improve $\mathrm{tr}(\mathscr{W}_A)$ relative to *EG*. The histogram shows that in $\simeq 95\%$ cases, no more than the top 1% of the edges selected using the proposed edge centrality measure are needed for the *RSEG* algorithm to reach the result of the corresponding *EG*. In our simulations, executing *RSEG* took on an average 25% less time (in CPU seconds) than executing *EG*. This time difference will grow as n increases. This is a remarkable enhancement in computational cost which illustrates the utility of the proposed edge centrality measure.

For the same 10^3 Erdős-Rényi random networks, Figure 7 shows the result of using the *RSEG* algorithm with $N_S = 30$ (5% of total edges) and $\log\det(\mathscr{W}_A)$ as objective function. Finally, Figure 8 shows the histogram of the minimum size of the set $\mathscr{E}_S$, i.e., minimum N_S for the algorithm *RSEG* required to match or improve $\log\det(\mathscr{W}_A)$ relative to *EG*. Unlike the case of using $\mathrm{tr}(\mathscr{W}_A)$ in Figure 6, the histogram in Figure 8 is almost evenly spread from $N_S = 5\% - 90\%$ of total edges. This is due to the fact that ETEC is always positive while the gradient of $\log\det(\mathscr{W}_A)$ may be negative. Nonetheless, we see from Figure 8 that even though less than the corresponding improvement in $\mathrm{tr}(\mathscr{W}_A)$, we still obtain an average objective function increase of $\simeq 22\%$ in $\log\det(\mathscr{W}_A)$. It should also be noted that in the latter case when we used $\log\det(\mathscr{W}_A)$ as objective, *RSO* along with the use of OPTI TOOLBOX produced infeasible results.

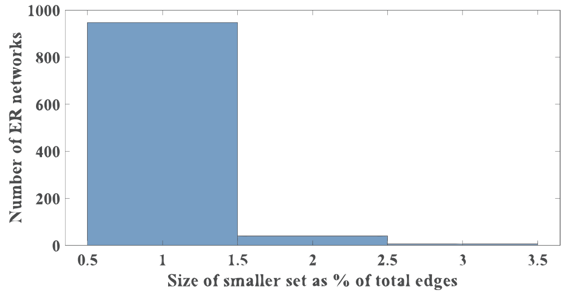


Fig. 6: The histogram of minimum N_S required for *RSEG* to match or better the corresponding *EG* value of the trace of the Gramian for 10^3 Erdős-Rényi random networks.

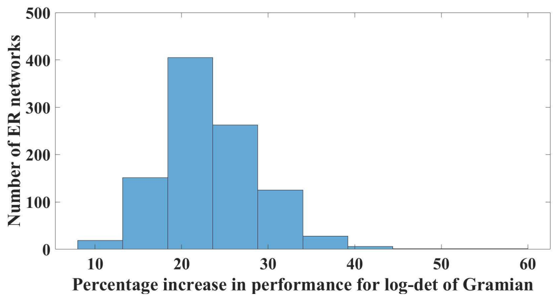


Fig. 7: The histogram of %-increase in $\log\det(\mathscr{W}_A)$ for 10^3 Erdős-Rényi random networks following the proposed *RSEG* edge modification algorithm.

6 Conclusions

We have studied the optimization problem that seeks to optimally modify the location and weights of edges to improve network controllability. To address the increasing computational intractability with the size of the network of the non-convex mixed-integer program, we have proposed reducing the search space by restricting the optimization to a subset of all possible edges. To select this restricted set, we have introduced the novel notion of energy-based edge centrality and characterized its relationship with the gradient of the trace, the $\log\det$ and the inverse of trace inverse of the network's controllability Gramian. We have also showed that the optimal solution of the edge modification algorithm lies on the boundary of the feasible region, in general for the trace and in particular in the case of systems with diagonal Gramian for $\log\det$ and the inverse of the trace inverse of Gramian. Building on these results, we have proposed the Restricted-Set Optimization and the Restricted-Set Exhaustive Greedy algorithms, which are computationally efficient edge modification procedures that yield approximate solutions to the original optimization problem. We have illustrated their performance on numerical simulations in a 10-node network and on random Erdős-Rényi 25-node networks. Future work will explore networks with negative edge weights, the analytical characterization of the sub-optimality gap of the proposed edge modification algorithms, method to select the size of the restricted feasible set, the consideration of stability constraints, the relationship between ETEC, network diameter, and the nonnormality of the adjacency matrix, the extension of the results on $\log\det$ and inverse of the trace inverse of Gramian to general network systems and the study of network resiliency in strategic scenarios with edge attacks. We also plan to apply our proposed notions to real-world networks to solve problems such as optimal rewiring of networks and optimal information transmission in communication networks.

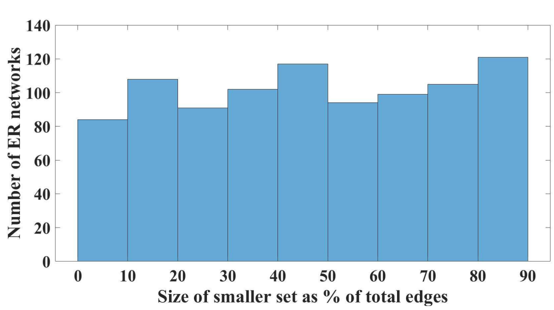


Fig. 8: The histogram of minimum N_S required for *RSEG* to match or better the corresponding *EG* value of the $\log\det$ of the Gramian for 10^3 Erdős-Rényi random networks.

References

[1] P. V. Chanekar, E. Nozari, and J. Cortés, "Network modification using a novel Gramian-based edge centrality," in *IEEE Conf. on Decision and Control*, Nice, France, Dec. 2019, pp. 1686–1691.

[2] B. Zhang, Y. Tian, and Z. Zhang, "Network biology in medicine and beyond," *Circulation: Cardiovascular Genetics*, vol. 7, no. 4, pp. 536–547, 2014.

[3] T. Gross and H. Sayama, *Adaptive Networks: Theory, Models and Applications*. MIT Press, 2009.

[4] R. E. Kalman, Y. Ho, and K. S. Narendra, "Controllability of linear dynamical systems," *Contributions to Differential Equations*, vol. 1, pp. 189–213, 1963.

[5] C. T. Chen, *Linear System Theory and Design*, 3rd ed. New York, NY, USA: Oxford University Press, Inc., 1998.

[6] P. C. Müller and H. I. Weber, "Analysis and optimization of certain qualities of controllability and observability for linear dynamical systems," *Automatica*, vol. 8, no. 3, pp. 237–246, 1972.

[7] G. Yan, J. Ren, Y. Lai, C. Lai, and B. Li, "Controlling complex networks: How much energy is needed?" *Physical Review Letters*, vol. 108, no. 21, p. 218703, 2012.

[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, 2016.

[10] N. Bof, G. Baggio, and S. Zampieri, "On the role of network centrality in the controllability of complex networks," *IEEE Transactions on Control of Network Systems*, vol. 4, no. 3, pp. 643–653, 2017.

[11] G. Lindmark and C. Altafini, "Combining centrality measures for control energy reduction in network controllability problems," in *European Control Conference*, Naples, Italy, June 2019, pp. 1518–1523.

[12] E. Nozari, F. Pasqualetti, and J. Cortés, "Heterogeneity of central nodes explains the benefits of time-varying control scheduling in complex dynamical networks," *Journal of Complex Networks*, vol. 7, no. 5, pp. 659–701, 2019.

[13] V. Tzoumas, M. A. Rahimian, G. J. Pappas, and A. Jadbabaie, "Minimal actuator placement with bounds on control effort," *IEEE Transactions on Control of Network Systems*, vol. 3, no. 1, pp. 67–78, 2016.

[14] P. V. Chanekar, N. Chopra, and S. Azarm, "Optimal actuator placement for linear systems with limited number of actuators," in *American Control Conference*, Seattle, WA, May 2017, pp. 334–339.

[15] A. Jadbabaie, A. Olshevsky, and M. Siami, "Deterministic and randomized actuator scheduling with guaranteed performance bounds," *arXiv preprint arXiv:1805.00606*, 2018.

[16] X. Chen, S. Pequito, G. J. Pappas, and V. M. Preciado, "Minimal edge addition for network controllability," *IEEE Transactions on Control of Network Systems*, vol. 6, no. 1, pp. 312–323, 2018.

[17] S. Hassan-Moghaddam, X. Wu, and M. R. Jovanović, "Edge addition in directed consensus networks," in *American Control Conference*, Seattle, WA, May 2017, pp. 5592–5597.

[18] M. Siami and N. Motee, "Growing linear dynamical networks endowed by spectral systemic performance measures," *IEEE Transactions on Automatic Control*, vol. 63, no. 7, pp. 2091–2106, 2018.

[19] C. O. Becker, S. Pequito, G. J. Pappas, and V. M. Preciado, "Network design for controllability metrics," *IEEE Transactions on Control of Network Systems*, vol. 7, no. 3, pp. 1404–1415, 2020.

[20] S. Zhao and F. Pasqualetti, "Networks with diagonal controllability Gramian: analysis, graphical conditions, and design algorithms," *Automatica*, vol. 102, pp. 10–18, 2019.

[21] H. Zhuge and J. Zhang, "Topological centrality and its e-Science applications," *Journal of the American Society for Information Science and Technology*, vol. 61, no. 9, pp. 1824–1841, 2010.

[22] U. Brandes, "A faster algorithm for betweenness centrality," *Journal of Mathematical Sociology*, vol. 25, no. 2, pp. 163–177, 2001.

[23] C. Klymko, "Centrality and communicability measures in complex networks: Analysis and algorithms," Ph.D. dissertation, Emory University, 2013.

[24] F. Arrigo and M. Benzi, "Edge modification criteria for enhancing the communicability of digraphs," *SIAM Journal on Matrix Analysis and Applications*, vol. 37, no. 1, pp. 443–468, 2016.

[25] ——, "Updating and downdating techniques for optimizing network communicability," *SIAM Journal on Scientific Computing*, vol. 38, no. 1, pp. B25–B49, 2016.

[26] K. Fitch and N. E. Leonard, "Joint centrality distinguishes optimal leaders in noisy networks," *IEEE Transactions on Control of Network Systems*, vol. 3, no. 4, pp. 366–378, 2016.

[27] Y. Ghaedsharaf, M. Siami, C. Somarakis, and N. Motee, "Centrality in time-delay consensus networks with structured uncertainties," *arXiv preprint arXiv:1902.08514*, 2019.

[28] V. Preciado, M. Zargham, C. Enyioha, A. Jadbabaie, and G. Pappas, "Optimal resource allocation for network protection against spreading processes," *IEEE Transactions on Control of Network Systems*, vol. 1, no. 1, pp. 99–108, 2014.

[29] R. A. Fisher, "Statistical methods for research workers," in *Breakthroughs in statistics*. Springer, 1992, pp. 66–70.

[30] J. Gibbons and S. Chakraborti, *Nonparametric Statistical Inference: Revised and Expanded*. CRC press, 2014.

[31] G. Baggio, V. Rutten, G. Hennequin, and S. Zampieri, "Efficient communication over complex dynamical networks: The role of matrix non-normality," *Science Advances*, vol. 6, no. 22, p. eaba2282, 2020.

[32] S. Boyd and L. Vandenberghe, *Convex Optimization*. Cambridge University Press, 2009.

[33] C. D. Meyer, *Matrix Analysis and Applied Linear Algebra*. SIAM, 2001.

[34] S. D. Wang, T. S. Kuo, and C.-F. C. F. Hsu, "Trace bounds on the solution of the algebraic matrix Riccati and Lyapunov equation," *IEEE Transactions on Automatic Control*, vol. 31, no. 7, pp. 654–656, 1986.

[35] H. Wolkowicz and G. Styan, "Bounds for eigenvalues using traces," *Linear Algebra and its Applications*, vol. 29, pp. 471–506, 1980.

[36] S. G. Krantz and H. R. Parks, *A Primer of Real Analytic Functions*, 2nd ed., ser. Birkhäuser Advanced Texts. Boston, MA: Birkhäuser, 2002.

[37] F. Bullo, J. Cortés, and S. Martinez, *Distributed Control of Robotic Networks*, ser. Applied Mathematics Series. Princeton University Press, 2009.

[38] M. V. Srighakollapu, R. Kalaimani, and R. Pasumarthy, "Optimizing average controllability of networked systems," in *IEEE Conf. on Decision and Control*, Nice, France, Dec. 2019, pp. 2066–2071.

[39] M. L. Fisher, G. L. Nemhauser, and L. Wolsey, "An analysis of approximations for maximizing submodular set functions - II," *Mathematical Programming Studies*, vol. 8, p. 73, 1978.

[40] A. A. Bian, J. M. Buhmann, A. Krause, and S. Tschiatschek, "Guarantees for greedy maximization of non-submodular functions with applications," in *International Conference on Machine Learning*, vol. 70, Sydney, Australia, 2017, pp. 498–507.

[41] A. Clark, B. Alomair, L. Bushnell, and R. Poovendran, *Submodularity in Dynamics and Control of Networked Systems*, ser. Communications and Control Engineering. New York: Springer, 2016.

[42] S. T. McCormick, "Submodular function minimization," *Discrete Optimization*, vol. 12, pp. 321–391, 2005.

[43] S. Zhao and F. Pasqualetti, "Discrete-time dynamical networks with diagonal controllability Gramian," *IFAC-PapersOnLine*, vol. 50, no. 1, pp. 8297–8302, 2017.

[44] J. Currie and D. I. Wilson, "OPTI: lowering the barrier between open source optimizers and the industrial Matlab user," in *Foundations of Computer-Aided Process Operations*, Savannah, GA, 2012, electronic proceedings.

[45] P. Bonami, L. T. Biegler, A. R. Conn, G. Cornuéjols, I. E. Grossmann, C. D. Laird, J. Lee, A. Lodi, F. Margot, N. Sawaya, and A. Wächter, "An algorithmic framework for convex mixed integer nonlinear programs," *Discrete Optimization*, vol. 5, no. 2, pp. 186–204, 2008.

[46] R. T. Haftka and Z. Gürdal, *Elements of Structural Optimization*, ser. Solid Mechanics and Its Applications. New York: Springer, 2012, vol. 11.

[47] P. Erdös and A. Rényi, "On the evolution of random graphs," *Publications of the Mathematical Institute of the Hungarian Academy of Sciences*, vol. 5, pp. 17–61, 1960.

Prasad Vilas Chanekar received his undergraduate degree from the Indian Institute of Technology, Kharagpur, India, M.Sc. degree from the Indian Institute of Science, Bangalore, India and Ph.D. from the Department of Mechanical Engineering at the University of Maryland, College Park, USA. He is currently a Post-Doctoral researcher in the Department of Mechanical and Aerospace Engineering at the University of California, San Diego, USA. His research interests include dynamics, control and optimization of large-scale engineering systems.

Erfan Nozari received his B.Sc. degree in Electrical Engineering-Control in 2013 from Isfahan University of Technology, Iran, received his Ph.D. in Mechanical Engineering and Cognitive Science in 2019 from University of California San Diego, and was subsequently a postdoctoral researcher in the University of Pennsylvania Department of Electrical and Systems Engineering. He is currently an Assistant Professor in the Department of Mechanical Engineering, University of California, Riverside. He has been the (co)recipient of the 2019 IEEE Transactions on Control of Network Systems Outstanding Paper Award, the Best Student Paper Award from the 57th IEEE Conference on Decision and Control, the Best Student Paper Award from the 2018 American Control Conference, and the Mechanical and Aerospace Engineering Distinguished Fellowship Award from the University of California San Diego. His research interests include systems and control theory and machine learning and their applications in computational and theoretical neuroscience.

Jorge Cortés (M'02, SM'06, F'14) received the Licenciatura degree in mathematics from Universidad de Zaragoza, Zaragoza, Spain, in 1997, and the Ph.D. degree in engineering mathematics from Universidad Carlos III de Madrid, Madrid, Spain, in 2001. He held postdoctoral positions with the University of Twente, Twente, The Netherlands, and the University of Illinois at Urbana-Champaign, Urbana, IL, USA. He was an Assistant Professor with the Department of Applied Mathematics and Statistics, University of California, Santa Cruz, CA, USA, from 2004 to 2007. He is currently a Professor in the Department of Mechanical and Aerospace Engineering, University of California, San Diego, CA, USA. He is the author of Geometric, Control and Numerical Aspects of Nonholonomic Systems (Springer-Verlag, 2002) and co-author (together with F. Bullo and S. Martínez) of Distributed Control of Robotic Networks (Princeton University Press, 2009). He is a Fellow of IEEE and SIAM. At the IEEE Control Systems Society, he has been a Distinguished Lecturer (2010-2014), and is currently its Director of Operations and an elected member (2018-2020) of its Board of Governors. His current research interests include distributed control and optimization, network science, resource-aware control, nonsmooth analysis, distributed decision making, network neuroscience, and multi-agent coordination in robotic, power, and transportation networks.