EDGE DELETION ALGORITHMS FOR MINIMIZING SPREAD IN SIR EPIDEMIC MODELS[*]

YUHAO YI[†], LIREN SHAN[‡], PHILIP E. PARÉ[§], AND KARL HENRIK JOHANSSON[†]

Abstract. This paper studies algorithmic strategies to effectively reduce the number of infections in susceptible-infected-recovered (SIR) epidemic models. We consider a Markov chain SIR model and its two instantiations in the deterministic SIR (D-SIR) model and the independent cascade SIR (IC-SIR) model. We investigate the problem of minimizing the number of infections by restricting contacts under realistic constraints. Under moderate assumptions on the reproduction number, we prove that the infection numbers are bounded by supermodular functions in the D-SIR model and the IC-SIR model for large classes of random networks. We propose efficient algorithms with approximation guarantees to minimize infections. The theoretical results are illustrated by numerical simulations.

Key words. Markov chains, SIR epidemic, discrete optimization

AMS subject classifications. 60J10, 92D30, 68Q25

DOI. 10.1137/20M1377011

1. Introduction. Epidemic spreading processes can significantly disrupt the functioning of society and pose risks to the health of individuals. Therefore, the study of epidemic spread has a long history. One of the most fundamental models is the susceptible-infected-recovered (SIR) model, where each individual can be in one of three states: susceptible, infected, or recovered. Variations of the SIR model have been proposed, including population models and networked models [32]. An overview of SIR models is provided in [48]. The behaviors of the SIR spread dynamics are extensively studied [23, 1, 14, 10].

The problem of efficiently controlling epidemic spread has received increasing attention. One line of work has focused on minimizing the spectral norm of the transition matrix to suppress the process [45, 46, 44, 7, 39, 28]. Resource allocation [39] and network modifying [45, 37, 34] strategies have also been studied. Optimal control problems have been studied to minimize the cost over a given horizon [24]. Since the total number of infections is an important criterion, problems of minimizing infections have been proposed [37, 33].

In this paper, we consider optimization problems that minimize infections in a networked SIR Markov chain model. We provide efficient strategies of modifying the

[†]Division of Decision and Control Systems, School of Electrical Engineering and Computer Science (EECS), KTH Royal Institute of Technology and Digital Futures, Stockholm, SE-100 44, Sweden (yuhaoy@kth.se, kallej@kth.se).

[‡]Department of Computer Science, Northwestern University, Evanston, IL 60208 USA (lirenshan2023@u.northwestern.edu).

[§]School of Electrical and Computer Engineering, Purdue University, West Lafayette, IN 47907 USA (philpare@purdue.edu).

structure of the contact network to minimize the number of infections. We show the effectiveness of the proposed algorithms through theoretical characterizations and numerical evaluations.

1.1. Related work. Following from the Markov chain susceptible-infected-susceptible (SIS) model [43], continuous and discrete Markov chain SIR models have been proposed in [48] and [40]. Deterministic models based on mean-field approximations in the Markov chain models have also been studied in these papers. The independent cascade SIR (IC-SIR) model can be viewed as a networked extension to the Reed–Frost model [14], which is one of the earliest SIR models studied in depth [1]. The IC-SIR model was proposed in [18] and popularized in [22], and it has a rich volume of follow-up studies on efficient algorithms for influence maximization [9], network design [47], and inferring network structure [31].

Some researchers have focused on problems of minimizing the spectral norm of a parameter matrix by removing nodes or edges. These problems have been proven to be NP-complete and NP-hard, respectively [44]. Several heuristic algorithms are proposed based on betweenness centrality [20, 41] or convex relaxation of the original problem [7]. The continuous version of the problem has been discussed in [38, 39, 28, 35, 21].

As for minimizing the number of infections, a graph partitioning problem has been investigated in [15], without considering any dynamics or random initial conditions in the network. A budget-constraint resource allocation problem has been studied in [33]. The problem is formulated as a continuous optimization by modifying the infection rates and recovery rates for nodes under the assumption that the system is stable. To the best of our knowledge, no algorithm with provable approximation guarantee has been proposed for the node and edge removal problems to minimize infections.

1.2. Contributions. We start by proposing a general Markov chain SIR (G-SIR) model. Two common SIR models are simplifications of the G-SIR model, the deterministic SIR (D-SIR) model and the IC-SIR model.

Our main results focus on the problems of minimizing the number of infections by removing edges in the contact network from a given candidate set. The candidate set models the reality that only limited types of contacts can be removed. We prove that the considered problems are NP-hard. In contrast, with moderate restrictions on the reproduction number, we show that the problems can be efficiently solved by using greedy algorithms.

For the D-SIR model, we propose an upper bound for the number of infections as a surrogate objective function. We prove that when the system is exponentially stable, the surrogate function is a monotone supermodular function of the set of deleted edges, although in a relaxed problem the same function is not convex with respect to the edge weights. Then it is shown that the number of infections can be reduced by efficiently minimizing the upper bound.

For the IC-SIR model, we study the expected number of infections. In this model, the epidemic spreads over a contagion network randomly sampled from a given contact network. We consider contact networks generated from Erdős–Rényi (ER) graphs and the stochastic block model (SBM). The statistics of the contagion network encode many properties of the spread process. For example, the expected average degree of the contagion network can be interpreted as the reproduction number. When the average degree of the contagion network is less than one, and the number of initial infections is no larger than $O(n^{\frac{1}{3}-c})$, where n is the population size and c is any positive number, we design efficient mitigation algorithms. For a contact network generated

from an ER graph, we prove that with high probability, the expected number of infections is well approximated by a monotone supermodular function. We obtain similar results for SBMs, with the additional assumption that the number of blocks is no larger than $O(\log n)$.

Our theoretical results are supported by numerical simulations. We show the effectiveness of the proposed algorithms by running simulations on both synthesized random networks and a real contact network. We compare the results given by the algorithms using the D-SIR model and the IC-SIR model and corresponding results for the G-SIR model. We find that both algorithms effectively reduce the number of infections also in the G-SIR model under the proposed conditions.

1.3. Outline. The remainder of the paper is organized as follows. In section 2 we introduce notation and review some basic concepts. In section 3 we describe the considered models. In section 4 we provide the general framework of the algorithm and present the hardness results. In section 5 we present a supermodular upper bound for the infections in the D-SIR model when the system is stable. We provide conditions such that, with high probability, the number of infections approximates a supermodular function, addressing the cases where, in sections 6 and 7, the contact network is sampled from an ER graph and a SBM, respectively. Section 9 provides some numerical experiments and simulations to show the effectiveness of the proposed algorithms in all the considered models, followed by the conclusion.

2. Preliminaries. Let $G = (V, E)$ be a graph with vertex (node) set V and edge set E. G can be either directed or undirected unless noted specifically. For a digraph, an edge from node j to node i is denoted as an ordered pair (j, i); for an undirected graph the edge is denoted as an unordered pair $\{j, i\}$. We define the size of a graph or connected component as the number of vertices (nodes) in it. We further denote by $[n]$ the set $\{1, 2, \ldots, n\}$.

Let $\mathcal{G}(n, p)$ denote the ER graph with n vertices where each edge occurs with probability p. We define the SBM $\mathcal{SBM}(n, \kappa, \boldsymbol{Q})$ as follows: Let $\{V_1, V_2, \cdots, V_\kappa\}$ be κ communities with n vertices. For each pair of vertices $u \in V_i$ and $v \in V_j$, the edge (u, v) is sampled with probability Q_{ij} independently.

In this paper we use bold font for matrices and vectors. We denote $\boldsymbol{e}_i, i \in [n]$, the ith canonical basis of $\mathbb{R}^n$. We use $\mathbf{1}$ and $\mathbf{0}$ to denote all one and all zero vectors. Further, we recall the definition and several properties of M-matrices. An M-matrix is a square matrix whose off-diagonal entries are nonpositive and eigenvalues have nonnegative real parts. It has the following property.

LEMMA 2.1 ([36, Theorem 1]). *A nonsingular M-matrix is inverse-positive, which means the entries of the inverse are all nonnegative.*

We adopt the standard asymptotic notation. Given two nonnegative functions f and g of the variable n, we denote $f = O(g)$ if there exists an n_0 and a constant $c > 0$, such that for all $n \geq n_0$, $f \leq c \cdot g$. Further, we denote $f = o(g)$ if f/g tends to zero when n tends to infinity. The term *with high probability* is used when the probability is $1 - o(1)$.

The following definitions are frequently used.

DEFINITION 2.2 (plus and minus operation for graphs). *Given two graphs $G_1 = (V_1, E_1)$, $G_2 = (V_2, E_2)$, $G' = (V', E') \stackrel{\text{def}}{=} G_1 + G_2$ is a new graph with vertex set $V' = V_1 \cup V_2$ and edge set $E' = E_1 \cup E_2$, and $G'' = (V'', E'') \stackrel{\text{def}}{=} G_1 - G_2$ is a graph with $V'' = V_1 \cup V_2$ and $E'' = E_1 \backslash E_2$.*

DEFINITION 2.3 (monotonicity). *A set function $f : 2^\Omega \mapsto \mathbb{R}$ is monotonically nonincreasing if $f(P_1) \geq f(P_2)$ holds for all $P_1 \subseteq P_2 \subseteq \Omega$.*

DEFINITION 2.4 (supermodularity). *A set function $f : 2^\Omega \mapsto \mathbb{R}$ is supermodular if $f(P_1) - f(P_1 \cup \{e\}) \geq f(P_2) - f(P_2 \cup \{e\})$ for all $P_1 \subseteq P_2 \subseteq E$ and all $e \in (E \backslash P_2)$.*

A function $f(P)$ is submodular if $g(P) \stackrel{\text{def}}{=} -f(P)$ is supermodular. A function $f(P)$ is modular if it is both submodular and supermodular.

DEFINITION 2.5 (($1 \pm \epsilon$)-approximation). *Given two real numbers $a, b \geq 0$ and an approximation error $\epsilon \geq 0$, if a and b satisfy*

$$(1 - \epsilon)b \leq a \leq (1 + \epsilon)b,$$

then b is an $(1 \pm \epsilon)$-approximation of a, denoted $a \approx_\epsilon b$.

We note that $a \approx_{o(1)} b$ if and only if $b \approx_{o(1)} a$.

3. Model description. In networked SIR models, each node, at a given time, is exclusively in one of three states: *susceptible* (S), *infected* (I), or *removed* (R). For node i, we define the $\{0, 1\}$ indicator variables for the states at time step t as $S_i(t)$, $I_i(t)$, and $R_i(t)$. The variables take value 1 if the node is in that state, and otherwise take value 0. We note that $S_i(t) + I_i(t) + R_i(t) = 1$. Motivated by [2, 40], we propose the G-SIR model using a 3^n-state Markov chain. The recursive relations between the states are defined as

$$(3.1) \qquad I_i(t+1) = I_i(t)(1 - \delta_i(t)) + S_i(t) \left(1 - \prod_{j=1}^{n} (1 - \beta_{ij}(t)I_j(t)) \right),$$

$$(3.2) \qquad R_i(t+1) = R_i(t) + \delta_i(t)I_i(t),$$

where $\beta_{ij}(t)$ is the random indicator variable for the event that node i is infected from its infected neighbor j; $\delta_i(t)$ is the random indicator variable for recovering. We assume that $\beta_{ij}(t)$ for any i, j are mutually independent. In addition, we assume that for a fixed i, j pair, $\beta_{ij}(t)$ are i.i.d with respect to k, and for a fixed i, $\delta_i(t)$ are i.i.d. with respect to t.

We simplify the G-SIR model by using mean-field approximation in a similar manner as [42, 13]. We take expectation for both sides of (3.1) and arrive at

$$\mathbb{E}\left[I_i(t+1)\right] = \mathbb{E}\left[I_i(t)\right] \mathbb{E}\left[1 - \delta_i(t)\right] + \mathbb{E}\left[S_i(t)\right] - \mathbb{E}\left[S_i(t)\right] \prod_{j=1}^{n} \mathbb{E}\left[1 - \beta_{ij}(t)I_j(t)\right],$$

by assuming that the states of all nodes are mutually independent for any given t. Let $x_i(t) \in [0, 1]$ and $r_i(t) \in [0, 1]$ be the probabilities of node i being infected and being removed at time step t, respectively. Then, by using the approximation $1 - x \approx e^{-x}$ for small x, the mean-field approximation of the corresponding discrete-time SIR model can be written as

$$x_i(t+1) = x_i(t) + (1 - x_i(t) - r_i(t)) \sum_{j=1}^{n} B_{ij} x_j(t) - D_i x_i(t),$$

$$r_i(t+1) = r_i(t) + D_i x_i(t),$$

where $B_{ij} \stackrel{\text{def}}{=} \mathbb{E}\left[\beta_{ij}(t)\right]$ and $D_i \stackrel{\text{def}}{=} \mathbb{E}\left[\delta_i(t)\right]$. B_{ij} is the infection rate for edge (j, i) and D_i is the healing rate for node i. In some variants of this model, B_{ij} is interpreted as the product of the infection rate and the edge weight [2, 21].

We define the vectors $\boldsymbol{x}(t), \boldsymbol{r}(t)$ with entries $x_i(t)$ and $r_i(t)$ for all nodes i. Then the discrete-time SIR model can be expressed in matrix form as follows:

$$(3.3) \qquad \boldsymbol{x}(t+1) = \boldsymbol{x}(t) + (\boldsymbol{I} - \boldsymbol{X}(t) - \boldsymbol{R}(t))\boldsymbol{B}\boldsymbol{x}(t) - \boldsymbol{D}\boldsymbol{x}(t),$$

$$(3.4) \qquad \boldsymbol{r}(t+1) = \boldsymbol{r}(t) + \boldsymbol{D}\boldsymbol{x}(t),$$

where the entries of $\boldsymbol{B}$ are B_{ij}; $\boldsymbol{D}$ is a diagonal matrix with diagonal entry D_i; and $\boldsymbol{X}(t)$ and $\boldsymbol{R}(t)$ are the diagonal matrices $\mathrm{diag}\,(\boldsymbol{x}(t))$ and $\mathrm{diag}\,(\boldsymbol{r}(t))$, respectively.

In this paper we study the cumulative number of infections, defined as $\boldsymbol{m}(t) \stackrel{\mathrm{def}}{=} \boldsymbol{x}(t) + \boldsymbol{r}(t)$. We denote $\boldsymbol{m}^*$ with entries $m_i^* \stackrel{\mathrm{def}}{=} \sup_t(m_i(t))$. Then $\|\boldsymbol{m}^* - \boldsymbol{m}(0)\|_1$ can be interpreted as the number of increased infections in the network. This vector $\boldsymbol{m}(t)$ is known to be inside $[0, 1]^n$ under the assumption that $D_i < 1$, $\sum_{j=1}^n B_{ij} < 1$ for all $i \in [n]$ [21]. Therefore, we can treat $m_i(t)$ as the probability that node i is in I or R at time t. We refer to the model given by (3.3) and (3.4) as the D-SIR model. It is well known that the mean-field approximation gives an upper bound for the probabilities of infections in the G-SIR model [13].

We consider the IC-SIR model [18, 22, 31] as another simplification to the G-SIR model, where $\delta_i(t) = 1$ for all i and t. In this model, we use random indicator variable vectors $\tilde{\boldsymbol{x}}(t) \in \{0, 1\}^n$ and $\tilde{\boldsymbol{r}}(t) \in \{0, 1\}^n$ to denote whether the nodes are in states I and R, respectively. The nonzero entries in $\tilde{\boldsymbol{x}}(0)$ are the initial infected nodes, which are referred to as seeds. Let $s \stackrel{\mathrm{def}}{=} \|\tilde{\boldsymbol{x}}(0)\|_1$ be the number of initial infections. We also define a matrix $\widetilde{\boldsymbol{B}}(t)$, where the entries are random variables $\beta_{ij}(t)$ which are i.i.d. with respect to t. The success probability for each entry $\beta_{ij}(t)$ in $\widetilde{\boldsymbol{B}}(t)$ is $p_{ij} \stackrel{\mathrm{def}}{=} \mathbb{E}\,[\beta_{ij}(t)]$. Thus, the IC-SIR model can be expressed as follows:

$$(3.5) \qquad \tilde{\boldsymbol{x}}(t+1) = \mathrm{sat}\left(\widetilde{\boldsymbol{B}}(t)\tilde{\boldsymbol{x}}(t) + \tilde{\boldsymbol{x}}(t) + \tilde{\boldsymbol{r}}(t)\right) - \tilde{\boldsymbol{x}}(t) - \tilde{\boldsymbol{r}}(t),$$

$$(3.6) \qquad \tilde{\boldsymbol{r}}(t+1) = \tilde{\boldsymbol{r}}(t) + \tilde{\boldsymbol{x}}(t),$$

where the saturation function $\mathrm{sat}\,(\cdot) : \mathbb{R}^n \mapsto \mathbb{R}^n$ is defined as $\mathrm{sat}\,(\boldsymbol{a})_i = 1$ if $a_i \geq 1$, and $\mathrm{sat}\,(\boldsymbol{a})_i = a_i$ if $a_i \leq 1$. In an IC-SIR model, the reproduction number R_0 is defined as the expected average degree $(1/n) \sum_{i,j=1}^n p_{ij}$.

For each edge in the network, it can induce a new infection only when one of its incident nodes is infected and the other one is susceptible. Since any infected node recovers in one time step, each edge (j, i) can change the number of infections at the time that j is infected. Therefore, the random transition matrix $\widetilde{\boldsymbol{B}}(t)$ can be sampled once beforehand, denoted as $\widetilde{\boldsymbol{B}}$. This matrix $\widetilde{\boldsymbol{B}}$ can be viewed as a $\{0, 1\}$ adjacency matrix of a sampled network where the actual spreading happens. This network is called the contagion network. At the end of a spread process, nodes are either in the removed state or the susceptible state. Further, we use $\tilde{\boldsymbol{r}}^*$ to denote the nodes eventually in state R: $\tilde{r}_i^* \stackrel{\mathrm{def}}{=} \sup_t(\tilde{r}_i(t))$ for each node i. Then $\|\tilde{\boldsymbol{r}}^* - \tilde{\boldsymbol{x}}(0) - \tilde{\boldsymbol{r}}(0)\|_1$ is the number of increased infections.

Then we provide two propositions to relate the expected number of infections in G-SIR models to the expected number of infections in IC-SIR models. The proofs of Propositions 3.1 and 3.2 are given in Appendix A.

Proposition 3.1 shows that the G-SIR model has the same expected number of infections as an IC-SIR model with modified probability.

PROPOSITION 3.1. *Given an instance of a G-SIR model, the expected number of infections of the model is equal to the expected number of infections in an IC-SIR model with the same edge set and $p_{ij} = (1 - (1 - D_j)(1 - B_{ij}))^{-1} B_{ij}$ for all edges.*

We also consider the symmetric G-SIR model, which has the infection rates $B_{ij} = B_{ji}$ for each edge (i, j) and the recover rates D_i are same for all nodes $i \in [n]$. We call the IC-SIR model symmetric if $p_{ij} = p_{ji}$ for all i, j pairs. Proposition 3.2 shows that the symmetric G-SIR model has the same expected number of infections as an undirected IC-SIR model with modified activation probabilities for each edge.

PROPOSITION 3.2. *Given an instance of a symmetric G-SIR model, the expected number of infections of the model is equal to the expected number of infections in an undirected IC-SIR model by letting $p_{ij} = (1 - (1 - D_j)(1 - B_{ij}))^{-1} B_{ij}$ for all undirected edges in the IC-SIR model.*

4. Problem formulation. We consider the problems of minimizing the number of infections for the D-SIR model and the IC-SIR model.

Problem 4.1. Given a digraph $G = (V, E)$, the infection rate B_{ij} for each edge (j, i) in G and healing rate D_i for every node $i \in V$, an initial state vector $\boldsymbol{x}(0) \in [0, 1]^n$ and $\boldsymbol{r}(0)$ such that $\boldsymbol{x}(0) + \boldsymbol{r}(0) \in [0, 1]^n$, a candidate edge deletion set $Q \subseteq E$, $|Q| = q$, and an integer $0 < k \leq q$, find a set of edges $P^* \subseteq Q$, where $|P^*| \leq k$, such that

$$(4.1) \qquad P^* \in \operatorname*{arg\,min}_{P \subseteq Q, |P|=k} \sigma(P),$$

where $\sigma(P) \stackrel{\text{def}}{=} \|\boldsymbol{m}^* - \boldsymbol{m}(0)\|_1$ is the number of increased infections after deleting P.

Problem 4.2. Given an undirected graph $G = (V, E)$, the activation probability of each edge $\mathbb{P}_{p_{ij} \stackrel{\text{def}}{=}} [\widetilde{B}_{ij}(t) = 1]$, an initial state vector $\tilde{\boldsymbol{x}}(0) \in \{0, 1\}^n$ with s nonzero entries, a candidate edge deletion set $Q \subseteq E$, where $|Q| = q$, and an integer $0 < k \leq q$, find a set of edges $P^* \subseteq Q$, where $|P^*| \leq k$, such that

$$(4.2) \qquad P^* \in \operatorname*{arg\,min}_{P \subseteq Q, |P|=k} \mathbb{E}\left[\tilde{\sigma}(P)\right],$$

where $\tilde{\sigma}(P) \stackrel{\text{def}}{=} \|\tilde{\boldsymbol{r}}^* - \tilde{\boldsymbol{x}}(0)\|_1$ is the expected number of increased infections after deleting P.

We note that $\sigma(P)$ and $\tilde{\sigma}(P)$ are defined to measure the number of new infections.

Remark 4.3. An extension of Problem 4.2 is to consider $\tilde{\boldsymbol{x}}(0) \in \{0, 1\}^n$, where each $\tilde{x}_i(0)$ is a Bernoulli random variable with mean μ_i. Let $s = \sum_{i=1}^{n} \mu_i$ be the expected number of initial infections. By the Chernoff bound, the number of initial infections in $\tilde{\boldsymbol{x}}(0)$ is at most $s + 3\sqrt{s \log n}$ with high probability.

Before we introduce our algorithmic results, we show a computational hardness result to justify the investigation of approximation algorithms or well-motivated heuristics.

THEOREM 4.4. *Given an instance of Problem 4.1 or Problem 4.2, it is NP-hard to find an optimal solution P^*.*

The proof and detailed discussion of Theorem 4.4 are given in Appendix B. We note that our proof also shows that the deterministic partitioning problem studied in [15] is NP-hard. In addition, the hard instances can be constructed for any given $R_0 > 0$ for an IC-SIR model.

In this paper we find conditions for the D-SIR model and the IC-SIR model such that the objectives of Problems 4.1 and 4.2 are bounded by monotone supermodular functions of the edge deletion set. For Problem 4.1, when the dynamics of the D-SIR

Algorithm 1: Greedy algorithm

    Input: a function $f \in \{\hat{\sigma}, \mathbb{E}\left[\tilde{\sigma}'\right]\}$, a contact network G, initial states of
             nodes, a candidate edge set Q, an integer k;
    Output: a set of k edges $P \subseteq Q$;
    Initialize the set $P = \varnothing$;
    for *round i from 1 to k* **do**
        | Compute the corresponding function $f(P \cup \{e\})$ for any $e \in Q \setminus P$;
        | Set $e^* = \arg\max_{e \in Q \setminus P} f(P) - f(P \cup \{e\})$;
        | Update $P = P \cup \{e^*\}$;
    end
    Return P;

TABLE 1

Algorithmic results given in this paper with their corresponding assumptions, where †: directed or undirected; ∗: $(1-1/e)$-approximation for a surrogate upper bound function; section: $(1-1/e-\epsilon)$-approximation with high probability. Numbers given for algorithmic results are approximation ratios. The assumptions on the number of initial infections are omitted for Problem 4.2.

Problem	Algorithm	Regime	Graph	Result	Section
4.1	Greedy: $f = \hat{\sigma}$	exponential stable	†	∗	5
4.2	Greedy: $f = \mathbb{E}\left[\tilde{\sigma}'\right]$	$R_0 < 1$	$\mathcal{G}(n,p)$	§	6
		$R_0 < 1; \kappa = O(\log n)$	$\mathcal{SBM}(n,\kappa,\boldsymbol{Q})$		7

model are exponentially stable, we find a monotone supermodular upper bound $\hat{\sigma}$ of the objective function. Therefore, we use the upper bound as a surrogate to minimize. According to a well-known result of supermodular minimization [30, section 4], we obtain a greedy algorithm which gives a $(1 - 1/e)$ approximation for the optimum decrease of the surrogate.

For Problem 4.2, when the contact network is sampled from an ER graph or a SBM, the expected number of infections $\mathbb{E}\left[\tilde{\sigma}\right]$ is a $(1 \pm o(1))$-approximation to a monotone supermodular function $\mathbb{E}\left[\tilde{\sigma}'\right]$ with high probability. In this case we optimize the original objective function $\mathbb{E}\left[\tilde{\sigma}\right]$ and treat the differences as errors. A greedy algorithm gives a $(1 - 1/e - o(1) - \epsilon)$-approximation [22] for the optimal decrease in the expected number of infections. We describe the greedy algorithm with a given set function $f(P)$ in Algorithm 1. We note that f is defined as $f = \hat{\sigma}$ for Problem 4.1 and $f = \mathbb{E}\left[\tilde{\sigma}'\right]$ for Problem 4.2.

Table 1 summarizes the algorithmic results given in this paper.

5. Supermodularity for the D-SIR model. In this section, we consider the problem of deleting edges from the current network to mitigate epidemic spread in the D-SIR model. First, we show that the cumulative number of infections is upper bounded by a supermodular function of all the edges in the graph.

Let $\boldsymbol{M}_{-P} \stackrel{\text{def}}{=} \boldsymbol{I} - \boldsymbol{D} + (\boldsymbol{I} - \boldsymbol{X}(0) - \boldsymbol{R}(0))\boldsymbol{B}_{-P}$ denote the transition matrix between $\boldsymbol{x}(1)$ and $\boldsymbol{x}(0)$, where $\boldsymbol{B}_{-P}$ is the matrix obtained from $\boldsymbol{B}$ by setting entries $B_{ij} = 0$ for all edges $(j,i) \in P$.

THEOREM 5.1. *Suppose the D-SIR model satisfies $\|\boldsymbol{M}_{-P}\| < 1$ for all $P \subseteq Q$, where Q is the candidate edge deletion set. The number of infections is then at most*

$$(5.1) \qquad \sigma(P) \leq \hat{\sigma}(P) \stackrel{\text{def}}{=} \mathbf{1}^{\top}(\boldsymbol{M}_{-P} + \boldsymbol{D} - \boldsymbol{I})(\boldsymbol{I} - \boldsymbol{M}_{-P})^{-1}\boldsymbol{x}(0),$$

where $\hat{\sigma}(P)$ is a monotone supermodular function of the edge deletion set P.

Theorem 5.1 shows the existence of a supermodular upper bound for $\sigma(P)$, which motivates us to use it as a surrogate objective function. We employ this fact to design a greedy algorithm, given in Algorithm 1. The algorithm takes as input the function $\hat{\sigma}(P)$, the contact graph G, initial states of nodes, a candidate set Q, and an integer k. The algorithm returns a set P such that $\hat{\sigma}(\varnothing) - \hat{\sigma}(P) \geq (1 - 1/e)(\hat{\sigma}(\varnothing) - \hat{\sigma}(P^*))$, where P^* is the optimal solution for $\sigma(P)$. We note that the performance guarantee is given for $\hat{\sigma}$, instead of σ.

We show that the upper bound function $\hat{\sigma}(P)$ is monotone supermodular.

LEMMA 5.2. *The function $\hat{\sigma}(P)$ is monotone supermodular with respect to the edge deletion set $P \subseteq Q$ when $\|\boldsymbol{M}_{-P}\| < 1$ for all sets $P \subseteq Q$.*

The proof can be found in Appendix C. Then, we prove Theorem 5.1.

Proof of Theorem 5.1. By (3.3) in the D-SIR model, the infected state for any time step t is

$$
\begin{aligned}
\boldsymbol{x}(t+1) &= \boldsymbol{x}(t) - (\boldsymbol{D} + (\boldsymbol{X}(0) + \boldsymbol{R}(0))\boldsymbol{B}_{-P} - \boldsymbol{B}_{-P})\boldsymbol{x}(t) \\
&\quad - (\boldsymbol{X}(t) + \boldsymbol{R}(t) - (\boldsymbol{X}(0) + \boldsymbol{R}(0)))(\boldsymbol{B}_{-P})\boldsymbol{x}(t) \\
&\leq \boldsymbol{x}(t) - (\boldsymbol{D} + (\boldsymbol{X}(0) + \boldsymbol{R}(0))\boldsymbol{B}_{-P} - \boldsymbol{B}_{-P})\boldsymbol{x}(t) \\
&= (\boldsymbol{I} - \boldsymbol{D} + (\boldsymbol{I} - \boldsymbol{X}(0) - \boldsymbol{R}(0))\boldsymbol{B}_{-P})\boldsymbol{x}(t),
\end{aligned}
$$

where the inequality holds because the entries of $(\boldsymbol{X} + \boldsymbol{R})$ are monotonically non-decreasing as a function of time t. By applying this inequality recursively, we have for any time step t

$$
\tag{5.2} \boldsymbol{x}(t) \leq (\boldsymbol{M}_{-P})^t \boldsymbol{x}(0).
$$

Thus, the cumulative number of infections at any time step t is

$$
\begin{aligned}
\|\boldsymbol{m}(t) - \boldsymbol{m}(0)\|_1 &\leq \mathbf{1}^\top (\boldsymbol{I} - \boldsymbol{X}(0) - \boldsymbol{R}(0))(\boldsymbol{B}_{-P}) \sum_{t'=0}^{t-1} \boldsymbol{x}(t') \\
&\leq \mathbf{1}^\top (\boldsymbol{M}_{-P} + \boldsymbol{D} - \boldsymbol{I})(\boldsymbol{I} - \boldsymbol{M}_{-P})^{-1} \boldsymbol{x}(0),
\end{aligned}
$$

where the last inequality is due to the geometric series of matrices. $\square$

Now we provide a sufficient condition for $\|\boldsymbol{M}_{-P}\| < 1$ for all $P \subseteq Q$, which also ensures exponential stability of the healthy state starting from the initial state $\boldsymbol{x}(0)$ and $\boldsymbol{r}(0)$. A similar condition for the SIS model has been given in [19, Theorem 4].

THEOREM 5.3 (sufficient condition for exponentially stable). *If there exists a positive number ϵ such that for any $i \in [n]$, $(1 - x_i(0) - r_i(0)) \sum_{j=1}^n B_{ij} \leq D_i - \epsilon$, then the spectral norm $\|\boldsymbol{M}_{-P}\| \leq 1 - \epsilon$. Moreover, under this condition, $\hat{\boldsymbol{x}} = 0$ is the unique equilibrium and the system is exponentially stable.*

Proof. For simplicity of notation, we use $\boldsymbol{M}$ to denote the matrix $\boldsymbol{M}_{-P}$ for any edge deletion set P. For any $i \in [n]$, we have that the entries of $\boldsymbol{M}$ satisfy

$$
M_{ii} = 1 - D_i, \qquad \sum_{j \neq i} |M_{ij}| \leq (1 - x_i(0) - r_i(0)) \sum_{j=1}^n B_{ij}.
$$

By the Gershgorin circle theorem, the spectral norm of $\boldsymbol{M}$ is bounded by $1 - \epsilon$. Thus, this transition matrix is a contraction mapping. Since we have $\boldsymbol{x}(t+1) \leq \boldsymbol{M}\boldsymbol{x}(t)$

and $\boldsymbol{x}(t) \geq \boldsymbol{0}$ for any t, we obtain $\|\boldsymbol{x}(t+1)\| \leq (1-\epsilon)\|\boldsymbol{x}(t)\|$ for any t. Therefore, by (5.2), $\|\boldsymbol{x}(t)\| \leq (1-\epsilon)^t \|\boldsymbol{x}(0)\| \leq e^{-t\epsilon}\|\boldsymbol{x}(0)\|$, which means the state $\boldsymbol{x}$ converges to the origin with a rate of at least ϵ. Then we conclude that $\hat{\boldsymbol{x}} = 0$ is the unique equilibrium and the considered system is exponentially stable. $\qquad\square$

Remark 5.4. Alternative conditions can be given for Theorem 5.1. Given that $\|\boldsymbol{M}_{-\varnothing}\| < 1$ and $\boldsymbol{M}_{-\varnothing}$ is irreducible, the system is exponentially stable for any deletion edge set $P \subseteq Q$, following from $\|\boldsymbol{M}_{-P}\| < 1$. The reasoning is given as follows: According to the Perron–Frobenius theorem [26], the spectral norm of $\boldsymbol{M}_{-\varnothing}$ is its Perron root $\lambda_{\max}$. Due to Wielandt's theorem [26], the spectral norm, upon deleting edges, is less than or equal to $\lambda_{\max}$.

Convex relaxation is widely used as an approach to design heuristic algorithms or approximation algorithms by using a proper rounding. In addition, the supermodularity of a problem sometimes coincides with the convexity of a relaxed problem; for example, see [27]. However, we show that a direct relaxation of $\hat{\sigma}$ is not a convex function with respect to the edge weights.

Example 5.5. We consider two undirected (bidirectional) graphs G_1 and G_2 which have two copies of the same vertex set $V = \{1, 2, 3\}$. G_1 has edges $\{\{1, 2\}, \{1, 3\}\}$ and G_2 has edges $\{\{1, 2\}, \{2, 3\}\}$. Let the infection rate be $B_{ij} = 1/12$ for each edge (j, i) and the healing rate be $D_i = 1/4$ for each node i. Let $\boldsymbol{B}_{(1)}$ and $\boldsymbol{B}_{(2)}$ be the transition matrices of G_1 and G_2, respectively. We let $\boldsymbol{x}(0) = [1, 0, 0]^\top$, $\boldsymbol{r}(0) = [0, 0, 0]^\top$. Then we have $\boldsymbol{M}_{(i)} = \boldsymbol{I} - \boldsymbol{D} + (\boldsymbol{I} - \boldsymbol{X}(0) - \boldsymbol{R}(0))\boldsymbol{B}_{(i)}$ for $i \in \{1, 2\}$ and define $\hat{\sigma}(\boldsymbol{M}) = \boldsymbol{1}^\top (\boldsymbol{M} + \boldsymbol{D} - \boldsymbol{I})(\boldsymbol{I} - \boldsymbol{M})^{-1}\boldsymbol{x}(0)$. Thus, we obtain $\hat{\sigma}(\boldsymbol{M}_{(1)}) = 2/3$, $\hat{\sigma}(\boldsymbol{M}_{(2)}) = 1/2$, and $\hat{\sigma}(\frac{\boldsymbol{M}_{(1)} + \boldsymbol{M}_{(2)}}{2}) = 3/5$. Thus, $\frac{\hat{\sigma}(\boldsymbol{M}_{(1)}) + \hat{\sigma}(\boldsymbol{M}_{(2)})}{2} < \hat{\sigma}(\frac{\boldsymbol{M}_{(1)} + \boldsymbol{M}_{(2)}}{2})$ yields the nonconvexity result.

6. Supermodularity for the IC-SIR model in random graphs. In the IC-SIR model, the spreading process only uses the activated edges. The activated edges are sampled from the contact network G, which is given as the input of any algorithm. The contagion network $\widetilde{G}$, composed of the activated edges, is randomly generated from G. We assume the set of seeds is first fixed before the contact network and the contagion network are generated. We also assume the initial removed state $\boldsymbol{r}(0) = \boldsymbol{0}$ without loss of generality.

In subsection 6.1 we consider the case where the contact network is a complete graph. Then, the contagion network $\widetilde{G}$ is sampled from an ER random graph $\mathcal{G}(n, p)$, where p is the connection probability. In subsection 6.2 we consider the case where the contact network G is generated from an ER random graph $\mathcal{G}(n, p_1)$, and the contagion network $\widetilde{G}$ is sampled from G with probability p for each edge in G.

6.1. Expected number of infections for a complete contact network. Next we consider the expected number of infections $\mathbb{E}[\tilde{\sigma}(P)]$, which is the objective in Problem 4.2. The expectation is taken over the contagion network sampled from $\mathcal{G}(n, p)$, where the probability $p = d/n$ for some constant $d < 1$. We prove that $\mathbb{E}[\tilde{\sigma}(P)]$ is a $(1 \pm o(1))$-approximation of a supermodular function.

We use τ_1 to denote the event that the contagion network includes connected components with size greater than $L = 9(1-d)^{-2}\ln n$. Let τ_2 be the event that there is at least one component with more than one seed. The complements of these events are denoted as $\bar{\tau}_1$ and $\bar{\tau}_2$, respectively. Let $\rho(P)$ denote the number of infected vertices in connected components that are trees with at least one seed.

THEOREM 6.1. *If the number of seeds $s = O(n^{\frac{1}{3}-c})$ for any constant $c > 0$, then for all $P \subseteq Q$,*

$$\mathbb{E}_{\widetilde{G}}\left[\tilde{\sigma}(P)\right] \approx_{o(1)} \mathbb{E}_{\widetilde{G}}\left[\tilde{\sigma}'(P)\right],$$

where $\mathbb{E}_{\widetilde{G}}\left[\tilde{\sigma}'(P)\right] \stackrel{\text{def}}{=} \mathbb{E}_{\widetilde{G}}\left[\rho(P) \mid \bar{\tau}_1, \bar{\tau}_2\right]$ is a monotone supermodular function.

Theorem 6.1 shows that we can directly optimize the objective function as an approximation to a monotone supermodular function. We employ this fact to design a greedy algorithm, given in Algorithm 1. In particular, Algorithm 1 takes as inputs the function $\mathbb{E}_{\widetilde{G}}\left[\tilde{\sigma}'(P)\right]$, initial states of nodes, a candidate set Q, and an integer k. Assuming $\mathbb{E}_{\widetilde{G}}\left[\tilde{\sigma}(\varnothing)\right] - \mathbb{E}_{\widetilde{G}}\left[\tilde{\sigma}(P^*)\right]$ is at least 1, the algorithm returns a set P such that

$$\mathbb{E}_{\widetilde{G}}\left[\tilde{\sigma}(\varnothing)\right] - \mathbb{E}_{\widetilde{G}}\left[\tilde{\sigma}(P)\right] \geq \left(1 - 1/e - o(1) - \epsilon\right)\left(\mathbb{E}_{\widetilde{G}}\left[\tilde{\sigma}(\varnothing)\right] - \mathbb{E}_{\widetilde{G}}\left[\tilde{\sigma}(P^*)\right]\right),$$

where P^* is the optimum solution for $\mathbb{E}_{\widetilde{G}}\left[\tilde{\sigma}(P)\right]$, and the error term ϵ is from a $(1 \pm \epsilon)$-approximation of $\mathbb{E}_{\widetilde{G}}\left[\tilde{\sigma}'(P)\right]$ and its marginal gains. We will discuss the efficient computation of $\mathbb{E}_{\widetilde{G}}\left[\tilde{\sigma}'(P)\right]$ in section 8.1.

For the sake of analysis, we consider a modified process, which first generates the random graph and then chooses seeds randomly. As long as the information of the graph $\widetilde{G}$ is not revealed, this process is equivalent to the one where seeds are first chosen and then the random graph is generated.

To prove Theorem 6.1, we provide Lemmas 6.2, 6.3, and 6.4. The proofs of these lemmas can be found in Appendix D.

First, we show that the event τ_1 happens with a small probability.

LEMMA 6.2. *The probability that the random graph (with $p < 1/n$) contains a connected component with size greater than $L = 9(1-d)^{-2}\ln n$ is at most*

$$\mathbb{P}\left[\tau_1\right] \leq n^{-2}.$$

We also want to bound the probability of the seeds collision, i.e., the event that two seeds are sampled from the same component. This can be seen as a *balls into bins* problem.

LEMMA 6.3. *The probability that there is at least one component with more than one seed conditioned on event $\bar{\tau}_1$ is*

$$\mathbb{P}\left[\tau_2 \mid \bar{\tau}_1\right] \leq \frac{2s^2 L}{n},$$

where s is the number of seeds as defined in Problem 4.2.

We use $y(P) = \tilde{\sigma}(P) - \rho(P)$ to denote the number of vertices in connected components that contain at least one seed and one cycle. The following lemma follows in a similar manner to [8, Theorem 8.9].

LEMMA 6.4. *Given that the events $\bar{\tau}_1, \bar{\tau}_2$ happen, the expected number of vertices in connected components that contain at least one seed and one cycle satisfies*

$$\mathbb{E}\left[y(P) \mid \bar{\tau}_1, \bar{\tau}_2\right] \leq \frac{sL^2 d^3}{2n(1-d)} \cdot \frac{1}{\mathbb{P}\left[\bar{\tau}_1, \bar{\tau}_2\right]}.$$

Now we prove the main theorem.

Proof of Theorem 6.1. First, we show that $\mathbb{E}_{\widetilde{G}}[\rho(P) \mid \bar{\tau}_1, \bar{\tau}_2]$ is a good approximation of the expected number of infections $\mathbb{E}_{\widetilde{G}}[\tilde{\sigma}(P)]$.

By using the law of total expectation, the expected number of infections $\tilde{\sigma}(P)$ can be divided into three parts as follows,

$$\mathbb{E}_{\widetilde{G}}[\tilde{\sigma}(P)] = \mathbb{E}_{\widetilde{G}}[\tilde{\sigma} \mid \tau_1]\,\mathbb{P}[\tau_1] + \mathbb{E}_{\widetilde{G}}[\tilde{\sigma} \mid \bar{\tau}_1, \tau_2]\,\mathbb{P}[\bar{\tau}_1, \tau_2] + \mathbb{E}_{\widetilde{G}}[\tilde{\sigma} \mid \bar{\tau}_1, \bar{\tau}_2]\,\mathbb{P}[\bar{\tau}_1, \bar{\tau}_2].$$

Then, we give upper bounds of these three terms individually. Note that the first term corresponds to the event that large components exist. By Lemma 6.2, we have

$$\mathbb{E}_{\widetilde{G}}[\tilde{\sigma}(P) \mid \tau_1]\,\mathbb{P}[\tau_1] \leq n \cdot \frac{1}{n^2} = \frac{1}{n}.$$

When the largest component in $\widetilde{G}$ has size at most L, the number of infections is at most sL. By Lemma 6.3, the second term for the seeds collision case is

$$\mathbb{E}_{\widetilde{G}}[\tilde{\sigma}(P) \mid \bar{\tau}_1, \tau_2]\,\mathbb{P}[\bar{\tau}_1, \tau_2] \leq sL \cdot \mathbb{P}[\bar{\tau}_1, \tau_2] \leq \frac{2s^3L^2}{n}.$$

For the third term, we separate the infections in connected components with cycles from the infections in those without cycles, $\tilde{\sigma}(P) = \rho(P) + y(P)$. By Lemma 6.4, we derive

$$\mathbb{E}_{\widetilde{G}}[\tilde{\sigma}(P) \mid \bar{\tau}_1, \bar{\tau}_2]\,\mathbb{P}[\bar{\tau}_1, \bar{\tau}_2] \leq \mathbb{E}_{\widetilde{G}}[\rho(P) \mid \bar{\tau}_1, \bar{\tau}_2]\,\mathbb{P}[\bar{\tau}_1, \bar{\tau}_2] + \frac{sL^2d^3}{2n(1-d)}.$$

Thus, the expected number of infections $\tilde{\sigma}(P)$ is at most

$$\mathbb{E}_{\widetilde{G}}[\tilde{\sigma}(P)] \leq \mathbb{E}_{\widetilde{G}}[\rho(P) \mid \bar{\tau}_1, \bar{\tau}_2] + \frac{sL^2d^3}{2n(1-d)} + \frac{2s^3L^2}{n} + \frac{1}{n}.$$

By Lemmas 6.2 and 6.3, the expected number of infections can be lower bounded by

$$\mathbb{E}_{\widetilde{G}}[\tilde{\sigma}(P)] \geq \mathbb{E}[\tilde{\sigma}(P) \mid \bar{\tau}_1, \bar{\tau}_2]\,\mathbb{P}[\bar{\tau}_1, \bar{\tau}_2] \geq \left(1 - \frac{1}{n^2}\right)\left(1 - \frac{2s^2L}{n}\right)\mathbb{E}[\rho(P) \mid \bar{\tau}_1, \bar{\tau}_2].$$

If the number of seeds $s = O(n^{\frac{1}{3}-c})$ for any constant $c > 0$, then we have

$$\mathbb{E}_{\widetilde{G}}[\tilde{\sigma}(P)] \approx_{o(1)} \mathbb{E}_{\widetilde{G}}[\rho(P) \mid \bar{\tau}_1, \bar{\tau}_2].$$

Given that events $\bar{\tau}_1, \bar{\tau}_2$ happen, the connected components considered in $\rho(P)$ are all trees with exactly one seed. For such connected components, it is easy to check that the number of infected vertices $\rho(P)$ is a monotone supermodular function with respect to the deletion edges. Since the momotonicity and supermodularity of $\rho(P)$ holds for all instances conditioned on $\bar{\tau}_1$ and $\bar{\tau}_2$, it also holds for the expectation $\mathbb{E}[\rho(P) \mid \bar{\tau}_1, \bar{\tau}_2]$. □

6.2. Expected number of infections for a random contact network. In this section, we assume the contact network G is generated from a random graph $\mathcal{G}(n, p_1)$, and the contagion network $\widetilde{G}$ is again sampled from G with probability p for each edge. Further, the set of seeds is fixed before the contact network is generated. Combining these two random processes, the contagion network $\widetilde{G}$ can be seen as being generated from a random graph $\mathcal{G}(n, p_1p)$. If the probability $p_1p < 1$, we have already shown that $\mathbb{E}[\tilde{\sigma}] \approx_{o(1)} \mathbb{E}[\rho \mid \bar{\tau}_1, \bar{\tau}_2]$.

Suppose that the contact network G is realized and given as the problem input. We prove that the conditional expectation of infections given the contact network G satisfies $\mathbb{E}_{\widetilde{G}}[\tilde{\sigma} \mid G] \approx_{o(1)} \mathbb{E}_{\widetilde{G}}[\rho \mid G, \bar{\tau}_1, \bar{\tau}_2]$ with high probability.

THEOREM 6.5. *Suppose the contact network G is generated from a random graph model $\mathcal{G}(n, p_1)$. If the number of seeds $s = O(n^{\frac{1}{3}-c})$ for any constant c, then with high probability, the expected number of infections satisfies that for all $P \subseteq Q$*

$$\mathbb{E}_{\widetilde{G}}\left[\tilde{\sigma}(P) \mid G\right] \approx_{o(1)} \mathbb{E}_{\widetilde{G}}\left[\tilde{\sigma}'(P) \mid G\right],$$

where $\mathbb{E}_{\widetilde{G}}\left[\tilde{\sigma}'(P) \mid G\right] \stackrel{\text{def}}{=} \mathbb{E}_{\widetilde{G}}\left[\rho(P) \mid G, \bar{\tau}_1, \bar{\tau}_2\right]$ is a monotone supermodular function.

Theorem 6.5 shows that $\mathbb{E}_{\widetilde{G}}\left[\sigma(P) \mid G\right]$ can be effectively minimized by a greedy algorithm. In this case, Algorithm 1 takes as inputs the function $\mathbb{E}_{\widetilde{G}}\left[\tilde{\sigma}'(P) \mid G\right]$, a candidate set Q, initial states of nodes, and an integer k. It returns a set P such that $\mathbb{E}_{\widetilde{G}}\left[\tilde{\sigma}(\varnothing) \mid G\right] - \mathbb{E}_{\widetilde{G}}\left[\tilde{\sigma}(P) \mid G\right] \geq (1 - 1/e - o(1) - \epsilon)\left(\mathbb{E}_{\widetilde{G}}\left[\tilde{\sigma}(\varnothing) \mid G\right] - \mathbb{E}_{\widetilde{G}}\left[\tilde{\sigma}(P^*) \mid G\right]\right)$ holds with high probability, where P^* is the optimum solution of $\mathbb{E}_{\widetilde{G}}\left[\tilde{\sigma}(P) \mid G\right]$. Again, the computation of $\mathbb{E}_{\widetilde{G}}\left[\tilde{\sigma}'(P) \mid G\right]$ is discussed in section 8.1.

Proof. In Theorem 6.1, we show that

$$(6.1) \qquad \mathbb{E}\left[\tilde{\sigma}\right] - \mathbb{E}\left[\rho\right] \leq \mathbb{E}\left[\tilde{\sigma}\right] - \mathbb{E}\left[\rho \mid \bar{\tau}_1, \bar{\tau}_2\right] \mathbb{P}\left[\bar{\tau}_1, \bar{\tau}_2\right] \leq \frac{sL^2d^3}{2n(1-d)} + \frac{2s^3L^2}{n} + \frac{1}{n}.$$

By the law of total expectation, we know that

$$(6.2)\ \ \mathbb{E}\left[\tilde{\sigma}\right] - \mathbb{E}\left[\rho\right] = \mathbb{E}_G\left[\mathbb{E}_{\widetilde{G}}\left[\tilde{\sigma} \mid G\right] - \mathbb{E}_{\widetilde{G}}\left[\rho \mid G\right]\right]$$

$$= \mathbb{E}_G\left[\mathbb{E}_{\widetilde{G}}\left[\tilde{\sigma} \mid G\right] - \mathbb{E}_{\widetilde{G}}\left[\rho \mid G, \bar{\tau}_1, \bar{\tau}_2\right] \mathbb{P}_{\widetilde{G}}\left[\bar{\tau}_1, \bar{\tau}_2 \mid G\right]\right]$$

$$- \mathbb{E}\left[\rho \mid \bar{\tau}_1, \tau_2\right] \mathbb{P}\left[\tau_2, \bar{\tau}_1\right] - \mathbb{E}\left[\rho \mid \tau_1\right] \cdot \mathbb{P}\left[\tau_1\right]$$

$$\geq \mathbb{E}_G\left[\mathbb{E}_{\widetilde{G}}\left[\tilde{\sigma} \mid G\right] - \mathbb{E}_{\widetilde{G}}\left[\rho \mid G, \bar{\tau}_1, \bar{\tau}_2\right] \mathbb{P}_{\widetilde{G}}\left[\bar{\tau}_1, \bar{\tau}_2 \mid G\right]\right] - \frac{1}{n} - \frac{2s^3L^2}{n},$$

where the last inequality is due to Lemmas 6.2 and 6.3.

Let the random variable Z be $\mathbb{E}_{\widetilde{G}}\left[\tilde{\sigma} \mid G\right] - \mathbb{E}_{\widetilde{G}}\left[\rho \mid G, \bar{\tau}_1, \bar{\tau}_2\right] \mathbb{P}_{\widetilde{G}}\left[\bar{\tau}_1, \bar{\tau}_2 \mid G\right]$. By combining two inequalities (6.1) and (6.2), we have

$$\mathbb{E}_G\left[Z\right] \leq \mathbb{E}\left[\tilde{\sigma}\right] - \mathbb{E}\left[\rho\right] + \frac{1}{n} + \frac{2s^3L^2}{n} \leq \frac{sL^2d^3}{2n(1-d)} + 2\left(\frac{2s^3L^2}{n} + \frac{1}{n}\right).$$

Setting $\lambda_0 = s^3/n^{1-c}$ for any constant $c > 0$, by Markov's inequality, we have

$$\mathbb{P}\left[Z \geq \lambda_0\right] \leq \frac{\mathbb{E}_G\left[Z\right]}{\lambda_0} \leq \frac{L^2d^3}{2n^c(1-d)s^2} + \frac{6L^2}{n^c},$$

which implies, with high probability, that

$$0 \leq \mathbb{E}_{\widetilde{G}}\left[\tilde{\sigma} \mid G\right] - \mathbb{E}_{\widetilde{G}}\left[\rho \mid G, \bar{\tau}_1, \bar{\tau}_2\right] \mathbb{P}_{\widetilde{G}}\left[\bar{\tau}_1, \bar{\tau}_2 \mid G\right] \leq \lambda_0.$$

To show the approximation result, we need to show that $\mathbb{P}_{\widetilde{G}}\left[\bar{\tau}_1, \bar{\tau}_2 \mid G\right]$ is close to 1 with high probability. By Lemmas 6.2 and 6.3, we have

$$\mathbb{E}_G\left[1 - \mathbb{P}_{\widetilde{G}}\left[\bar{\tau}_1, \bar{\tau}_2 \mid G\right]\right] = 1 - \mathbb{P}\left[\bar{\tau}_1, \bar{\tau}_2\right] = \mathbb{P}\left[\tau_1\right] + \mathbb{P}\left[\tau_2, \bar{\tau}_1\right] \leq \frac{1}{n^2} + \frac{2s^2L}{n}.$$

By using Markov's inequality, we have

$$\mathbb{P}\left[1 - \mathbb{P}_{\widetilde{G}}\left[\bar{\tau}_1, \bar{\tau}_2 \mid G\right] \geq \lambda_0\right] \leq \frac{\mathbb{E}_G\left[1 - \mathbb{P}_{\widetilde{G}}\left[\bar{\tau}_1, \bar{\tau}_2 \mid G\right]\right]}{\lambda_0} \leq \frac{3L}{sn^c},$$

which implies that $\mathbb{P}_{\widetilde{G}}\left[\bar{\tau}_1, \bar{\tau}_2 \mid G\right] \geq 1 - \lambda_0$ with high probability.

Therefore, by using the union bound, we conclude, with probability at least $1 - O(L^2/n^c)$, that

$$\mathbb{E}_{\widetilde{G}}\left[\tilde{\sigma}(P) \mid G\right] \approx_{\lambda_0} \mathbb{E}_{\widetilde{G}}\left[\rho(P) \mid G, \bar{\tau}_1, \bar{\tau}_2\right],$$

where $\lambda_0 = s^3/n^{1-c}$ for any constant $c > 0$. Note that $\mathbb{E}_{\widetilde{G}}\left[\rho(P) \mid G, \bar{\tau}_1, \bar{\tau}_2\right]$ is a supermodular function of the set of removed edges P. □

7. Supermodularity for the IC-SIR model in SBMs. In this section, we consider the IC-SIR model on the network generated from SBMs. We assume there are $\kappa = O(\ln n)$ communities with n nodes. The set of seeds is still fixed before the contact network and the contagion network are generated.

In section 7.1 we investigate the expected number of infections in the IC-SIR model where the contact network is complete and the contagion network $\widetilde{G}$ is sampled from $\mathcal{SBM}(n, \kappa, \boldsymbol{Q})$. In section 7.2 we investigate the case where the contact network G is generated from $\mathcal{SBM}(n, \kappa, \boldsymbol{Q})$ and the contagion network $\widetilde{G}$ is again sampled from G with probability p for each edge.

7.1. Expected number of infections in SBMs. Given a complete contact network, we consider the contagion network $\widetilde{G}$ generated from $\mathcal{SBM}(n, \kappa, \boldsymbol{Q})$, where $\kappa = O(\ln n)$, matrix $\boldsymbol{Q}$ is symmetric, and $\boldsymbol{Q}\mathbf{1}$ is entrywise less than $\frac{1}{n}\mathbf{1}$. In this case, we prove that the expected number of infections $\tilde{\sigma}(P)$ over the contagion network $\widetilde{G}$ is a $(1 \pm o(1))$ approximation of a supermodular function. For simplicity, we define $d_{ij} = nQ_{ij}$ as the expected degree between block i and block j. We further let d_{init} be the maximum intrablock expected degree, defined as $d_{\text{init}} = n \cdot \max(\text{diag}(\boldsymbol{Q}))$, and d_{end} be the maximum expected degree, defined as $d_{\text{end}} = n \cdot \max(\boldsymbol{Q}\mathbf{1})$.

Similar to the analysis of ER random graphs, we define τ_1 as the event that the graph includes connected components with size greater than $L^* = 9(1 - d_{\text{end}})^2 \ln(n\kappa)$. Let τ_2 be the event that there is at least one component with more than one seed. Let events $\bar{\tau}_1$ and $\bar{\tau}_2$ be their complements, respectively. Since the set of seeds is fixed before the generation of the contagion network, the seeds in each community can be seen as sampled from each community uniformly at random afterward.

THEOREM 7.1. *If the number of blocks $\kappa = O(\ln n)$, the number of seeds $s = O(n^{\frac{1}{3}-c})$, and $d_{\text{end}} < 1 - c$ for any absolute constant $c > 0$, then for all $P \subseteq Q$,*

$$\mathbb{E}_{\widetilde{G}}\left[\tilde{\sigma}(P)\right] \approx_{o(1)} \mathbb{E}_{\widetilde{G}}\left[\tilde{\sigma}'(P)\right],$$

where $\mathbb{E}_{\widetilde{G}}\left[\tilde{\sigma}'(P)\right] \stackrel{\text{def}}{=} \mathbb{E}_{\widetilde{G}}\left[\rho(P) \mid \bar{\tau}_1, \bar{\tau}_2\right]$ is a monotone supermodular function of the edge deletion set P.

Theorem 7.1 shows that $\mathbb{E}_{\widetilde{G}}\left[\tilde{\sigma}(P)\right]$ is a close approximation to a monotone supermodular function. Therefore, similar to the result in section 6.1, a greedy algorithm returns a set P which gives a $(1 - 1/e - o(1) - \epsilon)$-approximation of the optimum decrease of the expected number of infections.

We start by proving that τ_1 happens with a small probability.

LEMMA 7.2. *The probability that the contagion network $\widetilde{G}$ has a connected component with size greater than L^* is bounded by*

$$\mathbb{P}\left[\tau_1\right] \leq n^{-2}.$$

Proof. The proof is the same as the proof of Lemma 6.2 except that the number of new leaves Z_t is the sum of $n\kappa$ independent Bernoulli random variables with various probabilities of success, instead of i.i.d. Bernoulli random variables. The result follows from the Chernoff bound by using d_{end} to upper bound the expected degree of each node. ☐

We also obtain the following lemma, which is almost the same as Lemma 6.3.

LEMMA 7.3. *The probability that there is at least one component with more than one seed conditioned on event $\bar{\tau}_1$ is*

$$\mathbb{P}\left[\tau_2 \mid \bar{\tau}_1\right] \leq \frac{2s^2 L^*}{n},$$

where s is the number of seeds.

Proof. We omit the proof since the approach is identical to that of Lemma 6.3. ☐

The key to the proof of Theorem 7.1 is an upper bound for the expected number of nodes in connected components with cycles, which is given by the following lemma.

LEMMA 7.4. *Given that the events $\bar{\tau}_1$ and $\bar{\tau}_2$ happen, the expected number of vertices in connected components which contain at least one seed and one cycle satisfies*

$$\mathbb{E}\left[y(P) \mid \bar{\tau}_1, \bar{\tau}_2\right] \mathbb{P}\left[\bar{\tau}_1, \bar{\tau}_2\right] = O\left(\frac{s(\ln n)^{4m_{\mathrm{bin}}}}{n}\right),$$

where $m_{\mathrm{bin}} \stackrel{\mathrm{def}}{=} 4\left\lceil (1-d_{\mathrm{init}})^2/(1-d_{\mathrm{end}})^2 \right\rceil$.

The proof of Lemma 7.4 is in Appendix E.

Proof of Theorem 7.1. Similar to the proof of Theorem 6.1, we use the bounds given by Lemmas 7.2, 6.3, and 7.4 to prove that $\mathbb{E}_{\widetilde{G}}\left[\tilde{\sigma}(P)\right] \approx_{o(1)} \mathbb{E}_{\widetilde{G}}\left[\rho(P) \mid \bar{\tau}_1, \bar{\tau}_2\right]$. The monotonicity and supermodularity of $\mathbb{E}_{\widetilde{G}}\left[\rho(P) \mid \bar{\tau}_1, \bar{\tau}_2\right]$ again follows from the fact that for all instances of $\widetilde{G}$ where $\bar{\tau}_1$ and $\bar{\tau}_2$ happen, the set function $\rho(P)$ is monotone supermodular. ☐

7.2. Expected number of infections for a random contact network in SBMs. In this section, we assume the contact network G is generated from $\mathcal{SBM}(n,\kappa,\boldsymbol{Q})$, and the contagion network $\widetilde{G}$ is again sampled from G with probability p for each edge. Thus, the contagion network $\widetilde{G}$ can be seen as generated from $\mathcal{SBM}(n,\kappa,p\boldsymbol{Q})$. We assume that the probability matrix $\boldsymbol{Q}$ satisfies that $p\boldsymbol{Q}\mathbf{1}$ is entrywise less than $\frac{1}{n}\mathbf{1}$. In this case, the maximum expected degree $d_{\mathrm{end}} = n \cdot \max(p\boldsymbol{Q}\mathbf{1})$ is less than 1.

We obtain a result similar to Theorem 6.5 for a given contact network generated from $\mathcal{SBM}(n,\kappa,\boldsymbol{Q})$.

THEOREM 7.5. *Suppose the contact network G is generated from a stochastic block model $\mathcal{SBM}(n,\kappa,\boldsymbol{Q})$. If the number of blocks $\kappa = O(\ln n)$, the number of seeds $s = O(n^{\frac{1}{3}-c})$, and $d_{\mathrm{end}} < 1 - c$ for any absolute constant $c > 0$, then with high probability, the expected number of infections satisfies that for all $P \subseteq Q$,*

$$\mathbb{E}_{\widetilde{G}}\left[\tilde{\sigma}(P) \mid G\right] \approx_{o(1)} \mathbb{E}_{\widetilde{G}}\left[\tilde{\sigma}'(P) \mid G\right],$$

where $\mathbb{E}_{\widetilde{G}}[\tilde{\sigma}'(P) \mid G] \overset{\text{def}}{=} \mathbb{E}_{\widetilde{G}}[\rho(P) \mid G, \bar{\tau}_1, \bar{\tau}_2]$ *is a monotone supermodular function of the edge deletion set P.*

Proof. The proof is omitted due to the similarity to the proof of Theorem 6.5. □

Similar to section 6.2, we conclude that a greedy algorithm can be used to effectively minimize $\mathbb{E}_{\widetilde{G}}[\tilde{\sigma}(P) \mid G]$ with high probability.

8. Discussion. In this section, we discuss the efficiency and the robustness of the proposed algorithm for the IC-SIR model and some issues related to real-world networks. In section 8.1, we present an algorithm to efficiently approximate the expected number of infections in the IC-SIR model. In section 8.2, we consider the performance of our algorithm for the case where adversarial nodes are present in the contact network. In section 8.3 we discuss the heterogeneity and temporal nature of the real-world networks that should be considered when applying the algorithms to real scenarios.

8.1. Approximating the expected number of infections in the IC-SIR model. In the IC-SIR model, we need to approximate the expected number of infections $\mathbb{E}[\tilde{\sigma}(P)]$ and its $(1 \pm o(1))$-approximation $\mathbb{E}[\tilde{\sigma}'(P)]$ on a contact network. We do so by sampling contagion networks, similar to the approach used in [9, 6]. For a given set of seeds and a contact network, we repeatedly execute the sampling process defined as follows:

We start by sampling a contagion network and a uniformly sampled terminal node. This process is defined as a success if this terminal node is reachable from the set of seeds. Among R rounds of this process, let R_s be the number of total successes. Then, we use $R_s n / R$ as the estimator of the expected number of infections $\mathbb{E}[\tilde{\sigma}]$.

PROPOSITION 8.1. *For $R = O(\epsilon^{-2} n \ln n)$ rounds of the sampling processes, we can approximate the expected number of infections $\mathbb{E}[\tilde{\sigma}(P)]$ for all sets P and its marginal gains upon edge deletion in the IC-SIR model within a factor of $(1 \pm \epsilon)$ with high probability.*

Proof. For $R = O(\epsilon^{-2} n \ln n)$ rounds, by the Chernoff bound, the estimator $R_s n / R$ is a $(1 \pm \epsilon)$-approximation to the expected number of infections with high probability. Then we consider updating the expected number of infections upon deleting a set P of edges from the contact network. Since all terminals and contagion networks are mutually independent, we can use the same pairs of terminal node and contagion network with different P. We check the reachability of a node in the corresponding contagion network without using edges in P. For a greedy algorithm it suffices to know whether the current edge $e \in Q \setminus P$ is a bridge that separates the terminal and the seeds in the contagion graph to compute the marginal gains $\mathbb{E}[\tilde{\sigma}(P)] - \mathbb{E}[\tilde{\sigma}(P \cup \{e\})]$. Then we use the number of successes without using edges in P and $P \cup \{e\}$ to estimate $\mathbb{E}[\tilde{\sigma}(P)]$ and the marginal gains for all remaining candidate edges. Due to the Chernoff bound, the resulting expectations and the marginal gains are well approximated with high probability. The overall high probability guarantee follows by a union bound. □

Remark 8.2. We can approximate the function $\mathbb{E}[\tilde{\sigma}'(P)] = \mathbb{E}[\rho(P) \mid \bar{\tau}_1, \bar{\tau}_2]$ for all sets P and its marginal gains by modifying the sampling process as follows. If the sampled contagion network has large connected components or seeds collision corresponding to events τ_1, τ_2, respectively, then we resample the contagion network. If the terminal node is contained in a connected component with cycles, then this process is seen as a failure.

8.2. Robustness of the algorithm for the IC-SIR model. We discuss the robustness of the algorithm in the IC-SIR model, in the presence of $O(\ln n)$ oblivious adversarial nodes, which add incident edges to the graph without information about the seeds and the graph except for its direct neighbors. To this end, we consider a contact network generated from an ER graph or a SBM with adversarial perturbation. After the generation of the contact network, the adversarial nodes randomly augment their connections to other nodes. In the following, we give an example of such an adversarial model.

Example 8.3. Consider a contact network constructed as follows: let $G_0 = (V, E_0)$ be a random graph generated from $\mathcal{G}(n, p_1)$, where $p_1 = (\ln n)/n$. Then, we choose $O(\ln n)$ adversarial nodes and add ν nonexisting edges to each of them uniformly at random. Let E_1 be the set of edges added. Thus, the contact network is $G = (V, E_0 \cup E_1)$. The contagion network $\widetilde{G}$ is sampled from G with probability $p < 1/\ln n$. Then, we consider the following two cases.

If the number of edges added to each adversarial node $\nu \geq 4 \ln n$, then by the Chernoff bound, the degree of a node in a random graph is at most $4 \ln n$ with probability at least $1 - (1/n)^{4/5}$. Thus, we can identify these adversarial nodes easily in this case by checking the degree of each node in the contact network. In this case we need to reduce the degree of these nodes to at most $4 \ln n$, which will reduce the network to the following situation.

If the number of edges added to each adversarial node $\nu < 4 \ln n$, then these adversarial nodes could potentially connect several connected components in $\widetilde{G}$. We note that the degree for any adversarial nodes after the edge addition is at most $8 \ln n$ with high probability. Given that $\widetilde{G} - (V, E_1)$ has only connected components of sizes $O(\ln n)$, these adversarial nodes can add a connected component with size at most $O(\ln^3 n)$. Since the number of seeds is $n^{\frac{1}{3} - c}$ for any $c > 0$, the probability that the aforementioned connected component has at least one seed is $O(n^{-\frac{2}{3} - c} \cdot \ln^3 n)$. Therefore, in the worst case, it will induce an $O(n^{-\frac{2}{3} - c} \cdot \ln^6 n) = o(1)$ expected difference between $\tilde{\sigma}$ and ρ.

8.3. Real-world networks. Real-world networks exhibit certain properties that can potentially affect the effectiveness of the algorithm.

Many real-world networks have a skewed degree distribution. Such networks are often well connected through a few hub nodes, which are harmful to the mitigation. In such a case we need to preprocess the network to decrease the maximum degree of the network. An example is given in section 9. By decreasing the maximum degree we can reduce the heterogeneity of the network and eliminate dense clusters.

Real-world contacts can be better characterized by time varying networks, also known as temporal networks. We recall that in our G-SIR model, the indicator random variables $\beta_{ij}(t)$ can be viewed as the product of two independent random variables $\hat{\beta}_{ij}(t)$ and $\widetilde{\beta}_{ij}(t)$, where $\hat{\beta}_{ij}(t)$ determines the contact work and $\widetilde{\beta}_{ij}(t)$ determines the infection in the contact network. Under such a view, the contact network, determined by indicator variables $\hat{\beta}_{ij}(t)$ for all edges, are time-varying. Specifically, these variables are defined to be probabilistic and i.i.d. with respect to time. For general temporal networks we can only take the union of all edges and run the algorithm on the union graph. For the IC-SIR model in temporal networks, we can sample each edge which has a timestamp independently. Each edge is activated when one of its neighbors is infected at the same time. However, for the G-SIR model, such an approach certainly affects the effectiveness of the algorithm. We also note that this approach requires tem-

poral information of all edges beforehand, while in real life such information is likely to be streaming. Further study should be given for temporal networks in future work.

9. Numerical results. In this section we run experiments to show the effectiveness of the proposed algorithms in all considered models. We first study the effectiveness of our greedy algorithms in the D-SIR and IC-SIR models by comparing it to two baseline algorithms, Random [12] and Max-Degree [3].

- Greedy: choose an edge greedily at each round with respect to the objective function, as shown in Algorithm 1.
- Max-Degree: remove an edge incident to the node with the maximum degree at each round.
- Random: remove edges chosen uniformly at random from the network.

9.1. D-SIR simulations. For the D-SIR model, we run experiments on an ER random graph with 500 nodes and the connection probability 0.0249. The edges in the network are always bidirectional, with possibly different infection rates for the two directions. For a generated graph $G(V, E)$ with 3142 edges, we uniformly at random pick the recovery rate $D_i \in [0.28, 0.35]$ for each i and the infection rate $B_{ij} \in [0.011, 0.034]$ for each i, j pair. Then, we uniformly at random sample five nodes from the network as the set of seeds $\mathcal{S}$, whose initial infection probabilities, $x_u(0)$ for all $u \in \mathcal{S}$, are uniformly at random chosen from $[0.8, 0.9]$. For any other node $u \notin \mathcal{S}$, its infection probability is $x_u(0) = 0$. Each node $v \in V$ has an initial removed probability $r_v(0)$ uniformly at random chosen from $[0, 0.05]$. The candidate set Q has 1571 edges which are randomly and uniformly chosen from the edge set E. Further, we run three algorithms to choose the edge deletion set $P \subseteq Q$ satisfying $|P| \leq k$ for $k = 523$. We show the results of the three algorithms for both the number of infections σ and the proposed upper bound of the number of infections $\hat{\sigma}$ in Figure 1.
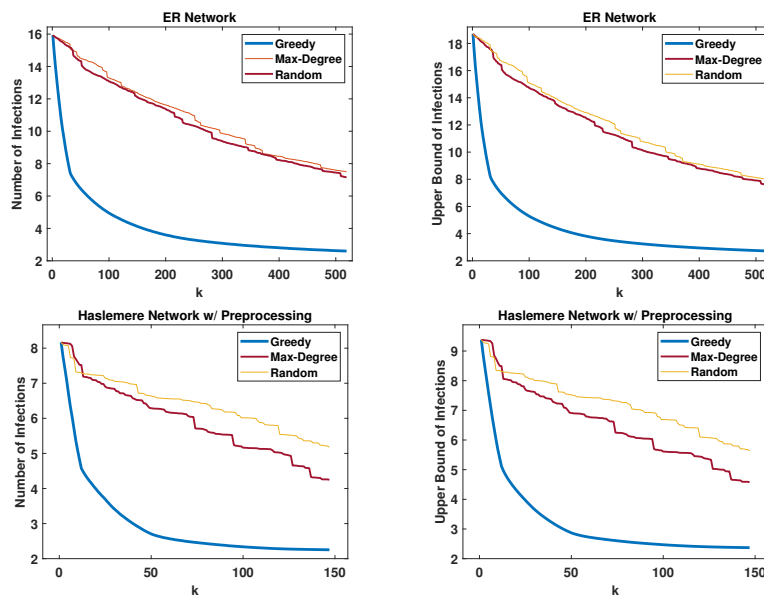

FIG. 1. *In the D-SIR model, the number of infections σ and the proposed upper bound of infections $\hat{\sigma}$ for the edge deletion sets given by Greedy, Max-Degree, and Random algorithms on the ER network and the preprocessed Haslemere contact network. The title of the horizontal axis k is the number of removed edges.*

Next we simulate the dynamics over a real contact network collected in Haslemere, England using mobile technologies [25, 17]. In this example, the contact network has 469 nodes and 1262 edges, with the maximum degree equal to 37. For this contact network, the condition for the exponential stable in Theorem 5.3 can only be guaranteed with small infection rates and large recovery rates. To address a more realistic setting, we preprocess the contact network by greedily removing some edges from the high degree nodes as the Max-Degree algorithm does. This constraint on the max degree of nodes can be implemented in reality by posing restrictions on gathering [29] or shelter-in-place orders [16]. The preprocessed network has 877 edges, with the maximum degree equal to 8. Then we set $D_i = 0.5$ for each i and $B_{ij} \in [0.056, 0.063]$ for each i, j pair. We choose the candidate set Q with 518 edges uniformly at random from the edge set. The results of three algorithms with $k = 219$ on the preprocessed network are shown in Figure 1.

The results clearly show that our algorithm outperforms the two heuristic algorithms in minimizing the number of infections on both the ER network and the Haslemere contact network. We can also observe that the results for the upper bound $\hat{\sigma}$ and the expected number of infections σ behave consistently in these examples.

9.2. IC-SIR simulations. For the IC-SIR model, we implement three algorithms on three different networks, the ER network, the SBM network, and the Haslemere network with preprocessing. The approximations of the expected number of infections after deleting k edges returned by three algorithms are shown in Figure 3.

We first run experiments on an ER network $G(V, E)$ with 1251 edges sampled from $\mathcal{G}(n, p_1)$, where $n = 500$ and $p_1 = 0.01$. The activation probability p for each edge in E is 0.16. The candidate set with 625 edges is chosen from the contact network uniformly at random. We randomly and uniformly choose 5 nodes in V as seeds.

Second, we consider a SBM network sampled from $\mathcal{SBM}(n, \kappa, \boldsymbol{Q})$, where $n = 100$, $\kappa = 5$, and the entries Q_{ij} for $i, j \in [\kappa]$ are set as $Q_{ij} = 0.023$ for $i = j$ and $Q_{ij} \in [0.0036, 0.0046]$, chosen uniformly at random. The generated contact network has 995 edges. The candidate set with 497 edges is chosen from the contact network uniformly at random. The activation probability p is set as 0.21. The set of seeds with 5 nodes is chosen from the network uniformly at random.

Third, we run experiments on the Haslemere network with preprocessing, introduced in section 9.1. For this network, we let the connection probability be 0.179 and uniformly sample half of the edges in the contact networks as the candidate sets.

The results show that our algorithm outperforms the two considered heuristics in minimizing the expected number of infections in all three networks. Since we have proved that the expected number of infections $\mathbb{E}[\tilde{\sigma}]$ is a $(1 \pm o(1))$-approximation of $\mathbb{E}[\tilde{\sigma}']$, we also apply Algorithm 1 with the input $\mathbb{E}[\tilde{\sigma}]$ to the preprocessed Haslemere network. In Figure 2, we compare the results returned by Algorithm 1 with the input $\mathbb{E}[\tilde{\sigma}']$ (Greedy) and $\mathbb{E}[\tilde{\sigma}]$ (Greedy-2), respectively. The results show that the performance of Greedy and Greedy-2 are almost identical.

Then we show the performance of Greedy-2 on the Haslemere network with and without preprocessing in Figure 3. In the Haslemere network without preprocessing, the Greedy-2 algorithm is marginally better than the two heuristics. Similar degeneration of the effectiveness of the greedy algorithm is observed in networks with power-law degree distributions. These examples highlight that our algorithm does not perform well on networks with relatively high max degree, as expected from the discussion in section 8.2.
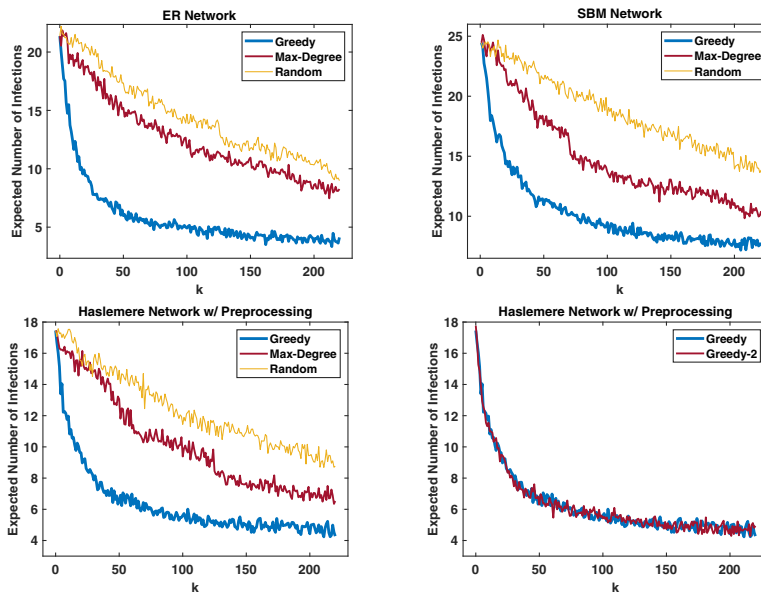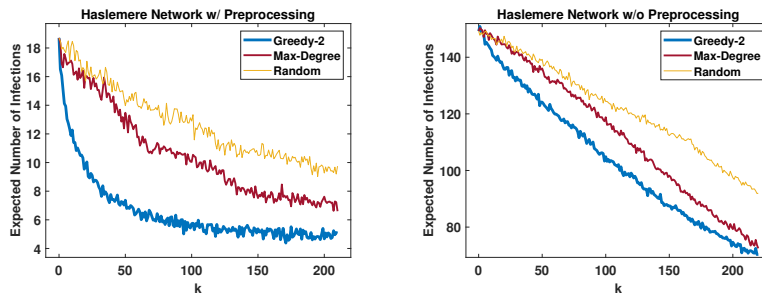
FIG. 2. *In the IC-SIR model, the expected number of infections $\mathbb{E}[\tilde{\sigma}]$ for the edge deletion sets given by Greedy, Max-Degree, and Random algorithms on the ER network, the contact network generated from the SBM model, the Haslemere contact network with preprocessing. The bottom-right figure compares the results of Algorithm 1 with the input $\mathbb{E}[\tilde{\sigma}']$ (Greedy) and $\mathbb{E}[\tilde{\sigma}]$ (Greedy-2). The title of the horizontal axis k is the number of removed edges.*

FIG. 3. *In the IC-SIR model, the expected number of infections $\mathbb{E}[\tilde{\sigma}]$ for the edge deletion sets given by Greedy-2, Max-Degree, and Random algorithms on the Haslemere contact network with and without preprocessing. The title of the horizontal axis k is the number of removed edges.*

9.3. Comparison of two proposed algorithms. Here we compare the performance of the greedy algorithm in Algorithm 1 when using objectives $\hat{\sigma}$ from the D-SIR model versus $\mathbb{E}[\tilde{\sigma}]$ from the IC-SIR model. We implement the results of the algorithms on the G-SIR model to evaluate their effectiveness, while also considering computation costs.

For both models we run the greedy algorithm on an ER network with 50 nodes and connection probability $p = 0.08$. The generated network has 98 edges. We choose 5 fixed nodes as the set of seeds. We run the G-SIR simulations 15000 times. According to the Hoeffding's bound, we have a probability of at least 0.9 to obtain an error less than or equal to 0.25. For the D-SIR model, Figure 4 shows the expected number of infections in the G-SIR model, the upper bound $\hat{\sigma}$, and the number of infections σ in
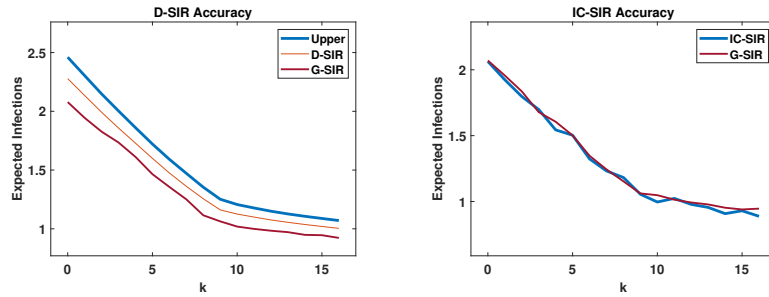
FIG. 4. *The expected number of infections for the edge deletion sets returned by the greedy algorithm in Algorithm* 1 *using objectives* $\hat{\sigma}$ *in the D-SIR model and* $\mathbb{E}[\tilde{\sigma}']$ *in the IC-SIR model on the ER network. We compare the results of the algorithm by implementing them on the G-SIR model with the approximations* $\hat{\sigma}$ *(Upper),* σ *(D-SIR), and the estimated* $\mathbb{E}[\tilde{\sigma}]$ *(IC-SIR).*

the D-SIR model. In the IC-SIR model, the activation probability of an edge (j,i) is set to be the cumulative infection rate from node j to i, which is $(1 - (1 - D_j)(1 - B_{ij}))^{-1}B_{ij}$. The figure shows the expected number of infections estimated in the G-SIR model and the corresponding IC-SIR model. We show that the two algorithms perform similarly in minimizing the number of infections. The IC-SIR model gives a better estimation for the G-SIR model in terms of the expected number of infections.

In our implementations, the running time of the algorithm for the D-SIR model is $O(n^3 + k(n^2 + qn))$; the running time of the algorithm for the IC-SIR model with the input $\mathbb{E}[\tilde{\sigma}']$ or $\mathbb{E}[\tilde{\sigma}]$ is $O(kqn^2 \ln n/\epsilon^2)$. Since we let $q = O(n)$ and $k = O(q)$ in all the experiments, the algorithm for the D-SIR model runs much faster than the algorithm for the IC-SIR model. We note that the running time of the latter algorithm can be improved by designing data structures to efficiently maintain the connectivity of the generated contagion networks upon edge deletion. Discussion about this topic is beyond the scope of this paper.

10. Conclusion. In this paper we studied the expected number of infections in a Markov chain SIR model and its two instantiations: the D-SIR model and the IC-SIR model. We focused on the problem of minimizing the number of infections by deleting edges from a given candidate edge deletion set in these models. We proposed a greedy algorithm for each model and provided guarantees for the performance of these algorithms under certain conditions. For the D-SIR model, we proposed a monotone supermodular upper bound for the number of infections when the system is exponentially stable. For the IC-SIR model we proved that the expected number of infections is a close approximation to a monotone supermodular function with high probability when (1) the contact network is sampled from an ER graph or a SBM with a small number of blocks; (2) the number of initial infections are relatively small; and (3) the average degree in the contagion network is less than 1. We validated the accuracy of the two models and the effectiveness of the proposed algorithms by running experiments on several synthesized networks and a real contact network. The results showed that the two models approximate the Markov chain SIR model well. The simulations show that the proposed algorithms can be used to minimize the number of infections effectively when the network does not contain highly connected nodes. Therefore, from this we learn that in order to enable effective targeted mitigation strategies based on contact tracing, the maximum degree in the network needs to be contained by implementing general measures.

Future work can consider the edge removing problem in a time-varying contact network. The problem can be formulated as an offline or an online problem. In addition, the contact network is assumed to be known without uncertainty in this paper. Designing adaptive algorithms to update the edge deletion set using new observations of the network or the process is another interesting future direction.

Appendix A. Proofs from section 3.

Proof of Proposition 3.1. Consider any directed edge from node j to node i in the G-SIR model. If the node j is infected, then this edge will be activated with probability $p_{ij} = \sum_{t=0}^{\infty}((1-D_j)^t(1-B_{ij})^t B_{ij}) = (1-(1-D_j)(1-B_{ij}))^{-1}B_{ij}$. We note that this activation probability is fixed regardless of the time or ordering of the infections. Thus, we can first sample each edge with probability p_{ij} independently. Then, the expected number of infected nodes in the sampled subgraph is exactly the same as the expected number of infected nodes in the G-SIR model. This second process is equivalent to an IC-SIR model with infected rates p_{ij} for all edges. □

Proof of Proposition 3.2. From Proposition 3.1 we know that a symmetric directed IC-SIR with $p_{ij} = (1-(1-D_j)(1-B_{ij}))^{-1}B_{ij}$ has the same expected number of infections as the symmetric G-SIR model. It remains to be shown that an undirected IC-SIR with p_{ij} for the undirected edge $\{i,j\}$ has the same number as the symmetric IC-SIR.

Given an undirected IC-SIR model, we consider an edge $\{i,j\}$. Since any infected node recovers in one time step, each edge can change the number of infections at the first time that any of its incident nodes are infected. If i and j are infected at the same time, then the activation probability of $\{i,j\}$ does not matter and can be set to p_{ij}. Otherwise only one of (i,j) and (j,i) in the symmetric IC-SIR model matters, with activation probability p_{ij}. Therefore, by setting p_{ij} as described in Proposition 3.2, the undirected IC-SIR model has the same expected number of infections as the symmetric directed IC-SIR model. □

Appendix B. Proof from section 4.
For the D-SIR model, we consider the case where

$$(B.1) \qquad D_i = 0 \ \forall i \in V; \quad B_{ij} > 0 \ \forall (j,i) \in E; \quad Q = E.$$

For the IC-SIR model, we consider the case where

$$(B.2) \qquad B_{ij} = 1 \ \forall (j,i) \in E; \quad Q = E,$$

such that the contagion network is equivalent to the contact network. With these parameters, both Problems 4.1 and 4.2 become equivalent to the graph partitioning problem formulated in [15].

In particular, we study the following decision problem.

Problem B.1. Given a graph $G = (V, E)$ with n nodes, a vector $\boldsymbol{s} \in \{0,1\}^n$ which defines the set of seeds $\mathcal{S} = \text{supp}(\boldsymbol{s})$, a candidate edge deletion set $Q \subseteq E$ with $|Q| = q$, and two integers $0 < k \le q$ and $0 \le z \le n$, decide whether or not there exists an edge set P, with $|P| = k$, such that after removing edges in P, the number of nodes that are reachable from the seeds is less or equal to z.

The hardness of Problem B.1 follows from the NP-completeness of the *minimum bisection of* 3-*regular graphs.*

LEMMA B.2 ([11, Theorem 2.12]).   *The problem of deciding whether or not a d-regular graph has a bisection of size b or less is* NP-*complete, whenever $d \geq 3$ and $b = n^\epsilon$ for any fixed $\epsilon \in (0, 1)$.*

Proof of Theorem 4.4. We consider Problem 4.1 under the condition (B.1) or Problem 4.2 under the condition (B.2).

We construct an instance as follows. For a 3-regular graph $\widehat{G} = (V, E)$ with n nodes, where n is a even number, we add 3 copies of a star graph G^* to it, denoted $G^*_{(1)} = (V_1, E_1)$, $G^*_{(2)} = (V_2, E_2)$, and $G^*_{(3)} = (V_3, E_3)$. Each one of the $(n + 1)$-node star graphs has n leaves supported on V. The central nodes of the 3 stars are denoted as $v_{(1)}$, $v_{(2)}$, and $v_{(3)}$, respectively.

Then we construct the graph $G' = \widehat{G} + G^*_{(1)} + G^*_{(2)} + G^*_{(3)}$. We consider an instance of Problem B.1 with the input $G = G'$, $\mathcal{S} = \{v_{(1)}, v_{(2)}, v_{(3)}\}$, $Q = E \cup E_1 \cup E_2 \cup E_3$, $k = b + \frac{3n}{2}$, and $z = 3 + \frac{n}{2}$.

We provide an interactive proof.

Completeness: If there is a bisection of $\widehat{G}$ with size b, then by definition there is a partition of nodes S and $\bar{S} \stackrel{\text{def}}{=} V \backslash S$, such that the number of edges in between is $b = |C(S, \bar{S})|$, and $|S| = |\bar{S}| = n/2$. By cutting b edges between S and $\bar{S}$, and all the edges $(u, v_{(i)})$, for all $u \in S$ and $i \in \{1, 2, 3\}$, we separate $n/2$ nodes from the rest of the graph. Therefore, the number of nodes reachable from the seeds is equal to $3 + \frac{n}{2}$.

Soundness: It suffices to prove that if there is no bisection of size b or less for $\widehat{G}$, then there is no such set P, with $|P| = b + \frac{3n}{2}$, such that by removing edges in P, $\frac{n}{2}$ nodes are no longer reachable from the seeds. We prove it by contradiction. We assume the bisection of size b for $\widehat{G}$ does not exist. If there exists such a set P for the constructed instance of Problem B.1, then P contains a cut $C(S, \bar{S})$ of graph G' with at most $b + \frac{3n}{2}$ edges, with $|S| \geq n/2$ and $S \cap \mathcal{S} = \varnothing$. If $|S| = n/2$, then $P \cap E$ contains a bisection of $\widehat{G}$ of size less than or equal to b. Therefore, we have constructed a contradiction. If $|S| > n/2$, one can always find an vertex $u \in S$, such that $|C(S \backslash \{u\}, \bar{S} \cup \{u\})| \leq b + \frac{3n}{2}$, since u has at most three neighbors in S and at least three neighbors in $\bar{S}$. By repeating this process we arrive at a new cut $C' \stackrel{\text{def}}{=} C(S', \bar{S}')$ of G' such that $|C(S', \bar{S}')| \leq b + \frac{3n}{2}$, $|S| = n/2$, and $S \cap \mathcal{S} = \varnothing$. Then $C' \cap E$ is a bisection of $\widehat{G}$ with size b or less. Again, we obtain a contradiction, concluding the proof. $\qquad \square$

Remark B.3. In the proof of Theorem 4.4 we constructed hard instances with $R_0 = 6n/(n + 3)$. We note that similar hard instances can be constructed when $\widehat{G}$ is a d-regular graph. In addition, we can attach new nodes to each node u in $\widehat{G}$ by adding probabilistic edges (which are not in Q). By applying these two techniques, we can construct hard instances for any given $R_0 > 0$ in an IC-SIR model.

Appendix C. Proof from section 5.

Proof of Lemma 5.2. Since $\|\boldsymbol{M}_{-P}\| < 1$ for any set $P \subseteq Q$, the inverse of matrix $(\boldsymbol{I} - \boldsymbol{M}_{-P})$ always exists. We note that $(\boldsymbol{I} - \boldsymbol{M}_{-P})$ is an M-matrix for any edge deletion set P, which implies that its inverse $(\boldsymbol{I} - \boldsymbol{M}_{-P})^{-1}$ is entrywise nonnegative by Lemma 2.1.

Consider adding an edge (j, i) to the deletion set P. By the Sherman–Morrison formula [5], we have

$$(\boldsymbol{I} - \boldsymbol{M}_{-P} + (1 - x_i(0) - r_i(0))B_{ij}\boldsymbol{e}_i\boldsymbol{e}_j^\top)^{-1}$$
$$= (\boldsymbol{I} - \boldsymbol{M}_{-P})^{-1} - c \cdot \frac{(\boldsymbol{I} - \boldsymbol{M}_{-P})^{-1}\boldsymbol{e}_i\boldsymbol{e}_j^\top(\boldsymbol{I} - \boldsymbol{M}_{-P})^{-1}}{1 + c\boldsymbol{e}_j^\top(\boldsymbol{I} - \boldsymbol{M}_{-P})^{-1}\boldsymbol{e}_i},$$

in which $c = (1 - x_i(0) - r_i(0))B_{ij}$ is a nonnegative number. Let $H(P) \overset{\text{def}}{=} (\boldsymbol{I} - \boldsymbol{M}_{-P})^{-1}$. Thus, we have that $H(P)$ is monotonically nonincreasing as a function of adding edges to the deletion set.

For an edge deletion set P and any edge $e = (j, i) \in Q \setminus P$, we derive

$$H(P) - H(P \cup (j, i)) = \int_0^1 c \cdot (\boldsymbol{I} - \boldsymbol{M}_{-P})^{-1} \boldsymbol{e}_i \boldsymbol{e}_j^\top (\boldsymbol{I} - \boldsymbol{M}_{-P})^{-1} t \, \mathrm{d}t \, .$$

Considering any set P_2 satisfying $P_1 \subseteq P_2$ and $e \notin P_2$, we have the same result except that $\boldsymbol{M}_{-P}$ is replaced by $\boldsymbol{M}_{-P_2}$. Since we proved that the function $H(S)$ is monotonically nonincreasing, we have that

$$H(P_1) - H(P_1 \cup (j, i)) \geq H(P_2) - H(P_2 \cup (j, i)) \, .$$

Therefore, we proved that function $H(P)$ is entrywise nonnegative, and entrywise monotone supermodular with respect to the set of deletion edges.

In addition, we observe that $(\boldsymbol{M}_{-P} + \boldsymbol{D} - \boldsymbol{I})$ is an entrywise nonnegative nonincreasing modular function of P. Therefore, the product of $(\boldsymbol{M}_{-P} + \boldsymbol{D} - \boldsymbol{I})$ and $H(P)$ is entrywise monotone supermodular. Since the entries of $\boldsymbol{1}$ and $\boldsymbol{x}(0)$ are all nonnegative, the function $\hat{\sigma}(P)$ is monotone supermodular. $\qquad \square$

Appendix D. Proofs from section 6.

Proof of Lemma 6.2. We couple the random graph model with the branching process [4, section 11.3], which starts from a root and samples n i.i.d. Bernoulli random variables with probability p for an active leaf to generate its children at each time step. Let Y_t be the number of active leaves at time step t. If $Y_{t-1} > 0$, then let Z_t be the number of new leaves generated at time t. Thus, each random variable Z_t is the sum of n i.i.d. Bernoulli random variables with probability p. Then, we consider the recursion $Y_0 = 1$ and $Y_t = Y_{t-1} + Z_t - 1$ for all $t \geq 1$. The number of vertices generated by this process is an upper bound of the size of the corresponding connected component in the random graph model.

Let T be the stopping time of this branching process, which is also the number of vertices generated in this process. Then, we have that the probability of generating more than t vertices is

$$\mathbb{P}\left[T > t\right] \leq \mathbb{P}\left[Y_t > 0\right] = \mathbb{P}\left[\sum_{i=1}^t Z_t > t\right] \, .$$

Since $\{Z_i\}_{i=1}^t$ are independent and each Z_i are the sum of n i.i.d. Bernoulli random variables, by the Chernoff bound, we have

$$\mathbb{P}\left[\sum_{i=1}^t Z_t > t\right] \leq \exp\left(-\frac{(1-d)^2 t}{3}\right) \, .$$

By setting $t = L$, we get $\mathbb{P}\left[T > L\right] \leq n^{-3}$. Using the union bound, we get the probability of any node in a connected component greater than L is at most n^{-2}. $\quad \square$

Proof of Lemma 6.3. Given that the event $\bar{\tau}_1$ happens, we can put all connected components in bins with the same capacity L. Let m be the number of bins we used. All connected components can be placed into less than $\lceil \frac{2n}{L} \rceil$ bins because we can place the connected components in a way that each bin is at least half full, with possibly one extra bin.

For any fixed bin, the probability that a seed is sampled from any components from this bin is at most L/n. By using the union bound, the probability of having at least one bin with more than one seed is

$$\mathbb{P}\left[\tau_2 \mid \bar{\tau}_1\right] \leq \binom{s}{2} m \left(\frac{L}{n}\right)^2 \leq \frac{2s^2 L}{n}. \qquad \square$$

Proof of Lemma 6.4. Since deleting edges might remove some cycles in the network, we can upper bound $y(P)$ by $y(\varnothing)$. We first estimate the number of cycles x in the random graph $\mathcal{G}(n, p)$. According to the analysis in [8, Theorem 8.9], the number of possible length-z cycles is $\binom{n}{z}\frac{(z-1)!}{2}$. The probability that such a length-z cycle appears in the random graph is p^z. Therefore, the expected number of cycles in the random graph is at most

$$(\text{D.1}) \qquad \mathbb{E}_{\widetilde{G}}\left[x\right] \leq \sum_{z=3}^{n} \binom{n}{z}\frac{(z-1)!}{2} p^z \leq \frac{1}{2}\sum_{z=3}^{n} d^z \leq \frac{d^3}{2(1-d)},$$

where the second inequality follows from the fact that $p = d/n$.

Suppose the event $\bar{\tau}_1$ happens, then all connected components in the contagion network have size at most L. Since all seeds are chosen uniformly at random, the probability that each seed is sampled from a connected component with cycles is at most xL/n. Thus, for any fixed random graph, the expected number of connected components with at least one seed and one cycle is at most sxL/n, where s is the number of seeds. Since the size of each component is at most L, the expectation of $y(P)$ over the set of seeds $\mathcal{S}$ is

$$\mathbb{E}_{\mathcal{S}}\left[y(P) \mid \widetilde{G}, \bar{\tau}_1\right] \leq L \cdot \frac{sxL}{n}.$$

Therefore, we have

$$\mathbb{E}\left[y(P) \mid \bar{\tau}_1, \bar{\tau}_2\right]\mathbb{P}\left[\bar{\tau}_1, \bar{\tau}_2\right] \leq \mathbb{E}\left[y(P) \mid \bar{\tau}_1\right]\mathbb{P}\left[\bar{\tau}_1\right] \leq \mathbb{E}_{\widetilde{G}}\left[L \cdot \frac{sxL}{n} \mid \bar{\tau}_1\right]\mathbb{P}\left[\bar{\tau}_1\right]$$

$$= \frac{sL^2}{n}\mathbb{E}_{\widetilde{G}}\left[x \mid \bar{\tau}_1\right]\mathbb{P}\left[\bar{\tau}_1\right] \leq \frac{sL^2}{n}\frac{d^3}{2(1-d)},$$

where the second inequality follows from taking the expectation over seeds and the last inequality holds because $\mathbb{E}_{\widetilde{G}}\left[x \mid \bar{\tau}_1\right]\mathbb{P}\left[\bar{\tau}_1\right] \leq \mathbb{E}_{\widetilde{G}}\left[x\right] \leq d^3/(2-2d)$. $\qquad \square$

Appendix E. Proof from section 7.

Proof of Lemma 7.4. An instance of the stochastic block model $\mathcal{SBM}(n, \kappa, \boldsymbol{Q})$ can be viewed as a composition of κ random graphs and $\kappa(\kappa-1)/2$ random bipartite graphs. First, we generate all the intrablock edges using probabilities given by $\text{diag}(\boldsymbol{Q})$. By Lemma 6.2 we obtain that each block contains a connected component of size greater than $L_{\text{init}} = 9(1-d_{\text{init}})^{-2}\ln n$ with probability at most n^{-2}. By the union bound, the probability that there exists a connected component greater than $L_{\text{init}} = 9(1-d_{\text{init}})^{-2}\ln n$ in the whole network is less or equal to κn^{-2}.

Similar to the analysis of the ER random graphs, we place the connected components into bins with the same capacity L_{init}, and we refer to these bins as small bins

hereafter. In addition, according to (D.1), the expected number of cycles given that $\bar{\tau}_1$ happens satisfies

$$\mathbb{E}\left[x \mid \bar{\tau}_1\right] \leq \kappa \cdot \frac{(d_{\text{init}})^3}{2(1 - d_{\text{init}})} \cdot \frac{1}{\mathbb{P}\left[\bar{\tau}_1\right]} .$$

We denote the current network by $\widetilde{G}^{(\kappa)}$, indicating that κ random graphs have been generated. Then we add the random bipartite graphs to the graph until we obtain $\widetilde{G} = \widetilde{G}^{(\kappa^2)}$.

Suppose at some point we have constructed the graph $\widetilde{G}^{(\kappa+\alpha-1)}$, with κ random graphs and $\alpha - 1$ random bipartite graphs generated. We consider the expected incremental number of cycles added to the graph when we add the αth random bipartite graph $\widetilde{G}^{(\alpha)}_{\text{bip}}$ with connecting probability $\boldsymbol{Q}_{ij}$. The set of new cycles includes cycles that only use edges in $\widetilde{G}^{(\alpha)}_{\text{bip}}$ and cycles formed by connecting existing connected components in $\widetilde{G}^{(\kappa+\alpha-1)}$.

The number of cycles using only edges in $\widetilde{G}^{(\alpha)}_{\text{bip}}$ is given by

$$\mathbb{E}\left[x^{\text{bip}}_{(\alpha+1)} \mid \bar{\tau}_1\right] \mathbb{P}\left[\bar{\tau}_1\right] \leq \sum_{z=2}^{\lceil L^*/2 \rceil} \binom{n}{z}^2 \frac{(z!)^2}{2} \left(\boldsymbol{Q}_{ij}\right)^z \leq \frac{1}{2} \sum_{z=2}^{\lceil L^*/2 \rceil} \left(n\boldsymbol{Q}_{ij}\right)^z \leq \frac{(d_{ij})^2}{2(1 - d_{ij})} .$$

Then we consider the expected number of cycles formed by connecting exactly z distinct existing connected components. For any z connected components $\{C_i\}_{i=1}^z$ with sizes $\{\ell_i\}_{i=1}^z$, respectively, let $\mathcal{C}(C_1, C_2, \cdots, C_z)$ be the event that they form a cycle. Then, we have

$$\mathbb{P}\left[\mathcal{C}(C_1, C_2, \cdots, C_z)\right] \leq (z-1)! \cdot \frac{\ell_z \ell_1}{n} \prod_{i=1}^{z-1} \frac{\ell_i \ell_{i+1}}{n}$$

$$\leq \frac{(z-1)!}{n^z} \prod_{i=1}^z \ell_i^2 \leq \frac{(z-1)!}{n^z} \left(\frac{\sum_{i=1}^z \ell_i}{z}\right)^{2z} \leq \frac{(z-1)!}{n^z} \left(\frac{L^*}{z}\right)^{2z},$$

where the last inequality is due to the generalized mean inequality. Note that this upper bound does not depend on the sizes of components, which can be denoted by $\mathbb{P}\left[\mathcal{C}_z\right]$.

For any fixed z, we use N_z to denote the number of choices for these z components. Let m_{bin} be the number of small bins needed to cover a large bin, which is at most

$$m_{\text{bin}} \leq \left\lceil \frac{L^*}{L_{\text{init}}/2} \right\rceil \leq 2 \left\lceil \left(\frac{1 - d_{\text{init}}}{1 - d_{\text{end}}}\right)^2 \frac{\ln(\kappa n)}{\ln n} \right\rceil \leq 4 \left\lceil \left(\frac{1 - d_{\text{init}}}{1 - d_{\text{end}}}\right)^2 \right\rceil .$$

Then, we consider two cases regrading the number of components z.

If the number of components $1 \leq z \leq m_{\text{bin}}$, the number of choices N_z satisfies

$$N_z \leq \binom{n\kappa}{z} \binom{L^*}{z},$$

where the first factor upper bounds the number of choices of z small bins in the graph, and the second factor upper bounds the number of choices of z connected components

from the z bins, which are all in a new connected component whose size is at most L^*. Then the expected number of cycles for this case is

$$\mathbb{E}\left[x'_{(\alpha)} \mid \bar{\tau}_1\right] \mathbb{P}\left[\bar{\tau}_1\right] \leq \sum_{z=1}^{m_{\text{bin}}} N_z \cdot \mathbb{P}\left[\mathcal{C}_z\right] \leq \sum_{z=1}^{m_{\text{bin}}} \binom{n\kappa}{z}\binom{L^*}{z} \cdot \mathbb{P}\left[\mathcal{C}_z\right]$$

$$\leq \sum_{z=1}^{m_{\text{bin}}} \frac{(n\kappa)^z}{z!} \cdot \frac{(L^*)^z}{z!} \cdot \frac{(z-1)!}{n^z}\left(\frac{L^*}{z}\right)^{2z} = \sum_{z=1}^{m_{\text{bin}}} \frac{(L^*)^{3z} \cdot \kappa^z}{z^{2z+1} \cdot z!} \leq \sum_{z=1}^{m_{\text{bin}}} (L^*)^{3m_{\text{bin}}} \cdot \kappa^{m_{\text{bin}}}.$$

If the number of components $z > m_{\text{bin}}$, then N_z satisfies

$$N_z \leq \binom{n\kappa}{m_{\text{bin}}}\binom{L^*}{z}.$$

The expected number of cycles in this case is

$$\mathbb{E}\left[x''_{(\alpha)} \mid \bar{\tau}_1\right] \mathbb{P}\left[\bar{\tau}_1\right] \leq \sum_{z=m_{\text{bin}}+1}^{L^*} N_z \cdot \mathbb{P}\left[\mathcal{C}_z\right]$$

$$\leq \sum_{z=m_{\text{bin}}+1}^{L^*} \frac{(n\kappa)^{m_{\text{bin}}}}{m_{\text{bin}}!} \cdot \frac{(L^*)^z}{z!} \cdot \frac{(z-1)!}{n^z}\left(\frac{L^*}{z}\right)^{2z}$$

$$= \sum_{z=m_{\text{bin}}+1}^{L^*} \frac{(L^*)^{3z} \cdot \kappa^{m_{\text{bin}}}}{n^{z-m_{\text{bin}}} \cdot z^{2z+1} \cdot m_{\text{bin}}!} \leq \sum_{z=m_{\text{bin}}+1}^{L^*} (L^*)^{3m_{\text{bin}}} \cdot \kappa^{m_{\text{bin}}}.$$

Combining these two cases, we have that the expected number of new cycles formed by connecting existing connected components is

$$\mathbb{E}\left[x^{\text{con}}_{(\alpha)} \mid \bar{\tau}_1\right] \mathbb{P}\left[\bar{\tau}_1\right] \leq \mathbb{E}\left[x'_{(\alpha)} + x''_{(\alpha)} \mid \bar{\tau}_1\right] \mathbb{P}\left[\bar{\tau}_1\right] \leq (L^*)^{3m_{\text{bin}}+1} \cdot \kappa^{m_{\text{bin}}},$$

and the total expected number of cycles that are created upon adding the random bipartite graph $\widetilde{G}^{(\alpha)}_{\text{bip}}$ is

$$\mathbb{E}\left[x_{(\alpha)} \mid \bar{\tau}_1\right] \mathbb{P}\left[\bar{\tau}_1\right] = \mathbb{E}\left[x^{\text{bip}}_{(\alpha)} + x^{\text{con}}_{(\alpha)} \mid \bar{\tau}_1\right] \mathbb{P}\left[\bar{\tau}_1\right] = \frac{(d_{ij})^2}{2(1-d_{ij})} + (L^*)^{3m_{\text{bin}}+1} \cdot \kappa^{m_{\text{bin}}}.$$

By adding all random bipartite graphs, we obtain that the expected number of cycles in $\widetilde{G}$ is

$$\mathbb{E}\left[x \mid \bar{\tau}_1\right] \mathbb{P}\left[\bar{\tau}_1\right] = \kappa \cdot \frac{(d_{\text{init}})^3}{2(1-d_{\text{init}})} + \kappa(\kappa-1)(L^*)^{3m_{\text{bin}}+1} \cdot \kappa^{m_{\text{bin}}} + \sum_{\substack{i,j\in[\kappa] \\ i\neq j}} \frac{(d_{ij})^2}{2(1-d_{ij})}$$

$$\leq \kappa \cdot \frac{(d_{\text{init}})^3}{2(1-d_{\text{init}})} + (L^*)^{3m_{\text{bin}}+1} \cdot \kappa^{m_{\text{bin}}+2} + \kappa^2 \frac{(d_{\text{end}})^2}{2(1-d_{\text{end}})}.$$

For any $\kappa = O(\ln n)$, we obtain

$$\mathbb{E}\left[y \mid \bar{\tau}_1, \bar{\tau}_2\right] \mathbb{P}\left[\bar{\tau}_1, \bar{\tau}_2\right] \leq \mathbb{E}\left[y \mid \bar{\tau}_1\right] \mathbb{P}\left[\bar{\tau}_1\right]$$

$$\leq \mathbb{E}\left[L^* \frac{sxL^*}{\kappa n} \mid \bar{\tau}_1\right] \mathbb{P}\left[\bar{\tau}_1\right] = O\left(\frac{s(\ln n)^{4m_{\text{bin}}}}{n}\right). \qquad \square$$

REFERENCES

[1] H. ABBEY, *An examination of the Reed-Frost theory of epidemics*, Human Biol., 24 (1952), pp. 201–233.

[2] H. J. AHN AND B. HASSIBI, *Global dynamics of epidemic spread over complex networks*, in Proceedings of the 52nd IEEE Conference on Decision and Control (CDC), IEEE, 2013, pp. 4579–4585.

[3] R. ALBERT, H. JEONG, AND A.-L. BARABÁSI, *Error and attack tolerance of complex networks*, Nature, 406 (2000), pp. 378–382.

[4] N. ALON AND J. H. SPENCER, *The Probabilistic Method*, 4th ed., John Wiley & Sons, Hoboken, NJ, 2015.

[5] M. S. BARTLETT, *An inverse matrix adjustment arising in discriminant analysis*, Ann. Math. Statist., 22 (1951), pp. 107–111.

[6] P. BHARDWAJ, *Disrupting Diffusion: Critical Nodes in Network*, Master's thesis, Iowa State University, 2018.

[7] A. N. BISHOP AND I. SHAMES, *Link operations for slowing the spread of disease in complex networks*, Europhys. Lett., 95 (2011), 18005.

[8] A. BLUM, J. HOPCROFT, AND R. KANNAN, *Foundations of Data Science*, Cambridge University Press, Cambridge, 2020.

[9] C. BORGS, M. BRAUTBAR, J. T. CHAYES, AND B. LUCIER, *Maximizing social influence in nearly optimal time*, in Proceedings of the 25th Annual ACM-SIAM Symposium on Discrete Algorithms (SODA), 2014, pp. 946–957.

[10] C. BORGS, J. CHAYES, A. GANESH, AND A. SABERI, *How to distribute antidote to control epidemics*, Random Structures Algorithms, 37 (2010), pp. 204–222.

[11] T. N. BUI, *Graph Bisection Algorithms*, Ph.D. thesis, Massachusetts Institute of Technology, 1986.

[12] D. S. CALLAWAY, M. E. NEWMAN, S. H. STROGATZ, AND D. J. WATTS, *Network robustness and fragility: Percolation on random graphs*, Phys. Rev. Lett., 85 (2000), pp. 5468–5471.

[13] E. CATOR AND P. VAN MIEGHEM, *Nodal infection in Markovian susceptible-infected-susceptible and susceptible-infected-removed epidemics on networks are non-negatively correlated*, Phys. Rev. E, 89 (2014), 052802.

[14] M. DRAIEF, A. GANESH, L. MASSOULIÉ, ET AL., *Thresholds for virus spread on networks*, Ann. Appl. Probab., 18 (2008), pp. 359–378.

[15] E. A. ENNS, J. J. MOUNZER, AND M. L. BRANDEAU, *Optimal link removal for epidemic mitigation: A two-way partitioning approach*, Math. Biosci., 235 (2012), pp. 138–147.

[16] Executive Order N-33-20, Executive Department, State of California, https://covid19.ca.gov/img/Executive-Order-N-33-20.pdf (26 October 2020).

[17] J. A. FIRTH, J. HELLEWELL, P. KLEPAC, S. KISSLER, A. J. KUCHARSKI, AND L. G. SPURGIN, *Using a real-world network to model localized COVID-19 control strategies*, Nat. Med., 26 (2020), pp. 1616–1622.

[18] J. GOLDENBERG, B. LIBAI, AND E. MULLER, *Talk of the network: A complex systems look at the underlying process of word-of-mouth*, Marketing Lett., 12 (2001), pp. 211–223.

[19] S. GRACY, P. E. PARÉ, H. SANDBERG, AND K. H. JOHANSSON, *Analysis and distributed control of periodic epidemic processes*, IEEE Trans. Control Netw. Syst., 8 (2021), pp. 123–134.

[20] P. HOLME, B. J. KIM, C. N. YOON, AND S. K. HAN, *Attack vulnerability of complex networks*, Phys. Rev. E, 65 (2002), 056109.

[21] A. R. HOTA, J. GODBOLE, P. BHARIYA, AND P. E. PARÉ, *A Closed-Loop Framework for Inference, Prediction and Control of SIR Epidemics on Networks*, preprint, arXiv:2006.16185, 2020.

[22] D. KEMPE, J. M. KLEINBERG, AND É. TARDOS, *Maximizing the spread of influence through a social network*, in Proceedings of the 9th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining (KDD), 2003, pp. 137–146.

[23] W. O. KERMACK AND A. G. MCKENDRICK, *A contribution to the mathematical theory of epidemics*, Proc. Roy. Soc. London Ser. A, 115 (1927), pp. 700–721.

[24] A. KHANAFER AND T. BASAR, *An optimal control problem over infected networks*, in Proceedings of the International Conference of Control, Dynamic Systems, and Robotics, 2014, pp. 1–6.

[25] P. KLEPAC, S. KISSLER, AND J. GOG, *Contagion! The BBC four pandemic–The model behind the documentary*, Epidemics, 24 (2018), pp. 49–59.

[26] C. R. MACCLUER, *The many proofs and applications of Perron's theorem*, SIAM Rev., 42 (2000), pp. 487–498.

[27] V. S. Mai and E. H. Abed, *Optimizing leader influence in networks through selection of direct followers*, IEEE Trans. Automat. Control, 64 (2019), pp. 1280–1287.

[28] V. S. Mai, A. Battou, and K. Mills, *Distributed algorithm for suppressing epidemic spread in networks*, IEEE Control Syst. Lett., 2 (2018), pp. 555–560.

[29] Ordinance on a Prohibition Against Holding Public Gatherings and Events, Ministry of Justice, Sweden, https://www.government.se/articles/2020/03/ordinance-on-a-prohibition-against-holding-public-gatherings-and-events/ (26 October 2020).

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

[31] P. Netrapalli and S. Sanghavi, *Learning the graph of epidemic cascades*, ACM SIGMETRICS Perform. Eval. Rev., 40 (2012), pp. 211–222.

[32] C. Nowzari, V. M. Preciado, and G. J. Pappas, *Analysis and control of epidemics: A survey of spreading processes on complex networks*, IEEE Control Syst., 36 (2016), pp. 26–46.

[33] M. Ogura and V. M. Preciado, *Efficient containment of exact SIR Markovian processes on networks*, in Proceedings of the 55th IEEE Conference on Decision and Control (CDC), 2016, pp. 967–972.

[34] P. E. Paré, C. L. Beck, and A. Nedić, *Epidemic processes over time-varying networks*, IEEE Trans. Control Netw. Syst., 5 (2018), pp. 1322–1334.

[35] P. E. Paré, J. Liu, C. Beck, A. Nedić, and T. Başar, *Multi-competitive viruses over time–varying networks with mutations and human awareness*, Automatica, 123 (2021), 109330.

[36] R. J. Plemmons, *M-matrix characterizations. I—Nonsingular m-matrices*, Linear Algebra Appl., 18 (1977), pp. 175–188.

[37] B. A. Prakash, L. Adamic, T. Iwashyna, H. Tong, and C. Faloutsos, *Fractional immunization in networks*, in Proceedings of the SIAM International Conference on Data Mining (SDM), SIAM, Philadelphia, 2013, pp. 659–667.

[38] V. M. Preciado, M. Zargham, C. Enyioha, A. Jadbabaie, and G. J. Pappas, *Optimal vaccine allocation to control epidemic outbreaks in arbitrary networks*, in Proceedings of the 52nd IEEE Conference on Decision and Control (CDC), IEEE, 2013, pp. 7486–7491.

[39] V. M. Preciado, M. Zargham, C. Enyioha, A. Jadbabaie, and G. J. Pappas, *Optimal resource allocation for network protection against spreading processes*, IEEE Trans. Control Netw. Syst., 1 (2014), pp. 99–108.

[40] N. A. Ruhi and B. Hassibi, *SIRS epidemics on complex networks: Concurrence of exact Markov chain and approximated models*, in Proceedings of the 54th IEEE Conference on Decision and Control (CDC), IEEE, 2015, pp. 2919–2926.

[41] C. M. Schneider, T. Mihaljev, S. Havlin, and H. J. Herrmann, *Suppressing epidemics with a limited amount of immunization units*, Physical Rev. E, 84 (2011), 061911.

[42] P. Van Mieghem, *Exact Markovian SIR and SIS Epidemics on Networks and an Upper Bound for the Epidemic Threshold*, preprint, arXiv:1402.1731, 2014.

[43] P. Van Mieghem, J. Omic, and R. Kooij, *Virus spread in networks*, IEEE/ACM Trans. Netw., 17 (2009), pp. 1–14.

[44] P. Van Mieghem, D. Stevanović, F. Kuipers, C. Li, R. Van De Bovenkamp, D. Liu, and H. Wang, *Decreasing the spectral radius of a graph by link removals*, Physical Rev. E, 84 (2011), 016101.

[45] Y. Wan, S. Roy, and A. Saberi, *Network design problems for controlling virus spread*, in Proceedings of the 46th IEEE Conference on Decision and Control (CDC), IEEE, 2007, pp. 3925–3932.

[46] Y. Wan, S. Roy, and A. Saberi, *Designing spatially heterogeneous strategies for control of virus spread*, IET Syst. Biol., 2 (2008), pp. 184–201.

[47] X. Wu, A. Kumar, D. Sheldon, and S. Zilberstein, *Robust optimization for tree-structured stochastic network design*, in Proceedings of the 31st AAAI Conference on Artificial Intelligence, 2017, pp. 4545–4551.

[48] M. Youssef and C. Scoglio, *An individual-based approach to SIR epidemics in contact networks*, J. Theoret. Biol., 283 (2011), pp. 136–144.