Optimization Algorithms for Epidemic Evolution in Broadcast Networks

Xiangping Zhai*, Liang Zheng†, Jianping Wang† and Chee Wei Tan†
*† City University of Hong Kong
* xpzhai2@student.cityu.edu.hk, †{lianzheng2,jianwang,cheewtan}@cityu.edu.hk

***Abstract*—Epidemic evolution is the spread of a computer or biological virus over a network. The goal is to control the speed of the epidemic evolution with limited network control resources and to study how users in the network can be infected. The epidemic evolution can be modeled by a probabilistic dynamical system over a connected graph. We consider several epidemic evolution models in the literature, and formulate their evolution control under a common framework that requires solving a nonconvex optimization problem with an objective that is the spectral radius function of a nonnegative matrix. We propose two algorithms to tackle this optimization problem. The first one is a suboptimal but computationally fast algorithm based on successive convex relaxation, while the second one can compute a global optimal solution using branch-and-bound techniques that leverage some key inequalities in nonnegative matrix theory.**

***Index Terms*—Epidemic evolution, spectral radius minimization, and nonnegative matrix theory.**

I. Introduction

Epidemic evolution is the spread of a computer or biological virus over a network. How to estimate and control the spread of an epidemic is of practical interests in cyber-physical networks or medical epidemiology. An epidemic evolution process can be modeled as a probabilistic dynamical systems over a connected graph, and controlling this process can then be suitably analyzed using systems control theory [1], [2]. In general, the stability of the evolution process is governed by the network connectivity and the parameters associated with the epidemic, e.g., propagation models, infection rates. Therefore, how do epidemics spread in bio-systems or cyber-physical networks and, how they can be controlled is a difficult stochastic nonlinear control problem. Network resources can be limited due to practical constraints such as deploying anti-virus measures or due to finite capacity. The static optimization solution is then applied as parameters to the probabilistic dynamical system.

In the literature, a typical propagation model classifies nodes in the network into three basic types: (i) *susceptible:* ones who are healthy and prone to getting infected, (ii) *infected:* ones who can infect others, and (iii) *recovered:* ones who were infected before but cured for now. Different epidemic models can be considered based on different combinations of these three basic types. When the model takes into account the network connectivity (graph topology), an epidemic threshold is used to evaluate the evolution of the epidemic [3]. The authors in [4] look at different threshold criteria and their relationship with graph structures for epidemic spreading. The authors in [5] introduce the epidemic threshold for a network that is closely related to the spectral radius of the topological matrix. The authors in [6] seek to minimize a general aggregate cost and characterize the optimal dynamic patching policies, relying on the homogeneous assumption that all pairs of nodes have identical contact rate. With an increasing popularity in mobile computing, particularly with the rise in Bluetooth-equipped mobile devices, the authors in [7] analyze the epidemic threshold on mobile ad-hoc networks. The authors in [8] propose an inference technique to find the source of a computer virus in a network. The authors in [9] study the fundamental spreading patterns that characterize a mobile phone virus outbreak. In these related work, the network topology is a fundamental property that is exploited to control epidemic spreading.

In [10], the authors formally introduce canonical epidemic control optimization problems that are deterministic. The authors propose minimizing the spectral radius (which captures the epidemic spreading rate) of the network topology matrix with a given amount of network control resources. However, the applicability is limited, for example, it is assumed that the abstracted interactive graph of the epidemic evolution has to be symmetric or diagonally symmetrized. These assumptions in [10] can be relaxed to broaden its applicability. It may be more useful to generalize the epidemic topological structure to include directed asymmetric interactive graphs since different nodes in the network may receive different levels of immunization.

Our approach in this paper is to use deterministic finite optimization theory to study the problem: given a set of limited network resources to control the epidemic evolution, how to optimize *the rate of epidemic evolution*? Motivated by [10], we study a deterministic epidemic evolution control framework that requires solving a nonconvex optimization problem by minimizing a spectral radius function subject to constraints on the control variables. We first show that the optimal solution of the spectral radius minimization problem resides on the boundary of the constraint set. Then, feasible suboptimal solutions of the spectral radius minimization problem are obtained using successive convex approximation based on geometric programming. Key tools in nonnegative matrix theory, e.g., the Perron-Frobenius theorem and the Friedland-Karlin inequalities, are then leveraged to find bounds that can be tightened using the branch-and-bound method. This leads to an algorithm that can compute the global optimal solution

978-1-4673-5939-9/13/$31.00 ©2013 IEEE

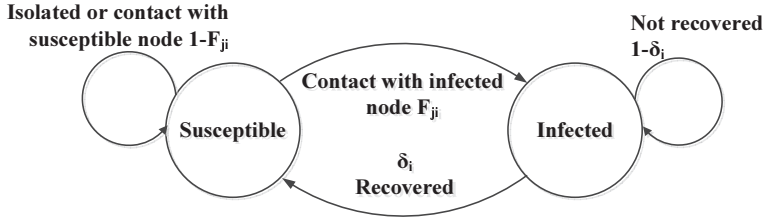

Fig. 1. Overview of the susceptible-infected model for the ith node. If the ith node is susceptible, it can be infected by receiving the virus from an infected node j with probability F_{ji}. On the other hand, if the ith node is infected, it can be recovered with probability δ_i.

of the spectral radius minimization problem.

The rest of this paper is organized as follows. We introduce the system model and formulate our epidemic evolution problem in Section II. We propose a fast successive convex approximation algorithm to solve the problem suboptimally in Section III. In Section IV, we propose a branch-and-bound algorithm that leverages some key inequalities in nonnegative matrix theory to compute the global optimal solution. In Section V, we evaluate the performance of our algorithms numerically. The conclusion of this paper is in Section VI. The following notation is used in our paper. Column vectors and matrices are denoted by boldfaced lowercase and uppercase respectively. $\rho(\mathbf{A})$ denotes the Perron-Frobenius eigenvalue of a nonnegative matrix $\mathbf{A}$. $\mathbf{x}(\mathbf{A})$ and $\mathbf{y}(\mathbf{A})$ denote the Perron right and left eigenvectors of $\mathbf{A}$ associated with $\rho(\mathbf{A})$, respectively. $\mathbf{I}$ denotes the identity matrix. For a given vector $\mathbf{x} = (x_1, \ldots, x_L)^\top$, $\mathrm{diag}(\mathbf{x})$ is a diagonal matrix $\mathrm{diag}(x_1, \ldots, x_L)$.

II. System Model and Problem Formulation

We consider a broadcast network with L users where a virus spreads from infected users to susceptible users. In this section, we review two discrete-time-based linear models, namely the multi-group model and the contact network model [10]. Let $\mathbf{F}$ represent the transmission coefficient matrix, where F_{ij} is the infectious ratio that the ith group infects the jth group in the multi-group model, or the ith node infects the jth node in the contact network model. $\boldsymbol{\delta} = (\delta_1, \ldots, \delta_L)^\top$ is the recovery vector where an infected node i recovers with probability δ_i. Figure 1 shows the susceptible-infected model on the ith node. We then construct a next-generation matrix, denoted by $\mathbf{G}$, that represents the transformation of nodes' types at each discrete time slot. By regarding different parameters in the next-generation matrix as control variables, we can control the rate of the epidemic evolution. Thus, in the following, we formulate the optimization problem that minimizes the spectral radius of the next-generation matrix.

A. Multi-group Model

Let us introduce the multi-group model with N groups, and the number of nodes in each group is n_i (which implies $\sum_{i=1}^{N} n_i = L$). In this case, an infected node does not recover. Assuming the average infectious time of a node in the ith group is τ_i, the infectious nodes $\eta_i(t+1)$ of the ith group at $(t+1)$th generation is given in terms of $\eta_i(t)$ as:

$$\frac{\eta_i(t+1)}{n_i} = \tau_i \left(F_{ii} \frac{\eta_i(t)}{n_i} + \sum_{j \neq i} F_{ij} \frac{\eta_j(t)}{n_j} \right). \tag{1}$$

Note that a node can be infected either internally by the infected nodes in the same group or externally by the infected nodes in the other groups. To control the epidemic spreading rate, we introduce the following variables:

1) $\hat{F}_{ij} = \frac{1}{\xi_i} F_{ij},\ \xi_i \in (0,1],\ i,j = 1, \ldots, N$. $\boldsymbol{\xi}$ scales the transmission coefficient matrix, and decreases the epidemic evolution rate in the whole network. e.g., the infected nodes are isolated.
2) $\bar{F}_{ij} = \frac{1}{\varphi_i} \hat{F}_{ij},\ \varphi_i \in (0,1],\ i \neq j,\ i,j = 1, \ldots, N$. Note that $\bar{\mathbf{F}}$ is a nonnegative matrix with $\bar{F}_{ii} = 0$ for all i. Thus, the contact rate between different groups is reduced, which means the restriction of nodes' movement within its own group.
3) $\hat{\tau} = \frac{1}{\omega_i} \tau_i,\ \omega_i \in (0,1],\ i = 1, \ldots, N$. We control the infectious time of nodes by introducing the variable ω_i in each group. In practice, when anti-virus software is used to overcome infection, ω_i can be set to be infinity.

The next-generation matrix for epidemic evolution is given by:

$$\mathbf{G} = \hat{\tau}\, \mathrm{diag}(\boldsymbol{\omega})\, \mathrm{diag}(\boldsymbol{\xi}) \big(\hat{\mathbf{F}}^{\mathbf{d}} + \mathrm{diag}(\boldsymbol{\varphi}) \bar{\mathbf{F}} \big), \tag{2}$$

where $\hat{\mathbf{F}}^{\mathbf{d}}$ is a diagonal matrix with $\hat{F}^{\mathbf{d}}_{ii} = \hat{F}_{ii}$ for all i.

Next, we introduce the contact network model in Section II-B. Similar to this section, we first construct the next-generation matrix, and then introduce the control variables to analyze the epidemic evolution.

B. Contact Network Model

In the contact network model, an epidemic spreads by nodes' communication, and nodes' movement does not depend on any group pattern. The ith node is infected at the $(t+1)$th time slot with probability $p_i(t+1)$, which is given by in terms of $\mathbf{p}(t)$ as:

$$p_i(t+1) = \left(1 - \prod_{j=1}^{L} \big(1 - F_{ij} p_j(t) \big) \right) + (1 - \delta_i) p_i(t), \tag{3}$$

TABLE I
SUMMARY OF THE NEXT-GENERATION MATRIX FOR EPIDEMIC EVOLUTION.

$\mathbf{G}(\mathbf{s})$	$\mathbf{s}$	Problem Parameters
$\mathrm{diag}(\mathbf{s})\mathbf{B}$	$\boldsymbol{\xi}$	$\mathbf{B}=\hat{\tau}\,\mathrm{diag}(\boldsymbol{\omega})\big(\hat{\mathbf{F}}^{\mathbf{d}}+\mathrm{diag}(\boldsymbol{\varphi})\bar{\mathbf{F}}\big)$
	$\boldsymbol{\omega}$	$\mathbf{B}=\hat{\tau}\,\mathrm{diag}(\boldsymbol{\xi})\big(\hat{\mathbf{F}}^{\mathbf{d}}+\mathrm{diag}(\boldsymbol{\varphi})\bar{\mathbf{F}}\big)$
$\mathrm{diag}(\mathbf{d})+\mathrm{diag}(\mathbf{s})\mathbf{B}$	$\boldsymbol{\varphi}$	$\mathrm{diag}(\mathbf{d})=\hat{\tau}\,\mathrm{diag}(\boldsymbol{\omega})\,\mathrm{diag}(\boldsymbol{\xi})\hat{\mathbf{F}}^{\mathbf{d}},\quad \mathbf{B}=\hat{\tau}\,\mathrm{diag}(\boldsymbol{\omega})\,\mathrm{diag}(\boldsymbol{\xi})\bar{\mathbf{F}}$
	$\boldsymbol{\beta}$	$\mathbf{d}=\mathbf{1}-\boldsymbol{\delta},\quad \mathbf{B}=\mathbf{F}$
$\mathrm{diag}(\mathbf{s})+\mathbf{B}$	$\mathbf{1}-\boldsymbol{\delta}$	$\mathbf{B}=\mathrm{diag}(\boldsymbol{\beta})\tilde{\mathbf{F}}$

where δ_i is the recovery rate. Note that $\prod_{j=1}^{L}(1-F_{ij}p_j(t))$ is the probability that the ith node does not communicate to any infected node, and its complement is the probability that the ith node is infected by at least one infected node. Note that $(1-\delta_i)p_i(t)$ is the probability that the infected node i does not recover at the $(t+1)$th time slot.

Controlling the immunity (denoted as $F_{ij}=\beta_i\tilde{F}_{ij},\ \beta_i\in[0,1],\ i\neq j,\ i,j=1,\ldots,L$) of each node, (3) can be reformulated as:

$$p_i(t+1)=\sum_{j=1}^{L}\beta_i\tilde{F}_{ij}p_j(t)+(1-\delta_i)p_i(t),$$

since $1-\prod_{j=1}^{L}(1-F_{ij}p_j(t))$ can be well approximated by $\sum_{j=1}^{L}F_{ij}p_j(t)$. Thus, the next-generation matrix is given by:

$$\mathbf{G}=\mathrm{diag}(\mathbf{1}-\boldsymbol{\delta})+\mathrm{diag}(\boldsymbol{\beta})\tilde{\mathbf{F}}. \tag{4}$$

Therefore, the epidemic evolution is dependent on the recovery rate $\boldsymbol{\delta}$ and the immunity $\boldsymbol{\beta}$ that can be optimized appropriately.

C. Spectral Radius Minimization Problem

The spectral radius of the next-generation matrix, denoted by $\rho(\mathbf{G})$, is a key metric that captures the spreading rate of a virus. We consider the next-generation matrix in terms of a control variable $\mathbf{s}$ which is summarized in Table I, i.e., $\mathbf{s}$ can be $\boldsymbol{\xi}$, $\boldsymbol{\varphi}$, or $\boldsymbol{\omega}$ in (2), and $\boldsymbol{\delta}$ or $\boldsymbol{\beta}$ in (4).

In addition, we assume that $\mathbf{s}$ is constrained by an upper bound $\mathbf{u}$ and a lower bound $\mathbf{b}$, and $\sum_{i=1}^{L}s_i$ cannot be less than Γ. For feasibility, these constraints on $\mathbf{s}$ should implicitly satisfy $\Gamma\leq\sum_{i=1}^{L}u_i$. Then, we formulate the minimizing speed of the epidemic evolution as the problem:

$$\begin{aligned}\text{minimize}\quad & \rho(\mathbf{G}(\mathbf{s}))\\ \text{subject to}\quad & \sum_{l=1}^{L}s_l\geq\Gamma,\\ & \mathbf{b}\leq\mathbf{s}\leq\mathbf{u},\\ \text{variables:}\quad & \mathbf{s}.\end{aligned} \tag{5}$$

In general, (5) is nonconvex (due to the spectral radius objective function), and thus it is difficult to solve.

Lemma 1: Assuming that $\mathbf{G}(\mathbf{s})$ is irreducible, the constraint $\mathbf{1}^\top\mathbf{s}\geq\Gamma$ in (5) is tight at optimality, i.e., $\mathbf{1}^\top\mathbf{s}^\star=\Gamma$.

Proof: From Perron-Frobenius theorem [11], $\mathbf{G}(\mathbf{s})$ has a real spectral radius, and the Perron right and left eigenvectors associated with the spectral radius are real and strictly positive. Let the associated right and left eigenvectors be $\mathbf{x}$ and $\mathbf{y}$ with $\mathbf{x}>\mathbf{0}$, $\mathbf{y}>\mathbf{0}$, respectively. From the eigenvalue sensitivity theory [12], [13], we have:

$$\partial\rho(\mathbf{G}(\mathbf{s}))/\partial s_i=\mathbf{y}^\top(\partial\mathbf{G}(\mathbf{s})/\partial s_i)^\top\mathbf{x}.$$

We first prove that $\mathbf{G}(\mathbf{s})$ is nondecreasing by inspecting the first order derivative of the special cases of $\mathbf{G}(\mathbf{s})$ in Table I which can be given respectively as follows:

1) For $\mathbf{G}(\mathbf{s})=\mathrm{diag}(\mathbf{s})\mathbf{B}$, we have:
$$\partial\rho(\mathbf{G}(\mathbf{s}))/\partial s_i=\rho(\mathbf{G}(\mathbf{s}))x_iy_i\geq 0.$$
2) For $\mathbf{G}(\mathbf{s})=\mathrm{diag}(\mathbf{d})+\mathrm{diag}(\mathbf{s})\mathbf{B}$, we have:
$$\partial\rho(\mathbf{G}(\mathbf{s}))/\partial s_i=x_i\sum_{j=1}^{L}B_{ji}y_j\geq 0.$$
3) For $\mathbf{G}(\mathbf{s})=\mathrm{diag}(\mathbf{s})+\mathbf{B}$, we have $\partial\rho(\mathbf{G}(\mathbf{s}))/\partial s_i\geq 0$ according to Theorem A.2 in [10].

Therefore, $\rho(\mathbf{G}(\mathbf{s}))$ is nondecreasing. Since minimizing $\rho(\mathbf{G}(\mathbf{s}))$ chooses a feasible $\mathbf{s}$ as small as possible, this makes $\mathbf{1}^\top\mathbf{s}\geq\Gamma$ tight at optimality. ■

In the following, we first solve (5) by successive convex approximation in Section III. Then, we propose a branch-and-bound method [14] to calculate the global optimal solution for (5) in Section IV.

III. Successive Convex Approximation

In this section, we propose to approximate (5) using a successive convex approximation technique based on geometric programming (similarly to the condensation method in signomial programming).

Algorithm 1 (Successive Convex Approximation):

Choose a feasible initial point $\mathbf{s}(0)$.

1) Set $\alpha_l(k)=\frac{s_l(k)}{\sum_{j=1}^{L}s_j(k)},\quad l=1,\ldots,L.$
2) Solve the kth approximation problem by geometric programming:

$$\begin{aligned}\text{minimize}\quad & \lambda\\ \text{subject to}\quad & \mathbf{G}(\mathbf{s})\mathbf{v}\leq\lambda\mathbf{v},\\ & \prod_{l=1}^{L}\left(\frac{s_l}{\alpha_l(k)}\right)^{\alpha_l(k)}=\Gamma,\\ & \mathbf{b}\leq\mathbf{s}\leq\mathbf{u},\\ \text{variables:}\quad & \mathbf{s},\mathbf{v},\lambda,\end{aligned} \tag{6}$$

whose optimal solution is denoted as $\mathbf{s}(k+1)$.

Theorem 1: Starting from any feasible initial point $\mathbf{s}(0)$, $\mathbf{s}(k)$ converges to the local optimal solution of (5).

Proof: Using Lemma 1 and a basic result in the Perron-Frobenius eigenvalue (cf. Chapter 4.5 in [15]):

$$\rho(\mathbf{G}) = \inf\{\lambda | \mathbf{G}\mathbf{v} \leq \lambda \mathbf{v} \text{ for some } \mathbf{v} > 0\},$$

we reformulate (5) to the following problem by introducing the auxiliary variables λ and $\mathbf{v}$:

$$\begin{array}{ll} \text{minimize} & \lambda \\ \text{subject to} & \mathbf{G}(\mathbf{s})\mathbf{v} \leq \lambda \mathbf{v}, \\ & \sum_{l=1}^{L} s_l = \Gamma, \\ & \mathbf{b} \leq \mathbf{s} \leq \mathbf{u}, \\ \text{variables:} & \mathbf{s}, \mathbf{v}, \lambda. \end{array} \tag{7}$$

However, (7) still cannot be turned into a geometric program due to the posynomial equality $\sum_{l=1}^{L} s_l = \Gamma$ in the constraint set of (7). Let $g(\mathbf{s}) = \sum_{l=1}^{L} s_l$, then we approximate $g(\mathbf{s})$ with a monomial $\tilde{g}(\mathbf{s})$, such that:

$$g(\mathbf{s}) \geq \tilde{g}(\mathbf{s}) = \prod_{l=1}^{L} \left(\frac{s_l}{\alpha_l}\right)^{\alpha_l}. \tag{8}$$

We set $\alpha_l(k) = s_l(k-1)/\sum_i s_i(k-1)$, then $\tilde{g}(\mathbf{s})$ is the best local monomial approximation to $g(\mathbf{s})$ near $\mathbf{s}$ in the sense of first order Taylor approximation. This leads to the use of successive convex approximation (6).

This approximation satisfies the following properties:

1) For any fixed positive $\mathbf{s}(k)$, $g(\mathbf{s}(k)) = \tilde{g}(\mathbf{s}(k))$ guarantees that any solution of the approximated problem (6) is a feasible point of the original problem (7).

2) $g(\mathbf{s}(k)) = \tilde{g}(\mathbf{s}(k))$ also guarantees that the solution of each approximated problem decreases the object function.

3) $\nabla g(\mathbf{s}(k)) = \nabla \tilde{g}(\mathbf{s}(k))$ guarantees that the Karush-Kuhn-Tucker conditions of the original problem are satisfied after the series of approximations converges.

Therefore, the solutions of these sequences of convex approximations given in (6) converge to a stationary point of (5). ■

Remark 1: The optimization problem at Step 2 is a geometric program, which can be solved by an interior point method [15]. The convergence of Algorithm 1 depends on the initial point $\mathbf{s}(0)$. The output of Algorithm 1 is an upper bound to (5), which facilitates the branch-and-bound method proposed in the next section.

IV. Global Optimization Algorithm

Although the suboptimal solution computed in Algorithm 1 by a successive convex approximation technique only provides an upper bound for (5), the global optimal solution can be obtained by a branch-and-bound method that uses Algorithm 1 as a submodule. We consider an initial box $\{\mathbf{s} \,|\, \mathbf{b} \leq \mathbf{s} \leq \mathbf{u}\}$ which is subdivided iteratively according to the upper bound and the lower bound of (5). In the following, we discuss the lower bounds of the cases summarized in Table I.

1) For $\mathbf{G}(\mathbf{s}) = \text{diag}(\mathbf{s})\mathbf{B}$, using the Friedland-Karlin inequality in [16], we have:

$$\rho(\text{diag}(\mathbf{s})\mathbf{B}) \geq \rho(\mathbf{B}) \prod_{l=1}^{L} s_l^{x_l(\mathbf{B})y_l(\mathbf{B})},$$

where $\mathbf{x}(\mathbf{B})$ and $\mathbf{y}(\mathbf{B})$ are the Perron right and left eigenvectors of $\mathbf{B}$, normalized such that $\sum_{l=1}^{L} x_l(\mathbf{B})y_l(\mathbf{B}) = 1$. Hence, the optimal value of (5) is lower bounded by the optimal value of the following problem:

$$\begin{array}{ll} \text{minimize} & \sum_{l=1}^{L} x_l(\mathbf{B})y_l(\mathbf{B}) \log s_l \\ \text{subject to} & \sum_{l=1}^{L} s_l = \Gamma, \\ & \mathbf{b} \leq \mathbf{s} \leq \mathbf{u}, \\ \text{variables:} & \mathbf{s}. \end{array} \tag{9}$$

Although (9) is still a nonconvex problem, we replace the objective function of (9) by its convex envelope [17] over the box constraint set $\{\mathbf{s} \,|\, \mathbf{b} \leq \mathbf{s} \leq \mathbf{u}\}$ to obtain the linear program (12).

2) For $\mathbf{G}(\mathbf{s}) = \text{diag}(\mathbf{d}) + \text{diag}(\mathbf{s})\mathbf{B}$, we have:

$$\begin{aligned} \rho(\text{diag}(\mathbf{d}) + \text{diag}(\mathbf{s})\mathbf{B}) &\geq \rho\Big(\big(\min_{l=1,\ldots,L} d_l\big)\mathbf{I} + \text{diag}(\mathbf{s})\mathbf{B}\Big) \\ &= \rho(\text{diag}(\mathbf{s})\mathbf{B}) + \min_{l=1,\ldots,L} d_l. \end{aligned}$$

Hence, the optimal value of (5) is bounded by the optimal value of the following problem:

$$\begin{array}{ll} \text{minimize} & \min_{l=1,\ldots,L} d_l + \rho(\text{diag}(\mathbf{s})\mathbf{B}) \\ \text{subject to} & \sum_{l=1}^{L} s_l = \Gamma, \\ & \mathbf{b} \leq \mathbf{s} \leq \mathbf{u}, \\ \text{variables:} & \mathbf{s}. \end{array} \tag{10}$$

Thus, similar to the case of $\mathbf{G}(\mathbf{s}) = \text{diag}(\mathbf{s})\mathbf{B}$, the convex envelope of (10) is also obtained by the linear program in (12).

3) For $\mathbf{G}(\mathbf{s}) = \text{diag}(\mathbf{s}) + \mathbf{B}$, we first state the inequality in nonnegative matrix theory [18]:

$$\frac{\rho(\mathbf{D})}{\rho(\mathbf{A})} \geq \prod_{l=1}^{L}\prod_{j=1}^{L} \left(\frac{D_{lj}}{A_{lj}}\right)^{\frac{A_{lj}x_l(\mathbf{A})y_j(\mathbf{A})}{\rho(\mathbf{A})}},$$

where $\mathbf{A}$ and $\mathbf{D}$ are irreducible nonnegative matrices. Letting $\mathbf{A} = \mathbf{B}$ and $\mathbf{D} = \text{diag}(\mathbf{s}) + \mathbf{B}$ respectively, we have:

$$\rho(\text{diag}(\mathbf{s}) + \mathbf{B}) \geq \rho(\mathbf{B}) \prod_{l=1}^{L} \left(\frac{s_l + B_{ll}}{B_{ll}}\right)^{\frac{B_{ll}x_l(\mathbf{B})y_l(\mathbf{B})}{\rho(\mathbf{B})}}.$$

Likewise, the optimal value of (5) is bounded by the optimal value of the following problem:

$$\begin{aligned} \text{minimize} \quad & \sum_{l=1}^{L} B_{ll} x_l(\mathbf{B}) y_l(\mathbf{B}) \log(s_l + B_{ll}) \\ \text{subject to} \quad & \sum_{l=1}^{L} s_l = \Gamma, \\ & \mathbf{b} \le \mathbf{s} \le \mathbf{u}, \\ \text{variables:} \quad & \mathbf{s}. \end{aligned} \tag{11}$$

Following the same step for the case $\mathbf{G}(\mathbf{s}) = \mathrm{diag}(\mathbf{s})\mathbf{B}$, we also replace the objective function of (11) by its convex envelope over the box constraint set $\{\mathbf{s} \,|\, \mathbf{b} \le \mathbf{s} \le \mathbf{u}\}$ to obtain the linear program in (13).

By leveraging the above inequalities on the spectral radius and exploiting the branch-and-bound method [14], we propose the following algorithm to solve (5) optimally.

Algorithm 2 (Global Optimization Algorithm):

1) **Initialization**
Set $k = 0$ and $\mathcal{Q}_0 = \{\Omega_0\}$, where Ω_0 is the initial rectangular set $[\mathbf{b}, \mathbf{u}]$. Obtain the lower bound $L_0 = V_l(\Omega_0)$ for (5) by solving the following linear program.
If $\mathbf{G}(\mathbf{s}) = \mathrm{diag}(\mathbf{s})\mathbf{B}$ **or** $\mathbf{G}(\mathbf{s}) = \mathrm{diag}(\mathbf{d}) + \mathrm{diag}(\mathbf{s})\mathbf{B}$:

$$\begin{aligned} \text{minimize} \quad & \sum_{l=1}^{L} \frac{\log u_l - \log b_l}{u_l - b_l} x_l(\mathbf{B}) y_l(\mathbf{B}) s_l \\ \text{subject to} \quad & \sum_{l=1}^{L} s_l = \Gamma, \\ & \mathbf{b} \le \mathbf{s} \le \mathbf{u}, \\ \text{variables:} \quad & \mathbf{s}. \end{aligned} \tag{12}$$

else if $\mathbf{G}(\mathbf{s}) = \mathrm{diag}(\mathbf{s}) + \mathbf{B}$:

$$\begin{aligned} \text{minimize} \quad & \sum_{l=1}^{L} \frac{\log \frac{u_l + B_{ll}}{b_l + B_{ll}}}{u_l - b_l} B_{ll} x_l(\mathbf{B}) y_l(\mathbf{B}) s_l \\ \text{subject to} \quad & \sum_{l=1}^{L} s_l = \Gamma, \\ & \mathbf{b} \le \mathbf{s} \le \mathbf{u}, \\ \text{variables:} \quad & \mathbf{s}. \end{aligned} \tag{13}$$

end
Obtain the upper bound $U_0 = V_u(\Omega_0)$ by solving (5) with Algorithm 1.

2) **Stop Criterion**
If $U_k - L_k < \varepsilon$, then stop;
else go to the next step.

3) **Branching**
Pick a rectangular set $\bar{\Omega} \in \mathcal{Q}_k$ that satisfies $V_l(\bar{\Omega}) = L_k$. Split $\bar{\Omega}$ into Ω_{I} and Ω_{II} along one of its longest edges: $\mathcal{Q}_{k+1} = (\mathcal{Q}_k - \{\bar{\Omega}\}) \cup \{\Omega_{\mathrm{I}}, \Omega_{\mathrm{II}}\}$.
Update $L_{k+1} = \min_{\Omega \in \mathcal{Q}_{k+1}} V_l(\Omega)$ and $U_{k+1} = \min_{\Omega \in \mathcal{Q}_{k+1}} V_u(\Omega)$.

4) **Pruning**
Remove all Ω from $\mathcal{Q}_{k+1}$ where $V_l(\Omega) > U_{k+1}$.
Set $k \leftarrow k + 1$. Go to Step 2.

Theorem 2: Starting from an initial rectangular Ω_0, Algorithm 2 converges to the global optimal solution of (5).

Proof: We get the lower bound of (5) by solving the relaxed linear program and get the upper bound by Algorithm 1 for all the three types of $\mathbf{G}(\mathbf{s})$. Then, based on the upper bound and lower bound, Algorithm 2 terminates with a certificate proving that the suboptimal point found is ε-suboptimal [14]. ■

Remark 2: Algorithm 1 runs as an inner loop of Algorithm 2 at Step 2 to provide the upper bound for (5). L_k includes the child nodes across all the leaves in a binary tree. At the last two steps of Algorithm 2, namely "branching" and "pruning", minimizing $\rho(\mathbf{G}(\mathbf{s}))$ over all the lower bound is searching for a global lower bound on the optimal value of (5). When $b_l = 0$ for some l, we let $b_l = \epsilon$ to make Algorithm 1 run smoothly, where ϵ is a small enough positive value to approximate $\epsilon \to 0$. A practical stopping criterion for Algorithm 2 is $U_k - L_k < \varepsilon$ for a given small ε.

V. Numerical experiments

In this section, we provide numerical experiments to illustrate the performance of Algorithm 1 in Section III and Algorithm 2 in Section IV.

Example 1: We first consider the case that $\mathbf{G}(\mathbf{s}) = \mathrm{diag}(\mathbf{s})\mathbf{B}$, where $\mathbf{B}$ is a symmetric matrix given by:

$$\mathbf{B} = \begin{bmatrix} 0.620 & 0.310 & 0.124 \\ 0.310 & 0.620 & 0.310 \\ 0.124 & 0.310 & 0.620 \end{bmatrix}.$$

We set $\Gamma = 2.3$ and $[\mathbf{b}, \mathbf{u}] = [\mathbf{0}, \mathbf{1}]$, where $\mathbf{0}$ and $\mathbf{1}$ are the all zero and all one vectors, respectively. The first three iterations for the binary tree of Algorithm 2 are shown in Figure 2. Figure 3 (a) and (c) plot the evolutions of spectral radius that run Algorithm 1 and Algorithm 2 respectively. The optimal spectral radius is $\rho^\star = 0.8318$ and the optimal solution is $\mathbf{s}^\star = (0.9622, 0.3756, 0.9622)^\top$.

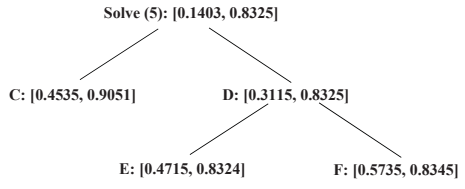

Fig. 2. We use a rectangular set $[\epsilon, 1]^3$ with $\epsilon = 0.001$. The lower bound and upper bound are obtained corresponding to Step 1 of Algorithm 2. At the root of the tree, we have $L_0 = 0.1403$ and $U_0 = 0.8325$. The rectangular set is then partitioned into two sets: C and D (C is the set $s_1 \in [\epsilon, (\epsilon+1)/2], s_2 \in [\epsilon, 1], s_3 \in [\epsilon, 1]$ and D is the set $s_1 \in [(\epsilon+1)/2, 1], s_2 \in [\epsilon, 1], s_3 \in [\epsilon, 1]$). We then have $L_1 = 0.3115$ and $U_1 = 0.8325$. In the third level of the binary tree, we partition the set D to obtain the bottom children E and F (E is the set $s_1 \in [(\epsilon+1)/2, 1], s_2 \in [\epsilon, (\epsilon+1)/2], s_3 \in [\epsilon, 1]$ and F is the set $s_1 \in [(\epsilon+1)/2, 1], s_2 \in [(\epsilon+1)/2, 1], s_3 \in [\epsilon, 1]$). We then have $L_2 = 0.4535$ and $U_2 = 0.8324$.

Next, we consider the case that $\mathbf{G}(\mathbf{s}) = \mathrm{diag}(\mathbf{s}) + \mathbf{B}$, where $\mathbf{B}$ is a nonnegative matrix given by:

$$\mathbf{B} = \begin{bmatrix} 0.000 & 0.124 & 0.124 \\ 0.310 & 0.000 & 0.310 \\ 0.124 & 0.124 & 0.000 \end{bmatrix}.$$

We set $\Gamma = 1.5$ and $[\mathbf{b}, \mathbf{u}] = [\mathbf{0}, \mathbf{1}]$. Figure 3 (b) and (d) plot the evolutions of the optimal value of (5) using Algorithm 1 and Algorithm 2, respectively. The optimal spectral radius is $\rho^\star = 0.8441$ and the optimal solution is $\mathbf{s}^\star = (0.525, 0.450, 0.525)^\top$.

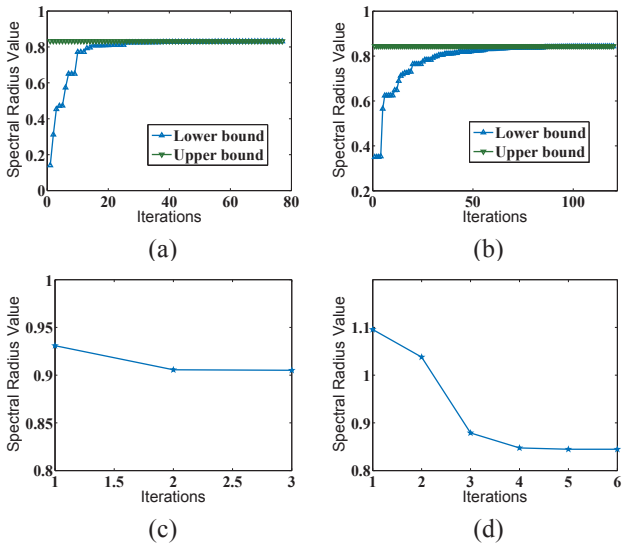

Fig. 3. Illustrations of convergence: (a) Algorithm 2 with $\mathbf{G}(\mathbf{s}) = \mathrm{diag}(\mathbf{s})\mathbf{B}$, (b) Algorithm 2 with $\mathbf{G}(\mathbf{s}) = \mathrm{diag}(\mathbf{s}) + \mathbf{B}$, (c) Algorithm 1 with $\mathbf{G}(\mathbf{s}) = \mathrm{diag}(\mathbf{s})\mathbf{B}$, and (d) Algorithm 1 with $\mathbf{G}(\mathbf{s}) = \mathrm{diag}(\mathbf{s}) + \mathbf{B}$.

Example 2: Using the Hong Kong SARS test case in [10], we compare three kinds of epidemic control on the 18 districts in Hong Kong for multi-group model with parameters: $f_{ii} = 3$, $f_{ij} = 0.57$ when districts i, j are neighbors and $f_{ij} = 0.02$ when districts i, j are not adjacent, $\Gamma = 3.6667$. The results in these three controls are 1) homogeneous control allocates the same resource to different districts yielding $\rho = 1.1323$, 2) heterogeneous control obtained from [10] yields $\rho = 1.0126$, and 3) we use Algorithm 2 with the parameter $\epsilon = 0.01$ to yield $\rho = 1.0088$.

The experiments demonstrate the value of taking advantage of the topology of interactive networks by placing more resources in more important districts. We also observe that the upper bound obtained by Algorithm 1 yields a value close to $\rho = 1.0088$ in the first few iterations of Algorithm 2, and Algorithm 1 typically converges in less than 100 iterations.

VI. Conclusion

We study a deterministic epidemic evolution control framework that requires solving a nonconvex optimization problem by minimizing a spectral radius function subject to constraints on the control variables. We first show that the optimal solution of the spectral radius minimization problem resides on the boundary of the constraint set. Then, feasible suboptimal solutions of the spectral radius minimization problem are obtained using successive convex approximation based on geometric programming. Key tools in nonnegative matrix theory, e.g., the Perron-Frobenius theorem and the Friedland-Karlin inequalities, are then leveraged to find bounds that can be tightened using the branch-and-bound method. This leads to an algorithm that can compute the global optimal solution of the spectral radius minimization problem. Numerical experiments show that our algorithms can be computationally fast in small to medium-sized networks.

References

[1] L. Chen, S. Roy, and A. Saberi, "On the information flow required for tracking control in networks of mobile sensing agents," *IEEE Trans. on Mobile Computing*, vol. 10, no. 4, pp. 519–531, April 2011.

[2] S. Roy, M. Xue, and S. K. Das, "Security and discoverability of spread dynamics in cyber-physical networks," *IEEE Trans. on Parallel and Distributed Systems*, vol. 23, no. 9, pp. 1694–1707, September 2012.

[3] D. J. Daley and J. Gani, *Epidemic modelling: An introduction*. UK: Cambridge University Press, 1999.

[4] L. Zager and G. Verghese, "Epidemic thresholds for infections in uncertain networks," *Complexity*, vol. 14, no. 4, pp. 12–25, Mar. 2009.

[5] D. Chakrabarti, Y. Wang, C. Wang, J. Leskovec, and C. Faloutsos, "Epidemic thresholds in real networks," *ACM Trans. on Information and System Security*, vol. 10, no. 4, pp. 1–26, Jan. 2008.

[6] M. Khouzani, S. Sarkar, and E. Altman, "Optimal control of epidemic evolution," *IEEE INFOCOM*, pp. 1683–1691, Apr. 2011.

[7] N. C.Valler, B. A. Prakash, H. Tong, M. Faloutsos, and C. Faloutsos, "Epidemic spread in mobile ad hoc networks: Determining the tipping point," *ECML PKDD*, vol. 6640, pp. 266–280, 2011.

[8] D. Shah and T. Zaman, "Detecting sources of viruses in network: Theory and application," *Proceedings of ACM SIGMETRICS*, pp. 14–18, Jun. 2010.

[9] P. Wang, M. C. Gonzalez, C. A. Hidalgo, and A.-L. Barabasi, "Understanding the spreading patterns of mobile phone viruses," *Science*, vol. 324, no. 5930, pp. 1071–1076, May 2009.

[10] Y. Wan, S. Roy, and A. Saberi, "A new focus in the science of networks: towards methods for design," *Proceedings of the Royal Society A: Mathematical, Physical and Engineering Science*, vol. 464, no. 2091, pp. 513–535, Mar. 2008.

[11] R. G. Gallager, *Discrete stochastic processes*. Norwell: Kluwer Academic Publishers, 1996.

[12] S. P. Banks, *Control systems engineering*. Englewood Cliffs: Prentice-Hall, 1986.

[13] K. Yeh and H. Lu, "Robust stability analysis for two-dimensional systems via eigenvalue sensitivity," *Multidimensional Systems and Signal Processing*, vol. 6, pp. 223–236, Jul. 1995.

[14] V. Balakrishnan, S. Boyd, and S. Balemi, "Branch and bound algorithm for computing the minimum stability degree of parameter-dependent linear systems," *International Journal of Robust and Nonlinear Control*, vol. 1, no. 4, pp. 295–317, Oct.-Dec. 1992.

[15] S. Boyd and L. Vandenberghe, *Convex optimization*. UK: Cambridge University Press, 2004.

[16] S. Friedland and S. Karlin, "Some inequalities for the spectral radius of non-negative matrices and applications," *Duke Mathematical Journal*, vol. 42, no. 3, pp. 459–490, Sep. 1975.

[17] J. E. Falk, "Lagrange multipliers and nonconvex programs," *SIAM Journal on Control*, vol. 7, no. 4, pp. 534–545, Nov. 1969.

[18] R. B. Bapat and T. E. S. Raghavan, *Nonnegative matrices and applications*. UK: Cambridge University Press, 1997.