Sparse Resource Allocation for Linear Network Spread Dynamics

Jackeline Abad Torres, *Member, IEEE*, Sandip Roy, *Member, IEEE*, and Yan Wan, *Member, IEEE*

***Abstract*—Sparse resource allocation to shape a network dynamical process is studied. Specifically, we consider allocating limited distributed control resources among a subset of a network's nodes, to minimize the dominant eigenvalue of a linear dynamical spread process associated with the network. Structural characterizations of the closed-loop dynamics at the optimum are obtained. These results are then used to 1) develop constructive algorithms for optimal resource allocation, 2) identify limits on the control performance, and 3) understand the relationship between the network's graph and the optimal resource profile. While the focus here is on a simplified linear model, an exploratory study of the design's applicability to realistic stochastic and nonlinear spread processes is undertaken, via simulation examples. As a whole, this study advances a research thrust on disease spread control in networks, toward the realistic paradigm that control resources can only be allocated at a subset of network locations.**

***Index Terms*—Control of networks, network theory (graphs), optimal control, spread processes.**

I. Introduction

A NUMBER of recent studies in the network controls community have considered optimal deployment of limited control resources to mitigate diffusion and spread processes in networks [1]–[8]. These efforts follow on an extensive literature on modeling and simulation of network spread dynamics, mostly published in applied-mathematics and computing forums (e.g., [9]); as well as a wide literature on diffusion and synchronization processes in the physical sciences and engineering (e.g., [10]). The study of spread management in the network-controls community was originated in [1], which formulated a basic design problem in the context of multi-group and contact-network models for spread, and provided structural characterizations and designs for symmetric and diagonally-symmetrizable interaction topologies. Several studies have addressed more general design problems and obtained results for broader graph classes, using a mix of numerical methods [2], [3], [5] and structural approaches [5]–[7]. In particular, [2]–[5] use optimization algorithms including geometric programming and the Distributed Euler Replicator Algorithm to minimize either the total resource cost or the spectral radius of the network's dynamics, under broad assumptions. The article [5] also describes a decentralized implementation of the resource-placement algorithm, and identifies an interesting structure (line-sum symmetry) that facilitates design. Meanwhile, the article [6] provides graph-theoretic results on the performance cost of using fair controllers and [7] suggests a dynamic (trend-driven) control scheme. Recently, our group has pursued application of these design techniques in managing the spread of zoonotic (mixed animal and human) diseases in agricultural communities [11] as well as hospital settings [12]. In complement with the network-controls approaches, optimal controls and game-theoretic approaches to spread management have also been proposed. Additionally, the important work [8] has proposed an efficient dynamic policy for containing epidemics in networks, which is near-optimal in the sense that the extinction time is within a multiplicative factor of the optimal for large networks.

The structural studies of optimal spread management in the network-control literature have assumed that resources can be applied at all network nodes. In many practical spread-management contexts, control capabilities are blunt and limited, effecting changes in spread patterns in a subset of network nodes, or in constrained ways. For example, when spread among multiple communities or states is considered, only some may have the wherewithal to deploy vaccination or control resources, of surveil for the disease [13]. Similarly, the ability to effect controls may differ among the multiple species involved in a spread [11], the ages or other personal characteristics of the infectives, or the locations visited by the infectives, among other factors. Also, in many circumstances, cost constraints dictate that resources can only be placed in a subset of network locations, although these location may be the choice of the control designers. This article addresses spread control when only a subset of network locations allow control.

While the design of sparse controls can be addressed using numerical methods (see e.g., the broad formulation in [2]), there is a need for 1) simple design algorithms that are catered for sparse-control problems and 2) simple structural insights into good control placements and performance bounds. The research presented here addresses these needs. Specifically, we study the optimal distributed allocation of control resources over a subset of network nodes to shape a linear network spread dynamics, subject to a total resource constraint. For this class of control problems, we characterize the dominant

Manuscript received July 7, 2015; revised July 8, 2015, February 23, 2016, and February 24, 2016; accepted July 8, 2016. Date of publication July 21, 2016; date of current version March 27, 2017. This work was supported in part by the National Science Foundation under Grant ECS-0901137, Grant CNS-1035369, and Grant CNS-1058124. Recommended by Associate Editor C. Altafini.

J. Abad Torres is with the Departamento de Automatización y Control Industrial, Facultad de Ingeniería Eléctrica y Electrónica, Escuela Politécnica Nacional, E11-253 Quito, Ecuador (e-mail: jackeline.abad@epn.edu.ec).

S. Roy is with the School of Electrical Engineering and Computer Science, Washington State University, Pullman, WA 99164 USA (e-mail: sroy@eecs.wsu.edu).

Y. Wan is with University of North Texas, Denton, TX 762013 USA (e-mail: yan.wan@unt.edu).

Color versions of one or more of the figures in this paper are available online at http://ieeexplore.ieee.org.

Digital Object Identifier 10.1109/TAC.2016.2593895

eigenvalue and corresponding eigenvectors of the closed loop, upon use of an optimal control scheme. These characterizations are then used to obtain explicit designs for the optimizing controller. The analyses and design procedures also provide some insight into the relationship between the network's graph and control performance. Our development draws on structural results given in [1] and [5], but generalizes these approaches to the sparse-control case. In the process, the work also obtains structural design results for more general (specifically, directed irreducible) graph topologies.

The design method presented here potentially can inform management of infectious disease outbreaks, as well as spread and diffusion processes in the cyber world. However, these processes are in reality much more complex than captured by the linear network model, involving significant nonlinearities, intrinsic stochastics, topological variations, delays, etc [14]. Thus, applying the optimization methods requires, at the least, evaluation of the design by detailed simulation, and consideration of whether the design provides simple structural insights that are informative to real-world management. In this article, an exploratory study of the method's applicability to infectious-disease management is undertaken, using two examples. First, the optimal control policy is implemented and evaluated in the context of a stochastic nonlinear multi-group model for spread [15], [16], for a 100-node network. Second, a small-scale case study of the spread of the vector-borne Chikungunya virus is developed, which gives some structural insight into resource placement when control locations are limited.

The research described here also connects to a much broader literature on decentralized and network control. Decentralized control of large-scale dynamical systems has been addressed extensively in the classical controls literature [17]–[20]; we particularly point out the important works giving non-conservative conditions for stabilization of general linear decentralized systems (e.g., [17]). However, much remains to be done in constructing practical high-performance decentralized controllers that exploit the network's interconnection topology, and also in relating control performance to topological parameters. In a complementary direction, there has been also an extensive graph-theoretic study of dynamical networks, centered on their modeling and analysis rather than control [21]–[23]. Over the last 15 years, an exciting literature has developed that meshes these two perspectives, and approaches decentralized control from a graph-theoretic perspective [24]–[28]. However, this literature on graph-exploiting decentralized controllers has primarily focused on autonomous-agent systems, where network nodes or agents are only interconnected via communications (not hard-wired), and controllers can be developed for all agents.

The results for autonomous-agent systems have found wide use (e.g., in robotics and sensor-network computations), but many infrastructural applications demand control of an existing hard-wired network process using sparse localized controls, perhaps subject to stringent resource constraints. Applications requiring such sparse control include not only the disease-spread processes that are our focus here [29], but also swing dynamics in the electric power grid [30], air traffic flow management [31], etc. Although control schemes have been implemented in the real world for many of these infrastructures, research on these hard-wired control problems for a graph-exploiting perspective is surprisingly sparse. The literature on *pinning control* is an important contribution in this direction [32]–[34], however these studies do not model resource constraints, and largely do not explicitly consider tuning of control gains. Recently, the problem of designing a subset of edge weights in a graph to shape a dynamics has been considered [35], which shows interesting relationships between the undesignable network structure and the network's spectrum. However, this study was focused on edge-weight design rather than control. The growing drive toward fine control of infrastructural processes (e.g., for building energy management, coordination of air traffic management initiatives, and epidemic mitigation) motivates the study of graph-exploiting control of hardwired processes under resource constraints. The research presented here contributes to this direction, with a focus particularly on controlling spread dynamics with limited control flexibility.

In terms of the methods used, this study particular draws on, and advances, a research effort on line-sum symmetry and dominant eigenvalue minimization under trace preserving diagonal perturbation [36], [37]. This work also builds on the broad literature related to nonnegative matrices and stability analysis of nonnegative systems (e.g., [38], [39]). Relative to these studies, our efforts are concerned with *optimizing* the dominant eigenvalue over a constrained set of diagonal perturbations, particularly for the special case where only a subset of diagonal entries can be designed.

The remainder of this article is organized as follows. In Section II, we formulate the problem of allocating control resources to manage a linear network spread dynamics. In Section III, results on the optimal allocation of control resources are presented. Finally, an exploratory study of the method's applicability to nonlinear infectious-disease spread processes is undertaken, in Section IV.

II. Model and Problem Formulation

A wide range of network spread and diffusion processes have been modeled, including disease spread, computer virus dissemination, population dynamics, oil spills on oceans, data multicast, etc. [40]–[44]. These studies have introduced nonlinear deterministic and stochastic models for network spread dynamics, and considered linear approximations of these models in detail. Since there is already an extensive literature, we directly present a generic linear model that encompasses a variety of linearized spread processes, without re-deriving the models and approximations. The described model is particularly closely connected with the models for infection spread given in [1], [2], and [5], see these works for detailed derivations of the linearized network dynamics.

Formally, let us consider a network with n nodes. Each network node, labeled $i \in 1, 2, \ldots, n$, has a non-negative scalar state $x_i(t)$ associated with it which indicates the prevalence of an *infective* (spreading quantity) in that node. Nominally, the states are approximated as evolving according to an autonomous linear differential equation

$$\dot{x}(t) = Ax(t) \tag{1}$$

where $t \geq 0$ and $x(t) = [x_1(t), \ldots, x_n(t)]'$. The off-diagonal entries $A_{ij} = \alpha_{ij}$ of the state matrix A, which indicate the unitized transmission rate of the infection from j to i, are assumed to be nonnegative ($\alpha_{ij} \geq 0 \; \forall \, i \neq j$). Meanwhile, the diagonal entries may be either positive or negative, reflecting the possibility for local spread as well as prevalence reduction due to local control or intrinsic "healing." State matrices of this form, known as *essentially-nonnegative* or *Metzler* matrices, arise in a wide range of network dynamics (including synchronization, disease spread, and flow processes) [1], [21], [45]. The model encompasses both stable and unstable processes. When essentially-nonnegative matrices are used to represent network processes, it is common to associate a directed graph with the matrix, which indicates direct influences among the nodes (as indicated by $\alpha_{ij} > 0$). Relationships between the connectivity properties of this graph and certain eigenvalues/eigenvectors of A are well-known, see e.g. [46]. For many of the analyses here, we will assume that the state matrix A is *irreducible*, i.e., corresponding to a strongly-connected graph. Notice that a nonnegative irreducible matrix has a real dominant eigenvalue with algebraic and geometric multiplicity one. We do not include a detailed review of the graph-theoretic analysis, see [47].

We are concerned with the allocation of distributed control resources to a subset of the nodes in a network. Our main goal is to allocate these resources, subject to limits, so as to optimize a performance measure. Specifically, without loss of generality, we assume that control resources may be placed in nodes $i \in 1, 2, \ldots, m$ ($m < n$). The resources placed at each node i are modeled as enacting a local linear feedback in node i ($i \in 1, \ldots, m$), whose strength is proportional to the amount of resource placed. Specifically, for a resource level D_i in node i, the ith row of the nominal state equation is modified by the additive feedback input $u_i(t) = -D_i x_i(t)$ for $i = 1, \ldots, m$. We note here that a more positive D_i indicates a larger resource allocation, which enacts a stronger negative feedback to curtail the spread.

Upon placement of the control resources, the closed-loop dynamics are

$$\dot{x}(t) = (A - K)x(t) \tag{2}$$

where

$$K = \begin{bmatrix} D & 0 \\ 0 & 0 \end{bmatrix}$$

and $D = \mathrm{diag}(D_1, \ldots, D_m)$. The closed-loop state matrix remains an essentially-nonnegative matrix, with the same graph (and hence same irreducibility property) as the original state matrix. In the case where the open-loop state matrix is irreducible, it follows that the closed-loop state matrix is irreducible for any K, and hence has a purely real eigenvalue whose real part is strictly larger than that of any other eigenvalues. We refer to this eigenvalue as the dominant eigenvalue, and use the notation $\lambda_{\max}$ for it. The dominant eigenvalue may be either positive or negative. We also refer to the corresponding left- and right-eigenvectors at $w_{\max}$, $v_{\max}$ and the ith node of these eigenvectors as $w_{\max,i}$, $v_{\max,i}$ respectively.

Our focus here is on reducing (making more negative) the dominant eigenvalue of the closed-loop dynamics subject to resource constraints, since this dominant eigenvalue determines stability/instability and specifies the convergence rate of the dynamics. Specifically, we seek to design the diagonal matrix D to minimize dominant (maximum) eigenvalue of $A - K$, $\lambda_{\max}(A - K)$, subject to the total resource constraint $0 \leq \sum_{i=1}^{m} D_i \leq \Gamma$. We also address this design problem under the further constraint that only negative feedback is allowed, i.e., $D_i \geq 0 \; \forall \, i = 1, 2, \ldots, m$. Formally, the design problem of interest is the following:

$$\begin{aligned} \min_{D_1,\ldots,D_m} \quad & \lambda_{\max}(A - K) \\ \text{s.t.} \quad & \sum_{i=1}^{m} D_i \leq \Gamma \\ & D_i \geq 0 \quad \forall \, i = 1, 2, \ldots, m \end{aligned} \tag{3}$$

where both the case with the individual constraints on D_i and the case without are considered. The eigenvalue-optimization problem considered here closely follows on the formulations given in [1], [2], and [5], but addresses the case that only a subset of network nodes can be allocated control resources from a structural perspective.

Although the above problem formulation focuses on a linear model, outcomes of the design have some bearing on more accurate nonlinear models for spread processes. Importantly, the stability vs. instability of the linear model is known to correspond to elimination vs. persistence of spread in the nonlinear model under broad conditions [9], [14], hence the asymptotics of spread for the closed-loop nonlinear model are characterized. Specifically, the stability of the linear model guarantees that the *basic reproduction ratio* (average number of new infections caused by a randomly-chosen infective) is less than one for the non-linear stochastic multi-group model, which is sufficient for stability of the origin for the nonlinear model; similarly, instability implies that the nonlinear model will have a stable fixed point away from the origin, i.e., a persistent spread. Also, for multi-group and certain contact-network models for spread, the state dynamics of the linear model are known to majorize those of the nonlinear one, so the design for the lincariation bounds the performance of the nonlinear model. However, it is important to stress that the transient dynamics of the nonlinear and linear models differ significantly, and thus evaluation of designed controls for the nonlinear model is necessary. An exploratory evaluation of the design in the context of a stochastic nonlinear model for spread is undertaken in Section IV.

III. Structural and Algorithmic Results

We refer to the design problem introduced in Section II as the *subset design problem* since only a subset of network nodes (states) have control input capabilities. We also distinguish the special case when $m = n$, i.e., all the network nodes have control input capabilities, as the *network-wide design problem*. Additionally, we find it convenient to distinguish: 1) *unconstrained design*, where the D_i, for $i = 1, \ldots, m$, are unconstrained except for the total resource constraint (Section III-A); and 2) *constrained design*, where the additional constraint that D_i is non-positive for $i = 1, \ldots, m$, is enforced (Section III-B).

For convenience, we refer to the diagonal matrix K that minimized the dominant eigenvalue of $A-K$ as $\bar{K}$ (unconstrained design) or K^* (constrained design). The dominant eigenvalue and corresponding eigenvectors of $A-K^*$ (or $A-\bar{K}$) are referred to as $\lambda^*_{\max}$, $w^*_{\max}$, and $v^*_{\max}$ (or equivalently as $\bar{\lambda}_{\max}$, $\bar{w}_{\max}$, and $\bar{v}_{\max}$).

Conceptually, the unconstrained design differs from the constrained one in that resources can be "taken away" (negative resources allocated) to some network locations, to allow provision of sufficient resources at other locations. In many cases, such reallocation of resources is not realistic, since the *ab initio* resources may reflect intrinsic healing capabilities or unchangeable policies. However, the unconstrained design may be relevant in some applications (e.g., computational ones), and it is also important as a performance bound and stepping stone for the constrained design.

A. Unconstrained Design: Optimal Resource Allocation

In this subsection, the unconstrained design is addressed from a structural perspective. Specifically, Lemma 1 characterizes the structure and sign pattern of the left- and right-eigenvectors associated with the dominant eigenvalue, for the optimal design. Using this result, an algebraic method for computing the subset optimal control resource allocation is developed in Theorem 2, for arbitrary essentially non-negative irreducible matrix A. Discussions about the network-wide optimal resource distribution and about the minimum resource level necessary for stabilization are also included.

Let us start by analyzing the dominant eigenvalue and eigenvectors of $A-K$, when an optimal control resource allocation $K=\bar{K}$ is used:

Lemma 1—(Unconstrained Optimization: Spectral Condition): Consider the matrix $A-K$, where $K=\mathrm{diag}(D_1,\ldots,D_m,0,\ldots,0)$ and A is an essentially-nonnegative matrix as defined in the problem formulation (which may or may not be irreducible). Consider any $K=\bar{K}$ that minimizes the dominant eigenvalue of $A-K$ subject to $\sum_{i=1}^m D_i\le\Gamma$. Assume that $A-\bar{K}$ has a real simple dominant eigenvalue. The optimizing $\bar{K}$ and the corresponding dominant eigenvalue/eigenvectors $\bar{\lambda}_{\max}$, $\bar{w}_{\max}$, and $\bar{v}_{\max}$ satisfy one of the following conditions:

1) $\sum_{i=1}^m \bar{D}_i=\Gamma$. In this case, $\forall\, i=1,\ldots,m$, we have that $\bar{w}_{\max,i}\bar{v}_{\max,i}=\bar{\mu}$.
2) $\sum_{i=1}^m \bar{D}_i<\Gamma$. In this case, $\forall\, i=1,\ldots,m$, we have that $\bar{w}_{\max,i}\bar{v}_{\max,i}=0$.

Furthermore, if A is irreducible, then $A-\bar{K}$ has a real simple dominant eigenvalue, and the optimizing $\bar{K}$ and the dominant eigenvectors always satisfy condition 1.

Proof: We use eigenvalue sensitivity analysis and constrained optimization concepts [48], [49] for this proof. To find the optimum $\bar{K}$ under the total resource constraint, we form the Lagrangian: $L=\lambda_{\max}(A-K)+\mu(\sum_{j=1}^m D_i-\Gamma+\eta^2)$, where μ is the Lagrange multiplier and η the slack variable. Then we recognize that the derivative of L with respect to each variable (D_i, μ, and η) is zero at the optimum (Karush-Kuhn-Tucker (KKT) conditions). Since the dominant eigenvalue of $A-\bar{K}$ is assumed real and nonrepeated, we can use the standard eigenvalue sensitivity formula for the derivative of $\lambda_{\max}$ with respect to the parameters, i.e., $(\partial\lambda_{\max}(A-K))/\partial D_i=-w_{\max,i}v_{\max,i}$ for $i=1,2,\ldots,m$. This leads to the following equations:

$$\begin{aligned}\bar{w}_{\max,i}\bar{v}_{\max,i}-\bar{\mu}&=0\quad\forall\, i=1,2,\ldots,m\\ \sum_{j=1}^m \bar{D}_i+\eta^2&=\Gamma\\ 2\bar{\mu}\eta&=0.\end{aligned}\tag{4}$$

The two cases in the Lemma follow directly from (4). First, we note that either $\bar{\mu}$ or η is zero. If $\eta=0$ and $\bar{\mu}>0$ then $\sum_{i=1}^m \bar{D}_i=\Gamma$ and $\bar{w}_{\max,i}\bar{v}_{\max,i}=\bar{\mu}$ whereas if $\eta=0$ and $\bar{\mu}=0$ then $\sum_{i=1}^m \bar{D}_i=\Gamma$ and $\bar{w}_{\max,i}\bar{v}_{\max,i}=0$. Otherwise, $\sum_{i=1}^m \bar{D}_i<\Gamma$ and $\bar{w}_{\max,i}\bar{v}_{\max,i}=0$. The two conditions in this Lemma provide a necessary condition on $\bar{K}$ to minimize the dominant eigenvalue of $A-K$. However, this condition is also sufficient since the dominant eigenvalue of an essentially-nonnegative matrix is a convex function of its diagonal entries [50]. Additionally, if A is irreducible, then the dominant eigenvalue is non-repeated, and the corresponding eigenvectors are entry-wise positive. Hence, condition 1 is always satisfied. ■

Remark 1: Condition 2 in Lemma 1 corresponds to the case that any increment/decrement in $\bar{D}_i$ for $i=1,\ldots,m$ will not change the maximum/dominant eigenvalue of $A-\bar{K}$. This case arises if the matrix A is reducible and the locations in the network of all the controls are such that they do not affect the dominant eigenvalue, i.e., $\bar{w}_{\max,i}\bar{v}_{\max,i}=0$ for $i=1,\ldots,m$. In such a case, any matrix K such that $\sum_{i=1}^m D_i<\Gamma$, is a solution, including the trivial one $K=0$. This is, however, an atypical case.

In the typical case that the control inputs can move the dominant eigenvalue, the entirety of the resources are allocated in such a way as to equalize the ***participation factors*** $\bar{w}_{\max,i}\bar{v}_{\max,i}$, i.e., to equalize the impact of each network node on the spread. In this case, the entirety of the available resources are necessarily used, i.e., $\sum_{i=1}^m \bar{D}_i=\Gamma$.

Remark 2: We stress that, for the case where A is irreducible, the dominant eigenvalue of $A-\bar{K}$ is necessarily simple and hence the theorem is applicable (as is explicit in the theorem statement). In the reducible case, a simple dominant eigenvalue is also typical and can be guaranteed in subcases, however specialized examples can be developed with repeated (semi-simple) dominant eigenvalue. The theorem can be generalized to encompass these cases, but the treatment becomes rather technical and is not included here.

Let us now present an algorithm for designing the optimal resource allocation, for irreducible state matrices A. The design result draws on an interesting transformation for irreducible essentially-nonnegative matrices. Specifically, there always exists a *unique* (up to scale factor) diagonal similarity transformation matrix P such that $PAP^{-1}\vec{1}=P^{-1}A'P\vec{1}$, where $\vec{1}$ is the all ones vector of the appropriate dimension. These kind of matrices, called *line-sum symmetric*, have interesting properties that are important in the development of our results (see [36], [51] for details on the existence and computation of P). We note that the line-sum symmetry concept differs from diagonal symmetrizability. Line-sum symmetry simply requires each row

sum to equal the corresponding column sum, and in fact every irreducible nonnegative matrix has a diagonal transformation that achieves line-sum symmetry. In contrast, a matrix is said to be diagonally symmetrizable only if it has a diagonal transformation which makes it truly symmetric. Also, we note that line-sum symmetry was exploited in optimizing a spread controller in [5], however their design did not achieve design for an arbitrary irreducible state matrix (arbitrary directed graph).

The following theorem gives an algorithm for computing the optimal resource allocation, for the unconstrained subset design problem. This development requires some further notation. Specifically, we partition the topology matrix into four submatrices A_{11}, A_{12}, A_{21}, and A_{22} which are commensurate with the designable and undesignable control channels (i.e., A_{11} is an $m \times m$ matrix).

Theorem 2—(Unconstrained Subset Optimal Design): Consider the matrix $A-K$, where $K=\mathrm{diag}(D_1,\ldots,D_m,0)$, and A is an irreducible essentially-nonnegative matrix as defined in the problem formulation. The control design $K=\bar{K}$ that minimizes the dominant eigenvalue of $A-K$ subject to the only constraint that $\sum_{i=1}^{m} D_i \leq \Gamma$ can be found as follows:

1) Solve the following system of equations for P and $\bar{\lambda}_{\max}$:

$$\begin{aligned} PA_r(\bar{\lambda}_{\max})P^{-1}\vec{1}-P^{-1}A_r(\bar{\lambda}_{\max})'P\vec{1}=0 \\ \vec{1}'PA_r(\bar{\lambda}_{\max})P^{-1}\vec{1}-\Gamma-\bar{\lambda}_{\max}\vec{1}'\vec{1}=0 \end{aligned} \quad (5)$$

where $A_r(\bar{\lambda}_{\max})=A_{11}+A_{12}(\bar{\lambda}_{\max}I-A_{22})^{-1}A_{21}$, and $P=\mathrm{diag}(p_1,\ldots,p_m)$

2) Calculate $\bar{D}=\mathrm{diag}(PA_r(\bar{\lambda}_{\max})P^{-1}\vec{1}-\bar{\lambda}_{\max}\vec{1})$ and

$$K=\begin{bmatrix}\bar{D} & 0\\ 0 & 0\end{bmatrix}.$$

Proof: Since the optimal solution satisfies that $\bar{w}_{\max,i}\bar{v}_{\max,i}$ are identical for all $i=1,\ldots,m$, we can use a diagonal similarity transformation such that ith right- and left-eigenvector component are identical, or equivalently

$$\begin{bmatrix}P & 0\\ 0 & I\end{bmatrix}\begin{bmatrix}A_{11}-\bar{D} & A_{12}\\ A_{21} & A_{22}\end{bmatrix}\begin{bmatrix}P^{-1} & 0\\ 0 & I\end{bmatrix}\begin{bmatrix}\vec{1}\\ \vec{v}\end{bmatrix}=\bar{\lambda}_{\max}\begin{bmatrix}\vec{1}\\ \vec{v}\end{bmatrix}$$

and

$$\begin{bmatrix}P^{-1} & 0\\ 0 & I\end{bmatrix}\begin{bmatrix}A_{11}-\bar{D} & A_{12}\\ A_{21} & A_{22}\end{bmatrix}'\begin{bmatrix}P & 0\\ 0 & I\end{bmatrix}\begin{bmatrix}\vec{1}\\ \vec{w}\end{bmatrix}=\bar{\lambda}_{\max}\begin{bmatrix}\vec{1}\\ \vec{w}\end{bmatrix}$$

where $\vec{w}$ and $\vec{v}$ are the entries of the left- and right-eigenvectors related with the entries $i=m+1,\ldots,n$. From the above eigenvalue equations, we find that $PA_r(\bar{\lambda}_{\max})P^{-1}\vec{1}=P^{-1}A_r(\bar{\lambda}_{\max})'P\vec{1}$, i.e., $PA_r(\bar{\lambda}_{\max})P^{-1}$ is line-sum symmetric. Furthermore, the last equation of the system of equations (5) and the expression for $\bar{D}$ are derived from $PA_r(\bar{\lambda}_{\max})P^{-1}\vec{1}-\bar{D}\vec{1}=\bar{\lambda}_{\max}\vec{1}$. ■

The system of equations (5) can be solved quickly using simple numerical methods. Specifically, the first m equations in (5) involve finding a diagonal matrix P that transforms the matrix $A_r(\bar{\lambda}_{\max})$ into a line-sum symmetric matrix: this problem can be solved using a fast algorithm [51], [52], and hence it only remains to scan over possible values of $\bar{\lambda}_{\max}$ to find the optimum.

Remark 3: If the matrix A is symmetric, it is clear that $P=I$ in (5), and hence the system of equations is reduced to only a single polynomial equation whose unknown is $\bar{\lambda}_{\max}$.

The above result specializes in the case that control resources are available everywhere, i.e., a network-wide design is possible.

Corollary 1—(Unconstrained Network-Wide Optimal Design): Consider the matrix $A-K$, where K is a diagonal matrix and A is an essentially-nonnegative irreducible matrix, as defined in the problem formulation. The matrix $\bar{K}$ that minimizes the dominant eigenvalue of $A-K$ subject to the constraint $\sum_{i=1}^{n} D_i=\Gamma$ can be found as follows.

1) Find a diagonal matrix P such that $PAP^{-1}\vec{1}=P^{-1}A'P\vec{1}$.
2) Compute $\bar{\lambda}_{\max}=(1/n)(\vec{1}'PAP^{-1}\vec{1}-\Gamma)$.
3) Compute $\bar{K}=\bar{D}=\mathrm{diag}(PAP^{-1}\vec{1}-\bar{\lambda}_{\max}\vec{1})$.

The design result also immediately allows us to determine the minimum resource level necessary for stability of the spread dynamics:

Corollary 2: The minimum resource level Γ such that a controller satisfying $\sum_{i=1}^{n} D_i=\Gamma$, achieves stability (i.e., makes $\lambda_{\max}\leq 0$) is $\Gamma_{\min}=\vec{1}'PAP^{-1}\vec{1}$. If $\vec{1}'PAP^{-1}\vec{1}\leq 0$, then the open-loop network model is already stable.

Conceptually, Corollary 2 indicates that the disease can be eliminated if enough resources are available to make the average row sum of the network matrix (average rate of new-infection generation compared to healing) negative, upon line-sum-symmetrization. We note that it is sufficient for the average row sum of the original (non line-sum-symmetrized) network matrix to be negative, although not necessary. For large networks with homogeneous inter-individual spread rates, it is easy to check that the required resource budget grows linearly with the density of interconnections, and the number of vertices; this is in contrast with the dynamic allocation considered in [8], which allows for a sublinear resource budget.

Theorem 2 and its associated lemmas/corollaries reflect a structural characteristic of the optimal design, namely that the optimal $\bar{D}$ equalizes the row sum of $PA_r(\lambda_{\max})P^{-1}-D$. Such an optimal design equalizes the participation of each network node on the dominant eigenvalue, and hence satisfies the conditions of the optimal solution given in Lemma 1. The minimal resource Γ required for stability therefore makes all of the row sums of the line-sum-symmetrized state matrix equal to zero.

The design presented in Theorem 2 also admits a graph-theoretic interpretation. To explain this, let us consider the matrix $A_r(\bar{\lambda}_{\max})=A_{11}+A_{12}(\bar{\lambda}_{\max}I-A_{22})^{-1}A_{21}$ as defined in the theorem statement. This Schur form-matrix is essentially nonnegative, and can be viewed as defining a graph with vertices corresponding to the designable subset, which we call the ***effective contact network graph***. We note that $A_r(\bar{\lambda}_{\max})$ encapsulates influences among designable nodes, but includes indirect influences via the undesignable part. Specifically, an edge (i,j) in the effective contact network graph describes the transmission rate of the infection from i to j considering not only the direct influence of i over j but also the influence through the undesignable part of the network. According to Lemma 1, the *participation factor* of every designable network

node (state) in the dominant mode is identical at the optimum. Then, Theorem 2 admits the following interpretation: in the contact network dynamics for virus spread, the optimal control resource distribution injects control resources in the designable subset such that influence of these nodes is equalized with respect to the effective contact network graph.

B. Constrained Design: Optimal Resource Allocation

Next, the constrained design problem is addressed, by building on the results for the unconstrained design. First, the structure and sign pattern of the left- and right-eigenvectors associated with the dominant eigenvalue at the optimum are characterized in Lemma 3. Next, in Lemma 5, we present intrinsic structural limitations on control performance, for the constrained subset design problem. A necessary and sufficient condition for checking whether a control resource allocation is globally optimal is then presented in Lemma 4, based on the element-wise product of the eigenvector components associated with the dominant eigenvalue of $A-K^*$. Finally, a finite-stage iterative method is developed based on these lemmas, to compute the optimal control resource allocation when the state matrix is diagonally symmetrizable (Theorem 6).

The following lemma characterizes the dominant eigenvectors of the optimal closed-loop state matrix $A-K^*$ for the constrained design problem, for both the network-wide and subset design cases.

Lemma 3—(Constrained Optimization: Spectral Analysis): Consider the matrix $A-K$, where $K=\mathrm{diag}(D_1,\ldots,D_m,0,\ldots,0)$ and A is an essentially-nonnegative matrix as defined in the problem formulation. Consider any $K=K^*$ that minimizes the eigenvalue of $A-K$ subject to the constraints: 1) $\sum_{i=1}^m D_i\le\Gamma$ and 2) $D_i\le 0\ \forall\, i=1,2,\ldots,m$, and assume that $A-K$ has a real simple dominant eigenvalue for $K=K^*$. The optimizing K^* and the corresponding dominant eigenvalue/eigenvectors $\lambda^*_{\max}$, $w^*_{\max}$, and $v^*_{\max}$ satisfy one of the following conditions.

1) $\sum_{i=1}^m D_i^*=\Gamma$. In this case, for each i, we either have that $D_i^*>0$ and $w^*_{\max,i}v^*_{\max,i}=\mu^*$, or $D_i^*=0$.
2) $\sum_{i=1}^m D_i^*<\Gamma$. In this case, for each i, we either have that $D_i^*>0$ and $w^*_{\max,i}v^*_{\max,i}=0$, or $D_i^*=0$.

Furthermore, if A is irreducible, then the optimizing K^* and the dominant eigenvectors always satisfy condition 1.

Proof: Let us express D_i as $D_i=d_i^2$, since $D_i\ge 0$ for all $i=1,2,\ldots,m$. We will use eigenvalue sensitivity analysis and constrained optimization [48], [49]. That is, to find the optimum K^* under the mentioned constraints, we form the Lagrangian: $L=\lambda_{\max}(A-K)+\mu(\sum_{j=1}^m d_i^2-\Gamma+\eta^2)$, where μ is the Lagrange multiplier and η the slack variable. Then we set the derivatives of L with respect to all variables (d_i, μ, and η) to zero (KKT conditions), which leads to the following equations:

$$\begin{aligned} -2d_i^*\left(w^*_{\max,i}v^*_{\max,i}-\mu^*\right)&=0\quad\forall\, i=1,2,\ldots,m\\ \sum_{j=1}^m {d_i^*}^2+{\eta^*}^2&=\Gamma\\ 2\mu^*\eta^*&=0. \end{aligned}\tag{6}$$

The two cases in the theorem follow directly from (6). First, we note that either η^* or μ^* is zero. If $\eta^*=0$ and $\mu^*>0$ then $\sum_{i=1}^m D_i=\Gamma$ and $w^*_{\max,i}v^*_{\max,i}=\mu^*$ for each i, where $D_i>0$ or $D_i=0$. If $\eta=0$ and $\mu=0$ then $\sum_{i=1}^m D_i=\Gamma$ and $w^*_{\max,i}v^*_{\max,i}=0$ for each i where $D_i>0$ or $D_i=0$. On the other hand, if $\eta\neq 0$ and $\mu=0$ then $\sum_{i=1}^m D_i<\Gamma$ and $w^*_{\max,i}v^*_{\max,i}=0$ for each i where $D_i>0$ or $D_i=0$. Additionally, if A is irreducible then the dominant eigenvectors are entry-wise positive and hence condition 1 is always satisfied. ■

Remark 4: Similar to the unconstrained case, the typical circumstance (which always holds for the irreducible case) is that the entirety of the available resource is used, whereupon the designable locations' participations are either equalized or zero resources are allocated. If none of the control variables affect the dominant eigenvalue, which can only possibly happen if the state matrix is reducible, then trivially the resources can be allocated at will to achieve the optimal.

In the following lemma, we give a further characterization of the dominant eigenvector's components. This analysis is a stepping stone toward our design since it allows us to check whether or not a solution K for the constrained design is optimal.

Lemma 4—(Global Optimum: Spectral Condition/Pattern): Consider the matrix $A-K$ where $K=\mathrm{diag}(D_1,\ldots,D_m,0)$, and A is the irreducible matrix defined in the problem formulation. A matrix $K=K^*$ minimizes the dominant eigenvalue of $A-K$ subject to $\sum_{i=1}^m D_i=\Gamma$ and $D_i\ge 0$ for all $i=1,\ldots,m$, if and only if the product of the ith entries of the left- and right-dominant eigenvectors $w^*_{\max,i}v^*_{\max,i}$ for all $i=1,\ldots,m$ has a special pattern: specifically, $w^*_{\max,i}v^*_{\max,i}<w^*_{\max,j}v_{\max,j}$ for all $i,j\in 1,\ldots,m$ such that $D_i^*=0$ and $D_j^*>0$.

Proof: Let K^* be the matrix that minimizes the optimal eigenvalue of $A-K$. Let $K-K^*=\Delta$, where $\Delta=\mathrm{diag}(\delta_1,\ldots,\delta_m,0,\ldots,0)$ such that $\sum_{i=1}^m\delta_i=0$. Let the set $\mathcal{I}_-$ be the index set containing the indices i such that $\bar{D}_i=0$ for $i=1,\ldots,m$. Without loss of generality, let $\delta_l=-\delta$ for an arbitrary $l\in\mathcal{I}_-$, i.e., the diagonal entry l of K is less than 0, and $\delta_i=0$ for $i\neq l$ and $i\in\mathcal{I}_-$. Then $\partial\lambda_{\max}(A-K^*)/\partial\delta=-w^*_{\max,l}v^*_{\max,l}+\sum_{i=1,i\neq l}^m w^*_{\max,i}v^*_{\max,i}\partial\delta_i(\delta)/\partial\delta=-w^*_{\max,l}v^*_{\max,l}+\mu^*$. Because of the optimality condition for convex functions, i.e., $\partial\lambda_{\max}(A-K^*)/\partial\delta\ge 0$, and the fact that $w^*_{\max,l}v^*_{\max,l}\neq\mu^*$, we have that K^* minimizes $\lambda_{\max}(A-K)$ if and only if $w^*_{\max,l}v^*_{\max,l}<\mu^*$. ■

When control resources are applied only to a subset of network nodes, there is always an intrinsic limit on the network's dynamical performance (and specifically on the dominant eigenvalue), regardless of the amount of resources available. The following lemma makes this limitation explicit:

Lemma 5—(Limits on the Optimum): Consider an irreducible essentially-nonnegative state matrix A. Let $\tilde{A}$ be the submatrix of A formed by the rows/columns $i=m+1,m+2,\ldots,n$. Let $\lambda_{\max}(A)$ and $\lambda_{\max}(\tilde{A})$ be the dominant eigenvalue of A and $\tilde{A}$ respectively. Then, the dominant eigenvalue $\lambda^*_{\max}$ of the optimal control resource allocation $(A-K^*)$ always satisfies: $\lambda_{\max}(\tilde{A})\le\lambda^*_{\max}\le\lambda_{\max}(A)$, for any amount

of total control resources Γ (and for either the constrained or unconstrained design).

Proof: Since the diagonal matrix K has nonnegative entries and the dominant eigenvalue of $A-K$ is convex with respect to the diagonal entries, a solution satisfying Lemma 3 (or Lemma 1) cannot increase the dominant eigenvalue ($\lambda^*_{\max} \leq \lambda(A)$).

Suppose that we have unlimited control resources. Then, we can set $D_i = O(f)$ for a big enough f and $i = 1, \ldots, m$. According to Gerschgorin's circle Theorem [53], $A-K$ has m negative eigenvalues of order $O(f)$. Clearly, none of them is the dominant eigenvalue of $A-K$ since the remaining $n-m$ eigenvalues are within the circles whose centers are at $A_{ii} > -O(f)$ and their corresponding radius are $\sum_{j=1, j\neq i}^{n} |A_{ij}|$ for $i = m+1, \ldots, n$. Further, since A is an irreducible essentially-nonnegative matrix, $\lambda_{\max}(\tilde{A}) \leq \lambda_{\max}(A)$, see [38], [54] for details. At the same time, $\lambda_{\max}(\tilde{A}) \leq \lambda_{\max}(A-K)$ since $\tilde{A}$ is also a submatrix of $A-K$. Hence, $\lambda_{\max}(\tilde{A}) \leq \lambda^*_{\max} \leq \lambda(A)$. ■

Remark 5: The above theorem not only provides a bound on the dominant eigenvalue of $A-K^*$ (optimal control resource distribution) for the subset design problem, but also informs sensor and actuation placement. For instance, if one has the freedom to choose the m network nodes that will receive the control resources, the best choice is to make the dominant eigenvalue of submatrix $\tilde{A}$ the minimum possible, which is a direct consequence of Lemma 5.

The following theorem gives a systematic algorithm for finding the optimal diagonal matrix K^* that minimizes the dominant eigenvalue of $A-K$ for the constrained subset design problem, through an iterative method. The concept of the algorithm is as follows. First, the unconstrained optimal design is obtained. Then, iteratively, the gains that violate the positivity constraint are set to 0, and the remaining gains are re-optimized (via solving a smaller unconstrained subset design problem). This process is continued until a strictly positive solution is obtained. The below theorem verifies that the algorithm necessarily finds the optimal design in the case that the topology matrix is diagonally symmetrizable. In practice, we have been able to modify the algorithm to solve the subset design problem for arbitrary topologies, see the discussion following the theorem.

Theorem 6—(Constrained Subset Design): Consider an irreducible essentially-nonnegative and diagonally symmetrisable state matrix A. Let K^* be the diagonal matrix (with first m diagonal entries non-zero) that minimizes the dominant eigenvalue of $A-K$, subject to $D_i \geq 0$ and $\sum_{i=1}^{m} D_i = \Gamma$. The diagonal matrix K^* can be found as follows.

1) *Initialization*: Prior to the iterative process, find a diagonal matrix Q such that $\hat{A} = Q^{-1}AQ$ is symmetric. Set the iteration counter to $k=1$. Also, define the set $\mathcal{I}_D$ to index the designable diagonal entries at each iteration, and initialize this set as containing all of the designable entries in the subset design problem (this set will be iteratively reduced). Denote the dimension of the designable set at iteration k as m_{k-1}. Then apply the following iteration:
2) First, re-arrange the rows/columns of $\hat{A}-K$ via permutation so that the designable entries are the first m_{k-1} entries, as specified in the following form:
$$\begin{bmatrix} \hat{A}_{11} - D^{(k)} & \hat{A}_{12} \\ \hat{A}_{21} & \hat{A}_{22} \end{bmatrix}.$$
Here $D^{(k)}$ is the $m_{k-1} \times m_{k-1}$ diagonal matrix containing the designable gains at the kth iteration (i.e., $D^{(k)}$ is designable part of matrix K). In the following steps, the unconstrained optimum gains $D^{(k)}$ will be found and then used for iteration.
3) Solve the following equation for $\lambda_{\max}$:
$$\vec{1}' A_r(\lambda_{\max})\vec{1} - \Gamma - \lambda_{\max}\vec{1}'\vec{1} = 0 \tag{7}$$
where $A_r(\lambda_{\max}) = \hat{A}_{11} + \hat{A}_{12}(\lambda_{\max} I - \hat{A}_{22})^{-1}\hat{A}_{21}$, $\vec{1}$ is the all ones vector of the appropriate dimension.
4) Calculate the optimal unconstrained design for this designable subset, as $D^{(k)} = \mathrm{diag}(A_r(\lambda_{\max})\vec{1} - \lambda_{\max}\vec{1})$ and check that $D^{(k)}_i > 0$ for all $i = 1, \ldots, m_{k-1}$. Let us denote the set of indices i such that $D^{(k)}_i < 0$ as $\mathcal{I}_-$, and the set of indices i such that $D^{(k)}_i = 0$ as $\mathcal{I}_0$. Also, we update $\mathcal{I}_D(k)$ to contain the indices i such that $D^{(k)}_i > 0$ for $i = 1, \ldots, m_{k-1}$. If the set $\mathcal{I}_-$ is empty then
$$K^* = \begin{bmatrix} D^{(k)} & 0 \\ 0 & 0 \end{bmatrix}.$$
5) If the set $\mathcal{I}_-$ is non-empty, we set $D^{(k)}_i = 0$ for all $i \in \mathcal{I}_-$. We stress that the set of designable gains D_i after the iteration only include those whose indices are in the updated $\mathcal{I}_D$, which is defined to have m_k elements. We also set $k = k+1$ and go to step 2.

Proof: The result is proved by showing that a sequence of increasingly-constrained optimization problems are solved by the iterative algorithm given in the theorem statement. Specifically, it is shown that, at each iteration, the original design problem with positivity constraints enforced on the nodes that are not in $\mathcal{I}_D(k)$ is solved. Without loss of generality, let A be a symmetric matrix. Clearly, step 3 provides an optimal solution for the unconstrained subset design (Theorem 2), for the specified designable set. It is clear that if $\mathcal{I}_-$ is empty after this step, the solution obtained in step 3 is the global optimal solution for the constrained case, i.e., $D_i \geq 0$. If not, further steps in the iteration need to be evaluated.

At each subsequent step of the iteration, we prove that the further-constrained optimization problem defined above (with all entries not in $\mathcal{I}_D(k)$ constrained to be nonnegative) is solved. This proof itself requires an inductive argument. In particular, we first show that setting any single gain D_i such that $i \in \mathcal{I}_-$ to zero and repeating step 3 will provide a global optimal with this entry further constrained to be positive, i.e., a solution that satisfies Lemma 4 (and hence KKT coditions are also satisfied). Via an induction argument, we then argue that we can set all the gains D_i such that $i \in \mathcal{I}_-$ to zero at once so as to get a solution to the further constrained design problem (with all entries not in $\mathcal{I}_D(k)$ constrained to be nonnegative), because of the directions in which the other gains move when

one gain is set to zero and the gains are reoptimized. Here is the argument.

Consider that $\mathcal{I}_-$ is non-empty. Specifically, let us say that $\mathcal{I}_-$ has u elements. Let us also consider the optimal solution when these gains are set to zero. We will prove that in this case Lemma 4 is satisfied. In other words, we will prove that $v^2_{\max,i} < v^2_{\max,j}$ for $i \in \mathcal{I}_- \cup \mathcal{I}_0$ and $j \in \mathcal{I}_D$, where $v_{\max}$ is the eigenvector associated with the dominant eigenvalue of the new matrix $A - K^{(k+1)}$, and hence this optimal solution is also global for the further-constrained optimization problem. Further, the designable gains for this new optimal are greater or equal to the previous ones. We will prove this using induction. We take for instance the kth step of the algorithm.

Basis Step: Let $Z_0 = A - K^{(k)}$ be the matrix that produces the unconstrained global optimal at the k-iteration and λ_0 be its dominant eigenvalue. Without loss of generality, assume that the gains $m_{k-1} - u + 1, \ldots, m_{k-1}$ are in $\mathcal{I}_-$. We first increase the gain $D_{m_{k-1}}$ to zero while optimizing the other gains (without any constraint).

Let $Z_1 = Z_0 + \Delta$, where $\Delta = \text{diag}(\Delta D_1, \ldots, \Delta D_{m_{k-1}-1}, -\alpha, 0, \ldots, 0)$ and $\sum_{i=1}^{m_{k-1}} \Delta D_i = \alpha$. We would like to prove: 1) the m^{th}_{k-1} eigenvector component (associated with $D_{m_{k-1}} < 0$) is less than the entries 1 to m_{k-1} and 2) the gains $\Delta D_i > 0$ for $i = 1, \ldots, m_{k-1} - 1$.

Since Z_0 is symmetric, we can write the dominant eigenvalue as the maximization of a function, via the Courant-Fisher Theorem [46]. Specifically, $\lambda_0 = \max_v v' Z_0 v = v_0' Z_0 v_0$, where $v_0' = [\vec{1}'_{m_{k-1}} \;\; \bar{v}_0']$ is an eigenvector of Z_0. Similarly, $\lambda_1 = \max_v v' Z_1 v = v_1' Z_1 v_1$, where $v_1' = [d\vec{1}'_{m_{k-1}-1} \;\; c \;\; \bar{v}']$ is an eigenvector of Z_1. Then $\lambda_1 = v_1'(Z_0 + \Delta)v_1 = v_1' Z_0 v_1 + \alpha d^2 - \alpha c^2$. Since v_0 maximizes the function $v' Z_0 v$, $\lambda_1 < \lambda_0 + \alpha d^2 - \alpha c^2$ and at the same time, $\lambda_0 < \lambda_1$. Hence, $d^2 > c^2$, or equivalently $d > c$, i.e., Lemma 4 is satisfied.

Let us prove now that the entries ΔD_i for $i = 1, \ldots, m_{k-1} - 1$ are positive. Because $d > c$, we can write the eigenvector corresponding to the maximum eigenvalue of Z_1 as $v_1' = v_0' + [d_1 \vec{1}'_{m_{k-1}-1} \;\; c_1 \;\; q_1']$, where $c_1 < 0 < d_1$ and q_1 is a $n - m_{k-1}$ vector. Let Z_1 have the following form:

$$Z_1 = \begin{bmatrix} Z_{11} + \Delta_1 & Z_{12} & Z_{13} \\ Z_{21} & Z_{22} - \alpha & Z_{23} \\ Z_{31} & Z_{32} & Z_{33} \end{bmatrix}$$

where $\Delta_1 = \text{diag}(\Delta D_1, \ldots, \Delta D_{m_{k-1}-1})$. Using the eigenvector equation and some algebra, we obtain that

$$\begin{aligned}(1 + d_1)\Delta_1 \vec{1}_{m_{k-1}-1} &= (\lambda_1 - \lambda_0)\Big(\vec{1}_{m_{k-1}-1} + Z_{13}Q_4\bar{v}_0 + Z_{13}Q_3 + Z_{12}Q_1 \\ &\quad + Z_{12}Q_2\bar{v}_0\Big) + \alpha(Z_{13}Q_3 + Z_{12}Q_1) \\ &\quad + d_1\left(\lambda_1 I - Z_{11} - \begin{bmatrix} Z_{12} & Z_{13}\end{bmatrix} M_3^{-1} \begin{bmatrix} Z_{21} \\ Z_{31}\end{bmatrix}\right)\vec{1}_{m_{k-1}-1} \end{aligned} \tag{8}$$

where

$$M_3 = \lambda_1 I - \begin{bmatrix} Z_{22} - \alpha & Z_{23} \\ Z_{32} & Z_{33}\end{bmatrix}$$

is a M-matrix. Further, the inverse of M_3 is entry wise non negative and has the following form:

$$M_3^{-1} = \begin{bmatrix} Q_1 & Q_2 \\ Q_3 & Q_4 \end{bmatrix}$$

where $Q_1 = P_2(1 + Z_{23}Q_4 Z_{32} P_2)$, $Q_2 = P_2 Z_{23} Q_4$, $Q_3 = Q_4 Z_{32} P_2$, $Q_4 = (\lambda_1 I - Z_{33} - Z_{32} P_2 Z_{23})^{-1}$ and $P_2 = (\lambda_1 - (Z_{22} - \alpha))^{-1}$.

The last expression in equation (8) contains a Schur complement of the M-matrix $\lambda_1 I - Z_1$. In particular, from the eigenvector equation $(\lambda_1 I - Z_1)v_1 = 0$, we get (with some algebra) that

$$\begin{aligned}(\lambda_1 I - (Z_{11} + \Delta_1))\,\vec{1}_{m_{k-1}-1} \\ = \left(\begin{bmatrix} Z_{12} & Z_{13}\end{bmatrix} M_3^{-1} \begin{bmatrix} Z_{21} \\ Z_{31}\end{bmatrix}\right)\vec{1}_{m_{k-1}-1}.\end{aligned}$$

Replacing the above equation in equation (8), we obtain that:

$$\begin{aligned}\Delta_1 \vec{1}_{m_{k-1}-1} &= (\lambda_1 - \lambda_0)\Big(\vec{1}_{m_{k-1}-1} + Z_{13}Q_4\bar{v}_0 + Z_{12}Q_2\bar{v}_0 \\ &\quad + Z_{13}Q_3 + Z_{12}Q_1\Big) \\ &\quad + \alpha(Z_{13}Q_3 + Z_{12}Q_1). \end{aligned} \tag{9}$$

We note that Q_i for $i = 1, \ldots, 4$ as well as the off-diagonal blocks of Z_1 and $\bar{v}_0$ are entry-wise nonnegative vector/matrices. Further, $\lambda_1 > \lambda_0$ and $\alpha > 0$. In other words, the gains ΔD_i for $i = 1, \ldots, m_{k-1} - 1$ are strictly positive.

Inductive Step: Suppose we have re-set l of the entries $D_i > 0$ to zero ($l < u$). Let us also suppose that $v_{\max,j} < v_{\max,i}$ for $j = m_{k-1} - l + 1, m_{k-1} - l + 2, \ldots, m_{k-1}$ and $i = 1, \ldots, m_{k-1} - l$ (global optimal condition). Additionally, suppose that the gains D_i for the remaining elements of $\mathcal{I}_-$ are at most 0. We would like to prove that after setting another gain $D_i < 0$ to zero, we still have the appropriate eigenvector pattern (Lemma 4) and condition on the gains.

Without loss of generality, consider setting the gain $D_{m_{k-1}-l}$ to zero. Let us denote the dominant eigenvalue in this step λ_l, and the one before setting this gain to zero as λ_{l-1}. Applying again Courant-Fisher Theorem, $\lambda_{l-1} = \max_v v'(A - K)v$ and $\lambda_l = \max_v v'(A - K + \Delta)v$, where $\Delta = \text{diag}(\Delta D_1, \ldots, \Delta D_{m-l-1}, -\beta, 0, \ldots, 0)$ is the matrix that set D_{m-l} to zero and change the other $m - l - 1$ gains accordingly. We note that $\sum_{i=1}^{m_{k-1}-l-1} \Delta D_i = \beta$. Repeating similar argument given in the basis step, we can show that: $v_{\max,j} < v_{\max,i}$ for $j = m_{k-1} - l, m_{k-1} - l + 1, \ldots, m_{k-1}$ and $i = 1, \ldots, m_{k-1} - l - 1$.

The proof that the remaining gains D_i decrease upon setting this gain to zero, and hence are at most zero if $i \in \mathcal{I}_-$, follows a similar argument to that given in the basis step. Further, since setting any gain D_i such that $i \in \mathcal{I}_-$ will only decrease the remaining designable gains, we can also set the gains D_i for $i \in \mathcal{I}_0$ to zero since they will violate the constraint. Thus, via an inductive argument, we have shown that the further-constrained optimization problem is solved (i.e., the KKT condition for this problem is achieved), via one step of the iteration in the theorem statement.

It remains to prove that the iteration given in the theorem statement eventually yields the solution to the fully-constrained design problem. To see this, notice that two outcomes are possible after each iteration stage. First, the designed unconstrained gains may be nonnegative, in which case the fully-constrained optimization problem has in fact been solved since the constraints are met. Alternately, some of these gains may be negative. In this case, the set $\mathcal{I}_D$ strictly decreases in cardinality in the following iteration. Thus, either the optimal solution is obtained, or within $m-1$ iterations only a single entry remains in $\mathcal{I}_D$; since the total resource allowed is assumed positive, the optimal design at this stage must be positive, and the fully-constrained optimization problem is solved. ■

The last theorem describes a finite iteration (requiring at most m stages) for computing the optimal control resource allocation K^* for the constrained subset-design problem, when the state matrix A defined in the problem formulation is diagonally-symmetrizable.

Remark 6: We have been able to modify this method to address the case when the topology matrix is an arbitrary irreducible matrix. This new algorithm replaces step 3 in the constrained subset design algorithm (Theorem 6) with the procedure in Theorem 2, which uses line-sum symmetrization to find the unconstrained optimal solution $\bar{K}$. In practice, this *generic* algorithm has been able to always find the optimal solution, but we have not yet been able to prove that it works generally. However, a special case can be proved: when the unconstrained optimum has only one gain $D_i < 0$ (one gain violating the constraint), then the procedure is guaranteed to find a global optimal, i.e., a solution that satisfies Lemma 4 [55]. The proof of this result builds on an important theorem on the diagonal scaling that yield line-sum symmetry for square nonnegative matrices [36]. We have excluded the details in interest of the space. We stress that the design obtained through this algorithm can be checked for optimality via Lemma 4.

Remark 7: It is worth commenting about computational complexity of the algorithm. Each stage of the algorithm requires: 1) solving a polynomial equation equation for $\lambda_{\max}$, for which many algorithms are available that are polynomial-time in dimension; and 2) computing the optimal gains thereof, which is at most quadratic in the problem dimension. The number of algorithm stages is at most $m_0 - 1$, where m_0 is the number of designable entries in the original design problem. Thus, the algorithm allows a fast (polynomial-time implementation). The special problem structure allows us to search over a nested sequence of subsets of the original designable set, rather than considering all possible subsets, and still guarantee optimality.

Remark 8: Geometric programming approaches can also permit fast solution to the formulated spread-control problem [2] via convex optimization. In comparison, our approach yields a specialized recursive algorithm that is guaranteed to find exactly the optimum within a certain number of iterations. Also, importantly, the approach gives interesting structural insights into the optimum, which can be applied even when the network is not precisely known, the problem dimension is too large, or other reasons. On the other hand, our method does not directly leverage the considerable machinery available for convex nonlinear optimization that geometric programming enables.

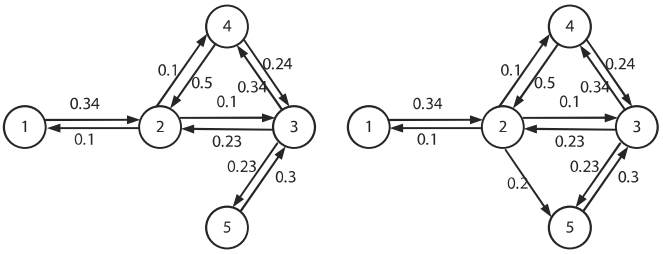

Fig. 1. Graph associated with the spread dynamics for example 1 (left) and example 2 (right).

TABLE I
OPTIMAL RESOURCE ALLOCATION FOR EXAMPLE 1

	1^{st} Iteration, $\lambda_{\max} = -0.0623$		2^{nd} Iteration, $\lambda_{\max} = -0.0551$	
i	D_i	$w_{\max,i} v_{\max,i}$	D_i	$w_{\max,i} v_{\max,i}$
1	0.1467	0.1549	0.0872	0.1801
2	-0.1448	0.1549	0	0.0924
3	1.0142	0.1549	0.9128	0.1801
4	-0.0161	0.1549	0	0.1409

In terms of numerical computation, examples indicate that our approach is fast compared to the geometric-programming approaches when the optimal design is at the boundary of the design space (specifically, some locations receive no resources at the optimum).

C. Examples

Two examples are given to illustrate the subset-design algorithm for the diagonally symmetrizable and irreducible non-diagonally symmetrizable cases.

Example 1 (Diagonally-Symmetrizable Matrix Topology): Let the 5×5 topology matrix of the spread dynamics be

$$A = \begin{bmatrix} -0.1 & 0.1 & & & \\ 0.34 & -0.62 & 0.23 & 0.5 & \\ & 0.1 & 0.1 & 0.24 & 0.3 \\ & 0.1 & 0.34 & -0.44 & \\ & & 0.23 & & -0.23 \end{bmatrix}. \quad (10)$$

The graph of the spread process is shown in Fig. 1. Consider the case that the total resource amount is $\Gamma = 1$ and the resources can only be applied to the nodes $i = 1, \ldots, 4$. In this case, the minimum amount of resources needed to stop the spread (achieve stability), considering no constraints, is $\Gamma_{\min} = 0.751$. Here, we consider the optimal subset design under constraint. In the following table, we summarize the results after each iteration of the algorithm in Theorem 6. The globally optimal constrained design is found after the second iteration, since a feasible solution has been found (and the algorithm is known to converge to the global optimum). The optimizing design, corresponding $\lambda_{\max}$, and left and right eigenvectors of the closed-loop, are shown in Table I. The constrained design achieves stability.

At the optimum, the participation factors for nodes 1 and 3 are greater those of nodes 2, 4, and 5, as expected from Lemma 3 and 4.

Example 2 (Non-Symmetrizable Irreducible Matrix Topology): For this example, we use a modification of the algorithm presented in Theorem 6. In this case, we substitute step 3 for the computation in Lemma 2, as described above.

TABLE II
OPTIMAL RESOURCE ALLOCATION FOR EXAMPLE 2

	1^{st} Iteration, $\lambda_{\max} = -0.0722$		2^{nd} Iteration, $\lambda_{\max} = -0.0713$	
i	D_i	$w_{\max,i} v_{\max,i}$	D_i	$w_{\max,i} v_{\max,i}$
1	0.1566	0.2014	0.1392	0.2148
2	-0.04435	0.2014	0	0.1781
3	0.8989	0.2014	0.8608	0.2148
4	-0.0112	0.2014	0	0.1938

Specifically, let us consider a spread dynamics with the following topology matrix:

$$A = \begin{bmatrix} -0.1 & 0.1 & & & \\ 0.34 & -0.62 & 0.23 & 0.5 & \\ & 0.1 & 0.1 & 0.24 & 0.3 \\ & 0.1 & 0.34 & -0.44 & \\ & 0.2 & 0.23 & & -0.43 \end{bmatrix}. \quad (11)$$

The network graph is shown in Fig. 1. For this example, we again consider $\Gamma = 1$, and consider the case that resources can only be applied at nodes $i = 1, \ldots, 4$. In this case, the minimum resource level needed for a stabilizing unconstrained design is $\Gamma_{\min} = 0.8574$.

The optimal constrained design was determined for this example. In Table II we summarize the results after each iteration of the algorithm, and present the design details once the optimum is reached:

We note that the resource allocation provided by this modified algorithm is provably optimal since the eigenvectors satisfy Lemma 4.

IV. EXPLORATORY APPLICATION TO DISEASE SPREAD CONTROL

Infectious disease outbreaks incur significant cost in terms of human morbidity/mortality as well as loss of productivity, and necessitate nimble management policies to limit impact. One important aspect of disease management is to decide where, and to what extent, surveillance and control capabilities should be put in place across a wide area. For example, for the Ebola outbreak of Winter 2015, there was a need to select certain ports-of-entry into the United States for further surveillance. Because these questions are related to the spatial patterns of movement and spread, it is natural to use dynamical network models to support policy-making. Indeed, a number of dynamical-network models have been proposed. These are of various resolutions, ranging from detailed simulation models representing individuals to delay differential equations for prevalence at the group level (e.g., [14]–[16].) While some models have been used to evaluate control policies in practice, to our knowledge systematic methods for designing surveillance and control have not been implemented in the field.

The research described here can potentially enable systematic design of management policies using network spread models. However, possible application to disease management must be considered with caution. In reality, disease spread processes and their management are much more sophisticated than the linearized models described here, involving nonlinearities, intrinstic stochastics, delays, time variation, highly specialized controls, environmental impacts, etc. Linearizations are used for evaluation of spread patterns and control policies in the epidemiology literature. This is primarily because stability vs. instability of the linear operator are known to correspond to elimination vs. persistence of the spread, respectively (see the wide literature on the next-generation operator). However, it unclear how effectively the linear models can capture transients of real-world disease spread processes, and hence whether the optimal solutions obtained here are in fact effective. Our belief is that the optimal design usually will not translate directly to quantitative design of policy details—there are too many unmodeled factors and operational specifics. Rather, we argue that the optimization framework can give wide-area insight into where, and to what extent, limited and costly surveillance/control resource should be placed. These broad insights can then guide policy-makers in selecting particular control actions and deciding their scope, at both operational and planning horizons. More detailed simulation models would then be used to evaluate designed policies.

In order for the optimization results to be effective for policy design, it is crucial that design insights obtained from the linearization be robust to the nonlinearities, stochastics, and other factors that are present in reality. We pursue two examples in order to evaluate, in a preliminary way, whether the obtained designs are sufficiently robust to support effective policy design. First, for a constructed 100-node network, the obtained design is simulated using a stochastic nonlinear multi-group model for spread, and shown to perform well. Second, a crude model for the spread of Chikungunya in the Latin Caribbean, which matches historical data, is used to understand in detail the policy structures suggested by the optimization. We also note that, in our previous work, network-design approaches have been used to support policy-development for management of zoonotic diseases [56], however the details are outside the scope of this paper.

The application examples pursued here focus on disease-spread control, but it is important to stress that similar subset-design problems arise in other application areas. In fact, the nonlinear multi-group model considered in the first example below also naturally captures spatial population dynamics and some biochemical reaction processes occurring at the cellular level. The subset design problem is also of significant interest in these contexts. Specifically, it allows study of whether control of pest animals at a few locations can be used to limit overall populations. Meanwhile, in a biochemical-reaction context, the subset design problem is relevant to designing targeted interventions to reaction processes, which may support e.g., design of medicines. Similar subset design problems also arise in control of diffusive processes, such as thermal management of a building, or monitoring and mitigation of pollution spread. Further details are omitted in the interest of space.

A. Design Simulation: 100-Node Example

The proposed design is tested using a stochastic nonlinear multi-group model for spread. Specifically, for the 100-node example shown in Fig. 2, the stochastic nonlinear susceptible-infectious-susceptible (SIS) multi-group model is simulated. We recall that the multi-group SIS model tracks infection counts or prevalence levels in multiple subpopulations. New

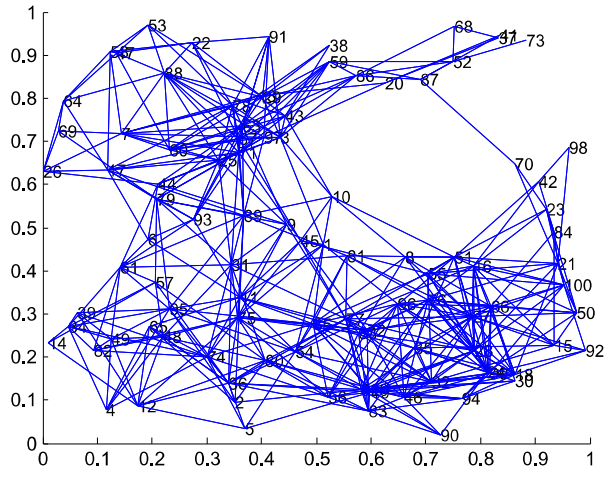

Fig. 2. Network Graph associated with the stochastic nonlinear susceptible-infectious-susceptible (SIS) multi-group model.

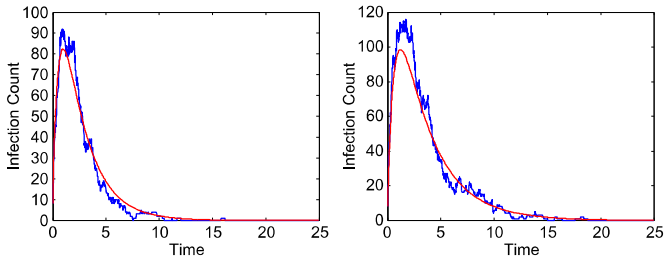

Fig. 3. Infection count variation at node 79 for network-wide optimal design (left) and uniform design (right). The solid lines consider the intrinsic variability in new infection and recovery rates. The dashed lines consider average new infection and recovery rates.

infections arise due to interactions among the subpopulations, with the rate scaling with the product of the infected population in the source group and uninfected population in the affected group—hence a quadratic nonlinearity is introduced. Linear local recovery processes are also captured, and controls are viewed as increasing recovery rates. When infection counts are small, intrinsic variability in new infections and recovery become important: such variability can be captured using Poisson-process models for new infections and recovery (whose rates are specified as above). Here, we have implemented this stochastic nonlinear network model for the 100-node network shown in Fig. 2. In the interest of space, the nonlinear model is not presented in detail here, see e.g., [15], [16] for similar dynamical models. We note that initial infection penetrations of 10%–50% have been assumed, which incur significant nonlinear effects.

For the nonlinear model described above, the linearization about the equilibrium at the origin has been found. This linearization is then used to design resource allocations at subsets of the network nodes. Broadly, two types of comparisons are undertaken here. First, the optimal resource allocation has been compared with a uniform resource allocation (with the same total resource level). Also, the optimal designs for different designable sets are compared.

In this example, we have considered the network-wide design where all 100 nodes are amenable to control. We have also considered four sets of *designable nodes* with 20 nodes, selected as

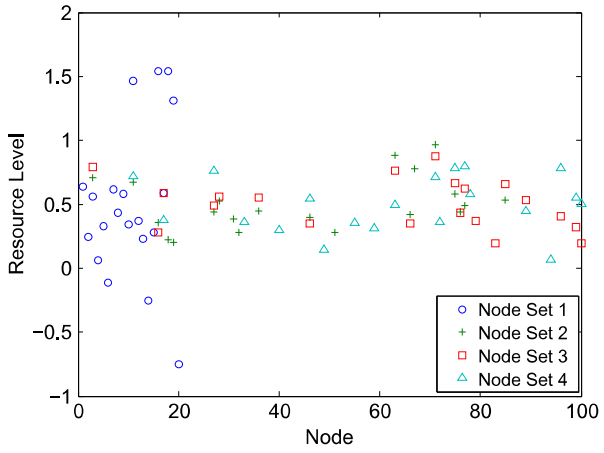

Fig. 4. Resource level distribution.

TABLE III
OPTIMAL MAXIMUM EIGENVALUE

Designable Nodes	Optimal Eigenvalue $\bar{\lambda}_{\max}$
All nodes	-0.4172
Node Set 1	-0.2581
Nose Set 2	-0.3233
Node Set 3	-0.3219
Node Set 4	-0.3094

follows: 1) Node Set 1 has 20 nodes chosen randomly, 2) Node Set 2 contains the 20 nodes with most number of connections, 3) Node Set 3 contains the 20 nodes with the highest total (local and inter) transmission rate, and 4) Node set 4 has 20 nodes with the highest local transmission rate.

Fig. 3 compares the performance of the network-wide optimal design and uniform design, in the context of the nonlinear stochastic model: a sample trajectory and the mean response are compared. The optimal design shows significantly better performance compared to the uniform design, including a reduced peak count and faster settling time. This is in correspondence with the performance of the linearized model, for which the dominant eigenvalue is -0.41 for the optimized design as compared to -0.28 for the uniform allocation. The example demonstrates that the optimal design for the linearized models also translates to an effective design for the linear model.

The optimized designs for the four subset-design cases considered here are shown in Fig. 4, while the optimized maximum eigenvalue $\bar{\lambda}_{\max}$ is shown in Table III. For each designable node set, the optimal resource allocation has significant variation across the network: some nodes are allocated significant resources, while others are not. The resource distribution also varies significantly among the Designable Node Sets, suggesting that different design strategies are needed for each case. For each designable node set, the dominant eigenvalue at the optimum shows a small but marked improvement compared to a uniform distribution (5%–10%). Fig. 5 shows the locations and the resource levels for the designable nodes corresponding to Node Set 1 (20 most connected nodes). The darkest squares in the Fig. 5 indicate the locations where more resources should be placed, while the whitest squares indicate the location where less resources should be applied, or resources should be taken.

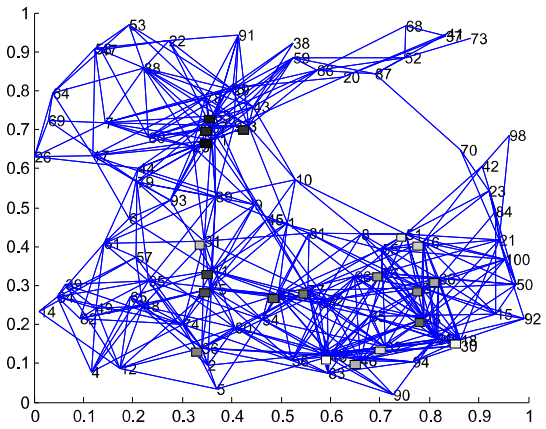

Fig. 5. Location and resource level distrbution for node set 2.

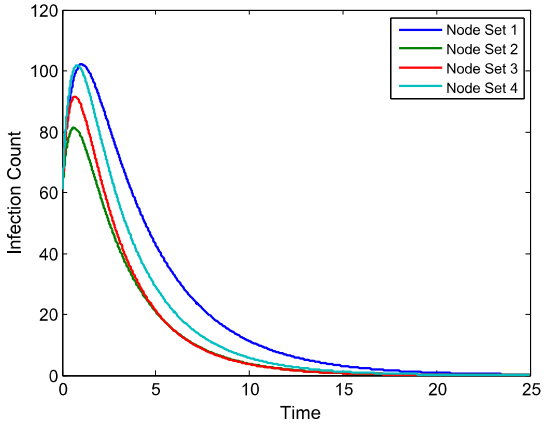

Fig. 6. Average infection count (state evolution) at node 27.

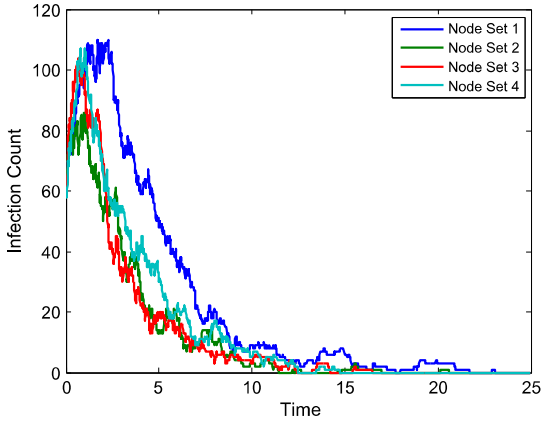

Fig. 7. Infection Count (state evolution) considering intrinsic variability in new infections and recovery at node 27.

Figs. 6 and 7 show the state evolution (infection count) at node 27 for each designable node set, as given by the nonlinear stochastic model. In particular, Fig. 6 presents the average response over many sample trajectories. On the other hand, Fig. 7 captures a single sample trajectory. Importantly, these figures show that the controlled nonlinear model displays similar performance characteristics as predicted by the optimally-controlled linearized model, in terms of the relative effectiveness of different designable node sets and the asymptotic performance of the design.

An important further direction is to evaluate the performance of the design in a detailed, individual-level susceptible-infectious-susceptible model of spread (e.g., [8]). Testing the design in a detailed model of this sort requires modeling the distribution of the resources allocated to a group to the individuals within the group. A natural approach is to probabilistically distribute the available curing resources among these individuals, to reflect regional control policies (e.g., public-health reminders or discounted treatments) that act on a random selection of individuals within a region. Noting the connection between individual-level models and diffusion models [57], we expect the performance comparisons to remain apt, however we leave detailed simulation to future work.

B. Chikungunya Spread in the Caribbean Countries

The Chikungunya virus is a vector-borne disease that is endemic to parts of Africa and Southeast Asia. Since late 2013, a significant epidemic of the disease has been underway in the Caribbean and parts of Central, South, and North America [58].[1] Chikungunga is primarily transmitted among human hosts by mosquitos, which are carriers of the virus. The virus causes fever, rash, long-duration joint pain, etc. in humans, and is exacting a significant economic and social cost in impacted countries. No vaccine or specific treatment for the virus is known. Thus, effective forecasting and strategic management of the disease (via control of mosquito populations, and reduction of contact) is crucial [58].

This example is concerned with how the Chikungunya virus could have been managed in its initial stages in the Latin Caribbean, with a particular focus on understanding how inter-country transmission rates would impact management policies. For the example developed here, we have developed a very simplified model of the spread of Chikungunya virus in six countries/territories of the Latin Caribbean. The model tracks total infection counts in the initial stages of the virus in these six countries, using a linearized multi-group model of the form (1). A simplified procedure has been used to parameterize the model. In particular, weekly growth rates in disease prevalence have been used to determine intra-country spread rates. Specifically, the diagonal entries of A have been chosen to achieve best fits of the exponential growth of the disease in each country in the initial stages of the disease. Meanwhile, low rates have been assumed for inter-country transmission based on passenger-transport and shipping patterns. Specifically, the off-diagonal entries in A have been selected as nonzero or zero depending on whether or not standard ferry/transport routes are present between the countries. The exact magnitudes of these off-diagonal entries are difficult to ascertain from data, but were guessed based on counts of new infections originating from other countries in the data record. Because of the difficulty in estimating inter-country transmission rates, we also study here how resource allocation patterns change if the inter-country rates are scaled. Data for parameterization were drawn from

[1] [Online]. Available: http://www.cdc.gov/chikungunya/

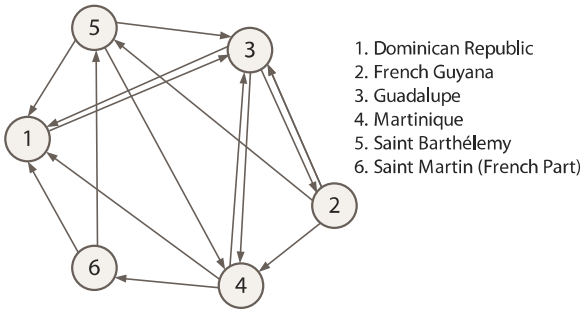

Fig. 8. Graph associated with the spread dynamics for Chikungunya spread example.

TABLE IV
OPTIMAL RESOURCE ALLOCATION FOR CHIKUNGUNYA SPREAD EXAMPLE

	Control everywhere $\lambda^*_{\max} = -0.1501$		Control at $1,3,4,5$ (1^{st} It.) $\lambda^*_{\max} = 0.0285$	
i	D_i	$w_{\max,i}v_{\max,i}$	D_i	$w_{\max,i}v_{\max,i}$
1	0.1831	0.1667	0.0099	1.09×10^{-7}
2	0.1776	0.1667	0	0.9999
3	0.1713	0.1667	0.8349	1.09×10^{-7}
4	0.1636	0.1667	0.1252	1.09×10^{-7}
5	0.1536	0.1667	0.0299	1.09×10^{-7}
6	0.1507	0.1667	0	4.07×10^{-11}

the Pan-American Health Organization.[2] The state matrix of the nominal linearized model was found to be

$$A = 10^{-3} \times \begin{bmatrix} 34.86 & & 0.095 & 0.709 & 0.0781 & 0.0336 \\ & 28.51 & 0.718 & & & \\ 0.336 & 0.0861 & 18.78 & 3.45 & 0.233 & \\ & 0.634 & 4.124 & 10.04 & 7.101 & \\ & 0.928 & & & 3.574 & 0.658 \\ & & & 0.802 & & 1.781 \end{bmatrix}.$$

Fig. 8 shows the network graph associated with the spread dynamics and specifies the countries/territories involved.

We consider that a total resource amount of $\Gamma = 1$ can be applied, and pursue the optimal constrained design. We note here that the control resources abstractly represent local control capabilities (e.g., mosquito-reduction efforts, quarantine, public-service announcements) that reduce spread rates within a country. Our main interest is to understand which regions (countries) in the network to target with control resources using this abstract formulation, rather than to compare detailed control strategies. We analyze two cases: 1) the case that resources can be applied everywhere in the network, and 2) the case that resources can be applied only in nodes $1,3,4,5$ of the network graph shown in Fig. 8. The optimal resource allocations are summarized in the following table.

Table IV shows that applying resources to all the nodes in the network is more effective than just in a subset of nodes, in that a more negative minimum dominant eigenvalue is achieved. In the network-wide design, the countries that receive more resources are the ones with high local (α_{ii}) and inter-country (α_{ij}) transmission rate in total. Clearly, this total transmission rate indicates high local and network impact. In the subset design, we observe an interesting phenomenon: the country with highest local transmission rate is not the one that receives more resources. Instead, the resources are focused in countries with higher inter-country transmission rates because of the influence of these nodes in the network as expressed in the participation factors. This reflects that the optimal resource distribution accounts for both local and network-wide impact. Because the inter-country transmission rates are so small, however, (i.e., α_{ij} are much smaller compared with the local rates), even this re-distributed resource profile has limited ability to stop spread across the network.

TABLE V
OPTIMAL RESOURCE ALLOCATION: SUBSET-DESIGN

	Control everywhere $\lambda^*_{\max} = 0.1266$		Control at $1,3,4,5$ (2^{nd} It.) $\lambda^*_{\max} = 0.1404$	
i	D_i	$w_{\max,i}v_{\max,i}$	D_i	$w_{\max,i}v_{\max,i}$
1	-0.0333	0.1667	0	0.0440
2	0.0476	0.1667	0	0.3547
3	0.3830	0.1667	0.0672	0.1336
4	0.4801	0.1667	0.3787	0.1336
5	0.1193	0.1667	0.5541	0.1336
6	0.0033	0.1667	0	0.2004

To get a better understanding of how the inter-country transmission rates influence resource allocation, we consider scaling up the inter-country transmission rates by 10^3 compared to the original fitted model. Even with the up-scaling, the inter-country transmission rates remain small compared to the local transmission rates (by a factor of about 50–100). The optimal resource allocation is summarized in the following table.

In the network-wide design, the country that receives more resources is again the one that has both more local and inter-country influence. Interestingly, when the inter-country transmission rates are stronger, and the local rates are the same, the country with largest local transmission rate does not receive more resources, because of the low impact that it has in the network. In the subset design, much stronger inter-country rates imply a more effective network-wide impact of the resource distribution, as observed in Table V. The allocation of resources is also modified somewhat compared to the low-intercountry-rate case, because different countries now have stronger network-wide impact. When the optimal subset design policy is put in place, the spread is seen to evolve in the following way. The spread dies out quickly in the locations that have been allocated resources. In the remaining (undesignable) locations, the spread evolves more slowly but eventually dies off because the local spread rate is small, and amplification via the rest of the network no longer occurs. An important implication is that spreads such as these can be controlled even if some communities are uncooperative or unable to provide control resources, albeit at a higher costs to the other players.

While the model and design here are simplistic, they reflect the operational recognition that controls should serve as barriers against spread to undesignable subpopulations/regions. As a further point, we notice that the designs bring forth interesting questions regarding fairness of the resource distribution: the optimal solution presented here is unfair in the sense this different regions receive different resource levels, but fair in the sense of equalization spread impact of designable regions (see also [6]).

[2][Online]. Available: http://www.paho.org/hq

REFERENCES

[1] Y. Wan, S. Roy, and A. Saberi, "Designing spatially heterogeneous strategies for control of virus spread," *Syst. Bio., IET*, vol. 2, no. 4, pp. 184–201, 2008.

[2] V. M. Preciado, M. Zargham, C. Enyioha, A. Jadbabaie, and G. Pappas, "Optimal Resource Allocation for Network Protection: A Geometric Programming Approach," *arXiv preprint arXiv:1309.6270*, 2013.

[3] Y. Ying and P. Li, "Distance metric learning with eigenvalue optimization," *J. Mach. Learn. Res.*, vol. 13, pp. 1–26, 2012.

[4] X. Zhai, L. Zheng, J. Wang, and C. W. Tan, "Optimization algorithms for epidemic evolution in broadcast networks," in *Proc. WCNC*, 2013, pp. 1540–1545.

[5] E. Ramirez-Llanos and S. Martínez, "A distributed algorithm for virus spread minimization," in *Proc. IEEE Amer. Control Conf.*, 2014, pp. 184–189.

[6] A. Vijayshankar and S. Roy, "Cost of fairness in disease spread control," in *Proc. IEEE 51st Annu. CDC*, 2012, pp. 4930–4935.

[7] X. Wang, S. Roy, and Y. Wan, "Using deliberate-delay decentralized controllers to stop spread dynamics in canonical network models," in *Proc. IEEE 1st ICCABS*, 2011, pp. 178–183.

[8] K. Drakopoulos, A. Ozdaglar, and J. N. Tsitsiklis, "An efficient curing policy for epidemics on graphs," *IEEE Trans. Netw. Sci. Eng.*, vol. 1, no. 2, pp. 67–75, 2014.

[9] O. Diekmann and J. Heesterbeek, *Mathematical Epidemiology of Infectious Diseases*. Chichester, U.K.: Wiley, 2000, vol. 146.

[10] L. Xiao and S. Boyd, "Fast linear iterations for distributed averaging," *Syst. Control Lett.*, vol. 53, no. 1, pp. 65–78, 2004.

[11] S. Roy, T. F. McElwain, and Y. Wan, "A network control theory approach to modeling and optimal control of zoonoses: Case study of brucellosis transmission in sub-saharan africa," *PLoS Neglected Tropical Diseases*, vol. 5, no. 10, 2011, Art. no. e1259.

[12] N. Suthar, S. Roy, D. R. Call, T. E. Besser, and M. A. Davis, "An individual-based model of transmission of resistant bacteria in a veterinary teaching hospital," *PloS One*, vol. 9, no. 6, 2014, Art. no. e98589.

[13] W. H. Organization *et al.*, "Who Report on Global Surveillance of Epidemic-Prone Infectious Diseases," 2000.

[14] R. M. Anderson, R. M. May, and B. Anderson, *Infectious Diseases of Humans: Dynamics and Control*. Hoboken, NJ, USA: Wiley, 1992, vol. 28.

[15] A. Lajmainovitch and J. Yorke, "A deterministic model for gonorrhea in a nonhomogeneous population," *Math. Biosci.*, vol. 28, no. 1, p. 221, 1978.

[16] C. Ji, D. Jiang, and N. Shi, "Multigroup sir epidemic model with stochastic perturbation," *Phys. A: Statist. Mech. Appl.*, vol. 390, no. 10, pp. 1747–1762, 2011.

[17] E. J. Davison, " Decentralized stabilization and regulation in large multivariable systems," in *Directions in Large-Scale Systems*. New York, NY, USA: Springer-Verlag, 1976, pp. 303–323.

[18] C. Leondes, *Control and Dynamic Systems: Decentralized/Distributed Control and Dynamic Systems. Part 3*. Orlando, FL, USA: Academic Press, 1986, vol. 24.

[19] E. Davison and T. Chang, "Decentralized stabilization and pole assignment for general proper systems," *IEEE Trans. Autom. Control*, vol. 35, no. 6, pp. 652–664, 1990.

[20] D. D. Siljak, *Large-Scale Dynamic Systems: Stability and Structure*. New York, NY, USA: North-Holland, 1978, vol. 310.

[21] C. W. Wu and L. O. Chua, "Synchronization in an array of linearly coupled dynamical systems," *IEEE Trans. Circuits Syst. I:, Fundam. Theory Appl.*, vol. 42, no. 8, pp. 430–447, 1995.

[22] I. Dobson, D. Newman, B. Carreras, and V. Lynch, "An initial model for complex dynamics in electric power system blackouts," in *Proc. 46th Hawaii Int. Conf. Syst. Sci.*, vol. 2, 2001, p. 2017.

[23] C. Asavathiratham, S. Roy, B. Lesieutre, and G. Verghese, "The influence model," *IEEE Control Syst.*, vol. 21, no. 6, pp. 52–64, 2001.

[24] R. Olfati-Saber and R. M. Murray, "Consensus problems in networks of agents with switching topology and time-delays," *IEEE Trans. Autom. Control*, vol. 49, no. 9, pp. 1520–1533, 2004.

[25] W. Ren *et al.*, "Consensus seeking in multiagent systems under dynamically changing interaction topologies," *IEEE Trans. Autom. Control*, vol. 50, no. 5, pp. 655–661, 2005.

[26] S. Roy, A. Saberi, and K. Herlugson, Formation and alignment of distributed sensing agents with double-integrator dynamics and actuator saturation," *Sensor Network Applications*. New York, NY, USA: IEEE Press, 2004.

[27] L. Xiao, S. Boyd, and S.-J. Kim, "Distributed average consensus with least-mean-square deviation," *J. Parallel Distrib. Comput.*, vol. 67, no. 1, pp. 33–46, 2007.

[28] R. L. Raffard, C. J. Tomlin, and S. P. Boyd, "Distributed optimization for cooperative agents: Application to formation flight," in *Proc. 43rd IEEE CDC*, vol. 3, 2004, pp. 2453–2459.

[29] R. Pastor-Satorras and A. Vespignani, "Epidemics and immunization in scale-free networks," *Handbook of Graphs and Networks: From the Genome to the Internet*, pp. 111–130, 2005.

[30] I. Kamwa, R. Grondin, and Y. Hebert, "Wide-area measurement based stabilizing control of large power systems-a decentralized/hierarchical approach," *IEEE Trans. Power Syst.*, vol. 16, no. 1, pp. 136–153, 2001.

[31] C. Taylor, C. Wanke, Y. Wan, and S. Roy, "A decision support tool for flow contingency management," in *Proc. Conf. AIAA Guidance Navigat. Control*, 2012, pp. 1–21.

[32] X. F. Wang and G. Chen, "Pinning control of scale-free dynamical networks," *Phys. A: Statist. Mech. Appl.*, vol. 310, no. 3, pp. 521–531, 2002.

[33] W. Yu, G. Chen, and J. Lü, "On pinning synchronization of complex dynamical networks," *Automatica*, vol. 45, no. 2, pp. 429–435, 2009.

[34] X. Li, X. Wang, and G. Chen, "Pinning a complex dynamical network to its equilibrium," *IEEE Trans. Circuits Syst. I, Reg. Papers*, vol. 51, no. 10, pp. 2074–2087, 2004.

[35] S. Roy, Y. Wan, and A. Saberi, "On time-scale designs for networks," *Int. J. Control*, vol. 82, no. 7, pp. 1313–1325, 2009.

[36] B. C. Eaves, A. J. Hoffman, U. G. Rothblum, and H. Schneider, "Line-sum-symmetric scalings of square nonnegative matrices," in *Mathematical Programming Essays in Honor of George B. Dantzig Part II*. New York, NY, USA: Springer-Verlag, 1985, pp. 124–141.

[37] C. R. Johnson, D. P. Stanford, D. Dale Olesky, and P. van den Driessche, "Dominant eigenvalues under trace-preserving diagonal perturbations," *Linear Algebra Appl.*, vol. 212, pp. 415–435, 1994.

[38] A. Berman and R. J. Plemmons, "Nonnegative matrices," ser. Mathematical Sciences, Classics Applied Mathematics, vol. 9. Philadelphia, PA, USA: SIAM, 1979.

[39] E. Kaszkurewicz, A. Bhaya, A. Bhaya, and E. Kaszkurewicz, *Matrix Diagonal Stability in Systems and Computation*. New York, NY, USA: Springer-Verlag, 2000.

[40] U. S. Bhalla and R. Iyengar, "Emergent properties of networks of biological signaling pathways," *Science*, vol. 283, no. 5400, pp. 381–387, 1999.

[41] J. W. Haefner, *Modeling Biological Systems: Principles and Applications*. New York, NY, USA: Springer Science & Business Media, 2005.

[42] P. Sebastiao and C. G. Soares, "Modeling the fate of oil spills at sea," *Spill Sci. Technol. Bulletin*, vol. 2, no. 2, pp. 121–131, 1995.

[43] P. Tkalich, K. Huda, and K. Y. Hoong Gin, "A multiphase oil spill model," *J. Hydraulic Res.*, vol. 41, no. 2, pp. 115–125, 2003.

[44] Y.-H. Chu, S. G. Rao, and H. Zhang, "A case for end system multicast (keynote address)," in *Proc. ACM SIGMETRICS Perform. Eval. Review*, vol. 28, no. 1, pp. 1–12, 2000.

[45] S. H. Strogatz, "Exploring complex networks," *Nature*, vol. 410, no. 6825, pp. 268–276, 2001.

[46] F. R. Chung, *Spectral Graph Theory*. Providence, RI, USA: American Mathematical Society, 1997, vol. 92.

[47] M. Fiedler, *Special Matrices and Their Applications in Numerical Mathematics*. New York, NY, USA: Courier Dover Publications, 2008.

[48] W. J. Rugh, *Linear System Theory*. Upper Saddle River, NJ, USA: Prentice Hall, 1996, vol. 2.

[49] D. P. Bertsekas, "Nonlinear Programming," 1999.

[50] J. E. Cohen, "Convexity of the dominant eigenvalue of an essentially nonnegative matrix," *Proc. Amer. Math. Society*, vol. 81, no. 4, pp. 657–658, 1981.

[51] M. H. Schneider and S. A. Zenios, "A comparative study of algorithms for matrix balancing," *Oper. Res.*, vol. 38, no. 3, pp. 439–455, 1990.

[52] T.-Y. Chen and J. W. Demmel, "Balancing sparse matrices for computing eigenvalues," *Linear Algebra Appl.*, vol. 309, no. 1, pp. 261–287, 2000.

[53] G. W. Stewart and J.-G. Sun, "Matrix Perturbation Theory," 1990.

[54] J. Abad Torres and S. Roy, "Graph-theoretic characterisations of zeros for the input-output dynamics of complex network processes," *Int. J. Control*, vol. 87, no. 5, pp. 940–950, 2014.

[55] J. Abad Torres and S. Roy, "Dominant eigenvalue minimization with trace preserving diagonal perturbation: Subset design problem," in *Proc. IEEE 54th Annu. CDC*, Dec. 2015, pp. 4208–4213.

[56] S. Roy, T. F. McElwain, and Y. Wan, "A Network Control Theory Approach to Modeling and Optimal Control of Zoonoses: Case Study of Brucellosis Transmission in Sub-Saharan Africa," 2011.

[57] M. Morris, "Epidemiology and social networks: Modeling structured diffusion," *Sociological Methods Res.*, vol. 22, no. 1, pp. 99–126, 1993.

[58] R. S. Nasci, "Movement of chikungunya virus into the western hemisphere," *Emerging Infectious Diseases*, vol. 20, no. 8, p. 1394, 2014.

Jackeline Abad Torres (M'15) received the B.S. degree from Escuela Politécnica Nacional, Quito, Ecuador, in 2008, and the M.S. and Ph.D. degrees in electrical engineering from Washington State University, Pullman, in 2012 and 2014, respectively.

She held a Fulbright grant in 2010. She is currently an Assistant Professor with the Departamento de Automatización y Control Industrial, Facultad de Ingeniería Eléctrica y Electrónica, Escuela Politécnica Nacional, Quito. Her current research interests include structural analysis and controller design of dynamical networks with applications to sensor/vehicle networking, epidemic control, and power systems network control.

Sandip Roy (M'04) received the B.S. degree in electrical engineering from the University of Illinois at Urbana-Champaign, Urbana, IL, USA, in 1998, and the M.S. and Ph.D. degrees in electrical engineering from the Massachusetts Institute of Technology, Cambridge, MA, USA, in 2000 and 2003, respectively.

He joined the School of Electrical Engineering and Computer Science, Washington State University, Pullman, in September 2003 as an Assistant Professor. He is currently a Professor and Associate Director of the School of Electrical Engineering and Computer Science, Washington State University, Pullman. He has also held visiting summer appointments with the University of Wisconsin-Madison and the Ames Research Center, National Aeronautics and Space Administration, Moffett Field, CA. His current interests include controller and topology design for dynamical networks, with applications to airtraffic control, computational biology, and sensor networking.

Yan Wan (S'08–M'09) received the B.S. degree from the Nanjing University of Aeronautics and Astronautics, Nanjing, China, in 2001, the M.S. degree from the University of Alabama, Tuscaloosa, AL, USA, in 2004, and the Ph.D. degree from Washington State University, Pullman, WA, USA, in 2009.

She was a Post-doctoral Scholar with the Control Systems Program, University of California at Santa Barbara, Santa Barbara, CA, USA. She is currently an Associate Professor with the Department of Electrical Engineering, University of North Texas, Denton, TX, USA. Her current research interests include decision-making tasks in large-scale networks, with applications to air traffic management, airborne networks, sensor networking, and biological systems.

Dr. Wan was a recipient of the prestigious National Science Foundation Early Career Development (CAREER) Award in 2015 and the Radio Technical Commission for Aeronautics (RTCA) William E. Jackson Award (excellence in aviation electronics and communication) in 2009.