SDP-Based Moment Closure for Epidemic Processes on Networks

Ximing Chen, Masaki Ogura, and Victor M. Preciado

Abstract—**In this paper, we analyze the (exact) stochastic dynamics of spreading processes taking place in complex networks. The analysis of this dynamics is, in general, very challenging since its state space grows exponentially with the network size. A common approach to overcome this challenge is to apply moment-closure techniques, such as the popular mean-field approach, to approximate the exact stochastic dynamics via ordinary differential equations. However, most existing moment-closure techniques do not provide quantitative guarantees on the quality of the approximation, limiting the applicability of these techniques. To overcome this limitation, we propose a novel moment-closure technique with explicit quality guarantees based on recent results relating the multidimensional moment problem with semidefinite programming. We illustrate how this technique can be used to derive upper and lower bounds on the exact (stochastic) dynamics of a variety of networked spreading processes, such as the SIS, SI, and SIR models. Moreover, we provide a simplified version of this moment-closure technique to approximate the dynamics of the probabilities of infection of each node using a linear number of piecewise-affine differential equations. Finally, we demonstrate the validity of our bounds via numerical simulations in a real-world social network.**

Index Terms—**Complex networks, epidemics, stochastic processes, moment closure, *K*-moment problem, semidefinite programming.**

I. Introduction

Modeling and analysis of spreading processes taking place in complex networks have found applications in a wide range of scenarios, such as modeling the propagation of malware in computer networks [1], failures in technological networks [2], memes in social networks [3], and diseases in human populations [4]–[6]. We find in the literature a wide variety of models to characterize the dynamics of spreading processes over networks. In the epidemiological literature, these models consider the spread of a disease in human contact networks in which individuals and their relationships are modeled via complex networks. Some of the most popular models in the literature are the *Susceptible-Infected-Susceptible* (SIS) [7], the *Susceptible-Infected-Recovered* (SIR) [8], and their variants [9], [10].

During the last decade, several mathematical techniques have been developed to determine whether a disease spreading over a network will be eradicated quickly or, in contrast, will spread widely over time, causing a large epidemic outbreak. These techniques can then be used to design efficient strategies to contain, or even eradicate, the spread of the disease by distributing medical resources throughout the network [11]–[14]. One of the most important characteristics in the global behavior of these models is the presence of phase transitions, or epidemic thresholds. These phase transitions can be described as a dynamical bifurcations in the dynamics, where the system transition from a single stable equilibrium at the origin (i.e., the disease-free state) towards the existence of (potentially many) nontrivial equilibria. The authors in [15] presented an approximate analysis to show that the networked SIS model presents a phase transition that can be characterized in terms of the largest eigenvalue of the adjacency matrix representing the network structure. A rigorous analysis of this phase transition was presented in [7], where the authors use Markov processes to model the exact stochastic dynamics of the networked SIS spreading process. Following this approach, the authors in [10] characterized the global dynamics of a more general spreading model which includes the SIS as a particular case.

A common idea behind the aforementioned results is to construct a Markov transition model and analyze the transition probabilities among network states. However, the number of possible network states grows exponentially with the number of nodes. Consequently, the analysis of the resulting Markov process is both computationally and analytically challenging to study. An alternative approach to overcome this challenge is to analyze the probability of each node being in a particular state at a given time. For example, in the SIS spreading model, it is mathematically convenient to analyze the time evolution of the probabilities of infection of each node in the network. However, as illustrated in [7], the ODEs describing the evolution of these infection probabilities depend on pairwise correlations (second-order moments) between the states of connected nodes in the network. As shown in [16], the governing dynamics of these second-order moments can also be described as ODEs involving third-order moments. In general, the ODEs describing the dynamics of k-th order moments depend on $(k+1)$-th order moment. Therefore, a complete characterization of the dynamics requires, in general, an exponential number of ODEs [17]. To address this issue, it is common to resort to moment-closure techniques—a method to

Manuscript received July 6, 2019; revised January 22, 2020; accepted April 1, 2020. Date of publication June 3, 2020; date of current version December 30, 2020. This work was partly supported by the Japan Society for the Promotion of Science KAKENHI under Grant JP18K13777 and the NSF under Grants CAREER-ECCS-1651433 and IIS-2008456. Recommended for acceptance by Dr. Gaoxi Xiao. *(Corresponding author: Ximing Chen.)*

Ximing Chen and Victor M Preciado are with the Department of Electrical and Systems Engineering, University of Pennsylvania, Philadelphia, PA 19104 USA (e-mail: ximingch@seas.upenn.edu; preciado@seas.upenn.edu).

Masaki Ogura is with the Graduate School of Information Science and Technology, Osaka University, Suita 565-0871, Japan (e-mail: m-ogura@ist.osaka-u.ac.jp).

Digital Object Identifier 10.1109/TNSE.2020.2998491

obtain a closed system of ODEs by approximating higher-order moments using lower-order ones [18]. Hence, it would be possible to use this technique to obtain a polynomial number of ODEs approximating the dynamics of moments of order k as a function of moments of order up to k. For instance, the popular mean-field approximation (MFA) is a moment-closure techniques in which pairwise correlations are approximated by the products of two first-order expectations [7], resulting in a linear number of ODEs. In [19], the authors proposed to close second-order moments using Fréchet inequalities; whereas the authors in [20] proposed to close third-order moments by products of first- and second-order moments.

Existing moment-closure techniques suffer from the following pitfalls. Firstly, there is no theoretical guarantee on the quality of the approximation obtained. Secondly, in the particular case of the SIS dynamics, these techniques often fail to lower bound the evolution of the probabilities of infection of each node. These pitfalls are partly addressed by [19], [21]; in particular, in [21] the authors showed that the moment-closure problem can be directly related to the *multidimensional moment problem* in functional analysis [22].

The objective of this paper is to provide a moment-closure framework with theoretical guarantee on the quality of the approximation for networked SIS model, along with several other disease spreading models. More specifically, in this paper, we further achieve the following contributions:

- We extend the work in [21] and propose a mathematical and computational framework to obtain a polynomial number of ODEs describing the dynamics of all k-th order moments of the SIS stochastic model for an arbitrary integer k and an arbitrary contact network.
- As part of this framework, we provide upper and lower bounds on the evolution of an arbitrary k-th moment of the SIS stochastic model.
- We provide a simplified expression for $k = 1$ to approximate the dynamics of the means of each node state using a linear number of piecewise-affine differential equations.
- Finally, we extend our framework to other compartmental spreading processes over networks, such as the SI and SIR models.

The rest of the paper is organized as follows. In Section II, we provide preliminaries and a description of the non-homogeneous SIS spreading process, as well as necessary background on the multidimensional moment problem. The proposed moment-closure framework is introduced in Section III, where we focus our attention on the networked SIS model. In Section IV, we discuss how to apply this moment-closure technique to both the SI and the SIR epidemic models. In Section V, we illustrate the performance of our framework by numerically analyzing several spreading processes taking place over a real-world social network. Finally, conclusions are presented in Section VI. All proofs of lemmas and theorems are available in Appendix A.

II. Preliminaries

We use bold fonts to represent vectors and capital letters to represent matrices. For a vector $\mathbf{x}$, we denote its transpose, i-th element, and 1-norm as $\mathbf{x}^\top$, x_i, and $|\mathbf{x}| = \sum_{i=1}^{n} |x_i|$, respectively. The n-dimensional vector of all ones is denoted by $\mathbf{1}_n \in \mathbb{R}^n$, and the m-by-n matrix of all ones by $J_{m,n} \in \mathbb{R}^{m\times n}$. The transpose of a matrix M is denoted by $M^\top$. If a matrix $M \in \mathbb{R}^{n\times n}$ is positive semi-definite, we write $M \succeq 0$. Given a positive integer $n \in \mathbb{N}$, we use the shorthand notation $[n]$ to represent the set of integers $\{1, \ldots, n\}$ and we let $\mathbb{N}_r^n = \{\mathbf{x} \in \mathbb{N}^n : |\mathbf{x}| \le r\}$.

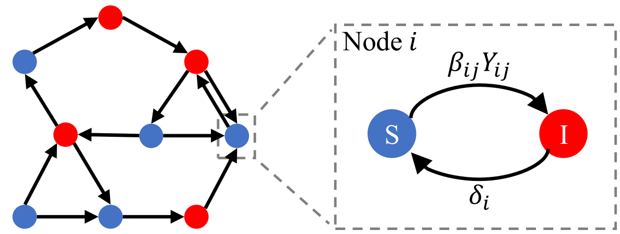


Fig. 1. Illustration of the *SIS* spreading model on a directed graph. In the left subfigure, the black arrows represent the edges in the digraph, whereas the red and blue circles represent the infected and susceptible nodes, respectively. In the right subfigure, we show the possible transitions between states of a node i. The variable Y_{ij} represents the event that an in-neighbor j of i is infected.

A. Heterogeneous Networked SIS Spreading Model

In this section, we describe the *Susceptible-Infected-Susceptible* (SIS) model, which is commonly used to characterize epidemics over networked populations. In the coming sections, we introduce a novel technique to analyze the stochastic dynamics of this, and other, epidemic models. For clarity in our exposition, we first illustrate the proposed technique using the SIS epidemic model; we then extend our analysis to other models, such as the SI and SIR models, in Section IV.

Let $G = (\mathcal{V}, \mathcal{E})$ be a directed graph, also called *digraph*, with node-set and edge-set denoted by $\mathcal{V} = [n]$ and $\mathcal{E} \subseteq \mathcal{V} \times \mathcal{V}$, respectively. The *out-neighborhood* (respectively, *in-neighborhood*) of node i is defined as $\mathcal{N}_i^+ = \{j \in \mathcal{V} : (i,j) \in \mathcal{E}\}$ (respectively, $\mathcal{N}_i^- = \{j \in \mathcal{V} : (j,i) \in \mathcal{E}\}$).

Next, we describe the continuous-time heterogeneous SIS spreading model on the graph G, [7]. In this model, at a given time $t \ge 0$, each node can be in one of the following two states: (i) '*Susceptible*,' representing the case of a healthy node, and (ii) '*Infected*,' in which the node is infected by a disease propagating through the network. On one hand, whenever node i is in the Susceptible state, i can be infected by one of its infected in-neighbour $j \in \mathcal{N}_i^-$ according to a Poisson process with parameter $\beta_{ij} > 0$, called the *infection rate* of edge (j,i). On the other hand, if node i is in the Infected state at a given time t, it cures itself according to a Poisson process with parameter δ_i, called the *recovery rate* of node i. We use a binary variable $x_i(t) \in \{0,1\}$ to represent the state of node $i \in \mathcal{V}$ at time $t \ge 0$. More specifically, $x_i(t) = 0$ if node i is Susceptible, and $x_i(t) = 1$ if it is Infected at time $t \ge 0$. We illustrate the SIS spreading model in Fig. 1.

The exact evolution of the random variables $x_i(t)$ can be characterized by a continuous-time Markov process with the following transition probabilities:

$$\begin{aligned}\mathbb{P}(x_i(t+h)=1\,|\,x_i(t)=0) &= \sum_{j\in\mathcal{N}_i^-}\beta_{ij}x_j(t)h+o(h),\\ \mathbb{P}(x_i(t+h)=0\,|\,x_i(t)=1) &= \delta_i h+o(h).\end{aligned} \tag{1}$$

Notice that the dimension of the state space of the Markov process in (1) is 2^n; hence, an exact analysis of the stochastic process is computationally challenging when the size of the underlying network is large. In what follows, we are interested in analyzing the dynamics of the probability of a node $i\in\mathcal{V}$ being infected at time t, i.e., $\mathbb{P}(x_i(t)=1)=\mathbb{E}[x_i(t)]$. As illustrated in [4], the governing equations for the evolution of the expectation of $x_i(t)$ is given by[1]

$$\frac{d}{dt}\mathbb{E}[x_i] = -\delta_i\mathbb{E}[x_i]+\sum_{j\in\mathcal{N}_i^-}\beta_{ij}\mathbb{E}[x_j]-\sum_{j\in\mathcal{N}_i^-}\beta_{ij}\mathbb{E}[x_ix_j]. \tag{2}$$

We refer to (2) as the *mean SIS dynamics* of node i. In order to solve (2), it is necessary to characterize the second-order moment $\mathbb{E}[x_ix_j]$ for all $j\in\mathcal{N}_i^-$. However, as shown in [16], the evolution of $\mathbb{E}[x_ix_j]$ depends on third-order moments of the form $\mathbb{E}[x_ix_jx_k]$, which in turn, forces us to characterize $\mathbb{E}[x_ix_jx_k]$. More generally, one can prove that, in order to characterize the evolution of a k-th order moment, one needs to characterize the time derivatives of moments of order $k+1$. As a result of this recursive dependency, the evolution of the mean SIS dynamics is fully characterized by 2^n ordinary differential equations.

In order to obtain a computationally tractable approximation of the mean SIS dynamics, it is common to use moment-closure techniques in which one approximates k-th order moments using lower-order moments (see for example [16]). In particular, the mean-field approximation (MFA) is a widely adopted moment-closure technique in which one assumes that $\mathbb{E}[x_ix_j]=\mathbb{E}[x_i]\mathbb{E}[x_j]$. Hence, defining the moment variable $\mu_i=\mathbb{E}[x_i]$, (2) turns into the following system of n nonlinear differential equations:

$$\dot{\mu}_i = -\delta_i\mu_i+\sum_{j\in\mathcal{N}_i^-}\beta_{ij}\mu_j-\sum_{j\in\mathcal{N}_i^-}\beta_{ij}\mu_i\mu_j. \tag{3}$$

Although, as shown in [23], this approximation provides an upper bound on the mean SIS dynamics of each node i, it is unclear how MFA can be generalized to provide upper or lower bounds for higher-order moment closures. In this paper, we develop a systematic framework to perform moment-closure with quality guarantees by using recent results on the K-moment problem [22]. The proposed framework is capable of providing both upper and lower bounds on the mean SIS spreading process. In the next subsection, we provide necessary background on the K-moment problem. We use the SIS model as a running example to illustrate the proposed technique. In Section IV, we will extend this technique to analyze other epidemic models, such as the networked SI and SIR models.

[1] Whenever clear from the context, we shall remove the time-dependent notation from the random variable $x_i(t)$.

B. The K-Moment Problem

To explain our approach, we first introduce the K-moment problem and related notions. Let $(\Omega,\mathcal{F},\mathbb{P})$ be a probability space, and $\mathcal{B}$ be a σ-algebra of $\mathbb{R}^n$ containing open sets. An $\mathbb{R}^n$-valued random variable is a function $X:\Omega\to\mathbb{R}^n$ such that for all $B\in\mathcal{B}$, $\{\omega: X(\omega)\in B\}\in\mathcal{F}$. Moreover, let $\lambda:\mathcal{B}\to[0,1]$ be a measure on $\mathbb{R}^n$; the support of λ, denoted as $\mathtt{Supp}(\lambda)$, is defined as the smallest closed set $C\subseteq\mathbb{R}^n$ such that $\lambda(\mathbb{R}^n\setminus C)=0$, [24]. Given an $\mathbb{R}^n$-valued random variable $\mathbf{x}\sim\lambda$ and an integer vector $\boldsymbol{\alpha}\in\mathbb{N}^n$, the $\boldsymbol{\alpha}$-moment of $\mathbf{x}$ is defined as $\mathbb{E}[\mathbf{x}^{\boldsymbol{\alpha}}]=\int_{\mathbb{R}^n}\prod_{i=1}^n x_i^{\alpha_i}d\lambda$. Moreover, the *order* of an $\boldsymbol{\alpha}$-moment is $|\boldsymbol{\alpha}|$. Finally, a sequence $\mathbf{y}=\{y_{\boldsymbol{\alpha}}\}_{\boldsymbol{\alpha}\in\mathbb{N}^n}$ indexed by $\boldsymbol{\alpha}$ is called a *multi-sequence*.

Definition 1: Let K be a closed set of $\mathbb{R}^n$. Let $\mathbf{y}=\{y_{\boldsymbol{\alpha}}\}_{\boldsymbol{\alpha}\in\mathbb{N}^n}$ be an infinite multi-sequence. A measure λ is said to be a K-representing measure for $\mathbf{y}$ if

$$y_{\boldsymbol{\alpha}}=\int_{\mathbb{R}^n}\mathbf{x}^{\boldsymbol{\alpha}}d\lambda,\ \text{for all } \boldsymbol{\alpha}\in\mathbb{N}^n, \tag{4}$$

and

$$\mathtt{Supp}(\lambda)\subseteq K. \tag{5}$$

If $\mathbf{y}$ has a K-representing measure, then we say that $\mathbf{y}$ is K-feasible.

In this paper, we are particularly interested in the case when K is a semi-algebraic set. A set $K\subseteq\mathbb{R}^n$ is called a *semi-algebraic* set if there exist a set of m polynomials $g_i:\mathbb{R}^n\to\mathbb{R}$ such that $K=\{\mathbf{x}\in\mathbb{R}^n: g_i(\mathbf{x})\geq 0 \text{ for all } i\in[m]\}$. A necessary and sufficient condition for the feasibility of the K-moment problem, restricted to the case when K is semi-algebraic and compact, can be stated in terms of linear matrix inequalities involving *moment matrices* and *localizing matrices*. In order to define these matrices, we introduce the following notions. Given an integer $r\in\mathbb{N}$, we define the vector

$$\mathbf{v}_r(\mathbf{x}) := \left[1,x_1,\ldots,x_n,x_1^2,x_1x_2,\ldots,x_1^r,\ldots,x_n^r\right]^\top, \tag{6}$$

i.e., the vector containing the monomials of the canonical basis of real-valued polynomials of degree at most r. Furthermore, given an integer vector $\boldsymbol{\alpha}=[\alpha_1,\ldots,\alpha_n]^\top\in\mathbb{N}_r^n$, we define $[\mathbf{v}_r]_{\boldsymbol{\alpha}}=\mathbf{x}^{\boldsymbol{\alpha}}$.

Definition 2: Given an $\mathbb{R}^n$-valued random variable $\mathbf{x}$, the moment matrix of $\mathbf{x}$ of order $2r$ is defined as $M_r=\mathbb{E}[\mathbf{v}_r(\mathbf{x})\mathbf{v}_r(\mathbf{x})^\top]$.

Let $\mathbf{y}=\{y_{\boldsymbol{\alpha}}\}_{|\boldsymbol{\alpha}|\leq 2r}$ be a finite t can index the entries of the moment matrix M_r using two elements of $\mathbb{N}_r^n$ as follows. Given two elements $\boldsymbol{\alpha},\boldsymbol{\beta}\in\mathbb{N}_r^n$, the $(\boldsymbol{\alpha},\boldsymbol{\beta})$-th entry of M_r, denoted by $[M_r]_{\boldsymbol{\alpha},\boldsymbol{\beta}}$, is equal to $\mathbb{E}[[\mathbf{v}_r]_{\boldsymbol{\alpha}}[\mathbf{v}_r]_{\boldsymbol{\beta}}]=\mathbb{E}[\mathbf{x}^{\boldsymbol{\alpha}}\mathbf{x}^{\boldsymbol{\beta}}]=y_{\boldsymbol{\alpha}+\boldsymbol{\beta}}$. For example, let $n=2$ and $r=1$; hence $\mathbf{v}_1(\mathbf{x})=[1,x_1,x_2]^\top$. Then, we have that the multi-sequence of moments is defined as $\mathbf{y}=\{y_{ij}\}_{i+j\leq 2}$, with $y_{ij}=\mathbb{E}[x_1^ix_2^j]$ for all $[i,j]^\top\in\mathbb{N}_2^2$. Hence, the moment matrix can be written as:

$$M_1(\mathbf{y})=\mathbb{E}\left[\begin{bmatrix}1\\x_1\\x_2\end{bmatrix}[1\ x_1\ x_2]\right]=\begin{bmatrix}y_{00}&y_{10}&y_{01}\\y_{10}&y_{20}&y_{11}\\y_{01}&y_{11}&y_{02}\end{bmatrix}.$$

Similarly, we define the *localizing matrices* as follows.

Definition 3: Given an $\mathbb{R}^n$-valued random variable $\mathbf{x}$, and a polynomial $g : \mathbb{R}^n \to \mathbb{R}$, we define the localizing matrix of $\mathbf{x}$ with respect to g as $L_r(g) = \mathbb{E}[g(\mathbf{x})\mathbf{v}_r(\mathbf{x})\mathbf{v}_r(\mathbf{x})^\top]$.

Let $deg(g)$ be the degree of the polynomial g. Then, g can be written as

$$g(\mathbf{x}) = \sum_{\boldsymbol{\gamma} \in \mathbb{N}^n_{deg(g)}} c_{\boldsymbol{\gamma}} \mathbf{x}^{\boldsymbol{\gamma}},$$

where $\mathbf{x}^{\boldsymbol{\gamma}}$ is a monomial (i.e., an entry in $\mathbf{v}_{deg(g)}(\mathbf{x})$) and $c_{\boldsymbol{\gamma}}$ is its corresponding coefficient. Let $\mathbf{y} = \{y_{\boldsymbol{\alpha}}\}_{|\boldsymbol{\alpha}| \le 2r+deg(g)}$ be the multi-sequence of moments such that $y_{\boldsymbol{\alpha}} = \mathbb{E}[\mathbf{x}^\alpha]$ for all $\boldsymbol{\alpha} \in \mathbb{N}^n_{2r+deg(g)}$. Hence, the entries of the localizing matrix can be indexed using two entries of $\mathbb{N}^n_{r+deg(g)}$, as follows. Given two elements $\boldsymbol{\alpha}, \boldsymbol{\beta} \in \mathbb{N}^n_r$, the $(\boldsymbol{\alpha}, \boldsymbol{\beta})$-th entry of $L_r(g)$, denoted by $[L_r(g)]_{\boldsymbol{\alpha},\boldsymbol{\beta}}$, is equal to $\mathbb{E}[g(\mathbf{x})[\mathbf{v}_r]_{\boldsymbol{\alpha}}[\mathbf{v}_r]_{\boldsymbol{\beta}}] = \mathbb{E}[\sum_{\boldsymbol{\gamma} \in \mathbb{N}^n_{deg(g)}} c_{\boldsymbol{\gamma}} \mathbf{x}^{\boldsymbol{\gamma}} \mathbf{x}^{\boldsymbol{\alpha}} \mathbf{x}^{\boldsymbol{\beta}}] = \sum_{\boldsymbol{\gamma} \in \mathbb{N}^n_{deg(g)}} c_{\boldsymbol{\gamma}} y_{\boldsymbol{\alpha}+\boldsymbol{\beta}+\boldsymbol{\gamma}}$. For example, let $n = 2$ and $r = 1$, and $g(\mathbf{x}) = 1 - x_1$, then the localizing matrix can be written as

$$L_1(g, \mathbf{y}) = \begin{bmatrix} y_{00} - y_{10} & y_{10} - y_{20} & y_{01} - y_{11} \\ y_{10} - y_{20} & y_{20} - y_{30} & y_{11} - y_{21} \\ y_{01} - y_{11} & y_{11} - y_{21} & y_{02} - y_{12} \end{bmatrix},$$

where $y_{\boldsymbol{\alpha}} = \mathbb{E}[x_1^i x_2^j]$, for $\boldsymbol{\alpha} = [i, j]^\top \in \mathbb{N}^2_3$.

In order to state a necessary and sufficient conditions for an infinite multi-sequence $\mathbf{y} = \{y_{\boldsymbol{\alpha}}\}_{\boldsymbol{\alpha} \in \mathbb{N}^n}$ to be K-feasible for a semi-algebraic set K, we first need to define the following notion [22].

Definition 4: A polynomial $p : \mathbb{R}^n \to \mathbb{R}$ is a sum-of-squares (SOS) if p can be written as

$$p(\mathbf{x}) = \sum_{j=1}^{J} p_j(\mathbf{x})^2, \tag{7}$$

for some finite set of polynomials $\{p_j : j \in [J]\}$.

A necessary and sufficient condition for an infinite multi-sequence $\mathbf{y} = \{y_{\boldsymbol{\alpha}}\}_{\boldsymbol{\alpha} \in \mathbb{N}^n}$ to be K-feasible, restricted to the case when K is both compact and semi-algebraic, is as follows.

Theorem 1: (Putinar's Positivstellensatz, [25]) Consider an infinite multi-sequence $\mathbf{y} = \{y_{\boldsymbol{\alpha}}\}_{\boldsymbol{\alpha} \in \mathbb{N}^n}$, and a collection of polynomials $g_i : \mathbb{R}^n \to \mathbb{R}$, for all $i \in [m]$. Define a compact semi-algebraic set $K = \{\mathbf{x} \in \mathbb{R}^n : g_i(\mathbf{x}) \ge 0,\ i \in [m]\}$. Assume that there exists a polynomial $u = u_0 + \sum_{i=1}^m u_i g_i$, where u_i are SOS polynomials for all $i \in [m]$, such that the set $\{\mathbf{x} : u(\mathbf{x}) \ge 0\}$ is compact. Then, the multi-sequence $\mathbf{y}$ has a K-representing measure, if and only if,

$$\begin{aligned} &M_r(\mathbf{y}) \succeq 0, \text{ and} \\ &L_r(g_j \mathbf{y}) \succeq 0, \text{ for all } j \in [m], \text{ and } r \in \mathbb{N}. \end{aligned} \tag{8}$$

Remark 1: Using (8) to verify the K-feasibility of a given multi-sequence requires checking the positive semi-definiteness of $m + 1$ matrices for each $r \in \mathbb{N}$. Moreover, the dimension of these matrices grows with r.

Based on Theorem 1, one can verify whether a given multi-sequence is a K-feasible moment sequence by solving an infinite sequence of semi-definite programs. On the other hand, given a finite moment sequence, up to a certain order, one can use Theorem 1 to derive conditions on higher-order moments for the multi-sequence of moments to be feasible. In the next section, we use this idea to provide upper and lower bounds on the evolution of $\mathbb{E}[x_i]$, described in (2).

III. SDP-Based Moment Closure

In this section, we first characterize the dynamics of the $\boldsymbol{\alpha}$-moment of the random vector describing the state of the SIS model, for an arbitrary $\boldsymbol{\alpha} \in \mathbb{N}^k$ (Section III-A). Then, we show that the problem of obtaining upper and lower bounds on the evolution of the $\boldsymbol{\alpha}$-moment is closely related to the K-moment problem (Section III-B). Finally, we obtain a closed-form expression for the mean dynamics of the SIS spreading process (Section III-C). In Section IV, we will extend our results to other networked epidemic models, such as the SI and SIR models.

A. Dynamics of the $\boldsymbol{\alpha}$-Moment in the SIS Spreading Process

As discussed in Section II-A, in order to close the system of differential equations describing the mean SIS dynamics (2), it is necessary to characterize the dynamics of second-order moments $\mathbb{E}[x_i x_j]$ for all $(i, j) \in \mathcal{E}$. More generally, in order to characterize the mean dynamics of any k-th order moment of the form $\mathbb{E}[x_{i_1} \cdots x_{i_k}]$, we need to obtain an expression for the $(k+1)$-th order differential $dx_{i_1} \cdots x_{i_{k+1}}$. To undertake the problem of finding a closed system of differential equations to describe the mean SIS spreading process, we propose the following three-step approach: First, we describe the stochastic dynamics of the networked SIS process using jump processes [26]. Second, we use Ito's formula for jump processes to obtain a governing equation for high-order differentials, i.e., an expression for $dx_{i_1} \cdots x_{i_k}$ for arbitrary k. Finally, we derive explicit differential equations allowing us to upper and lower bound the dynamics of any k-th order moment of the SIS spreading process. To achieve our goals, we first introduce related notions on Poisson jump processes.

Definition 5: [26] Given $\gamma > 0$, a stochastic process P_t^γ is called a Poisson jump process with rate γ if: (i) for every $s, t > 0$, the random variable $P_{s+t}^\gamma - P_s^\gamma$ is independent of $\{P_{t'}^\gamma : t' \le s\}$ and follows the same distribution as $P_t^\gamma - P_0^\gamma$, and (ii) the random variable $P_t^\gamma - P_0^\gamma$ follows a the Poisson distribution with mean γt, i.e., $\mathbb{P}(P_t^\gamma - P_0^\gamma = k) = e^{\gamma t} \frac{(\gamma t)^k}{k!}$.

In what follows, we abbreviate P_t^γ as P_γ for convenience. Using Poisson jump processes, the evolution of the states $x_i(t)$ in the SIS spreading process described in (1) can be characterized by the following set of stochastic differential equations:

$$dx_i = -x_i dP_{\delta_i} + (1 - x_i) \sum_{j \in \mathcal{N}_i^-} x_j dP_{\beta_{ij}}, \tag{9}$$

with $x_i(0) \in \{0, 1\}$ for all $i \in \mathcal{V}$. Notice that, we can recover the first-order mean dynamics of the SIS spreading process in (2) by taking expectation of (9). In order to obtain the dynamics of the second-order differential dx_ix_j, we use Ito's formula for jump processes, as stated below:

Theorem 2: [26] Let $\mathbf{x}(t)$ be an $\mathbb{R}^n$-valued random variable for all $t > 0$, and $\phi : \mathbb{R}^n \to \mathbb{R}$ be a twice continuously-differentiable function. If

$$d\mathbf{x}(t) = \sum_{k=1}^{n_p} \mathbf{h}_k(\mathbf{x}) dP_{\gamma_k}, \tag{10}$$

where $\mathbf{h}_k : \mathbb{R}^n \to \mathbb{R}^n$, for all $k \in [n_p]$, then

$$d\phi(\mathbf{x}) = \sum_{k=1}^{n_p} [\phi(\mathbf{x} + \mathbf{h}_k(\mathbf{x})) - \phi(\mathbf{x})] dP_{\gamma_k}. \tag{11}$$

As an example, we let $\phi(\mathbf{x}) = x_ix_j$, and apply Theorem 2 on (9). Subsequently, after tedious (but simple) algebraic manipulations, we obtain

$$\begin{aligned} dx_ix_j = & -x_ix_j(dP_{\delta_i} + dP_{\delta_j}) + (1 - x_i)x_j \sum_{k \in \mathcal{N}_i^-} x_k dP_{\beta_{ik}} \\ & + (1 - x_j)x_i \sum_{k \in \mathcal{N}_j^-} x_k dP_{\beta_{jk}}. \end{aligned} \tag{12}$$

If the SIS spreading process is homogeneous, i.e., $\delta_i = \delta$ for all $i \in [n]$ and $\beta_{ij} = \beta$ for all $(i, j) \in \mathcal{E}$, then taking the expectation of (12) results in:

$$\begin{aligned} \frac{d\mathbb{E}[x_ix_j]}{dt} = & -2\delta \mathbb{E}[x_ix_j] - \beta \sum_{k=1}^{n} (a_{jk} + a_{ik}) \mathbb{E}[x_ix_jx_k] \\ & + \beta \left[\sum_{k \in \mathcal{N}_i^-} \mathbb{E}[x_ix_k] + \sum_{k \in \mathcal{N}_j^-} \mathbb{E}[x_jx_k] \right], \end{aligned}$$

which reduces to the result in [16]. More generally, we can use (11) to derive explicit expressions of higher-order differentials, i.e., $dx_1^{\alpha_1} x_2^{\alpha_2} \cdots x_n^{\alpha_n}$ for arbitrary $\alpha_1, \ldots, \alpha_n \in \mathbb{N}$, for the SIS spreading model, as stated in the following theorem.

Theorem 3: Given a collection of k integers $i_1, \ldots, i_k \in [n]$ and a vector of positive integers $\boldsymbol{\alpha} \in \mathbb{N}^k$, we define the following monomials $\phi_{\boldsymbol{\alpha}}(\mathbf{x}) = x_{i_1}^{\alpha_1} \cdots x_{i_k}^{\alpha_k}$, $\phi_{\boldsymbol{\alpha}}^{-s}(\mathbf{x}) = x_{i_1}^{\alpha_1} \cdots x_{i_{s-1}}^{\alpha_{s-1}} x_{i_{s+1}}^{\alpha_{s+1}} \cdots x_{i_k}^{\alpha_k}$, and $\phi_{\mathbf{1}}(\mathbf{x}) = x_{i_1} \cdots x_{i_k}$. Consider a directed graph $G = (\mathcal{V}, \mathcal{E})$, and the set of stochastic differential equations described in (9). Then, the evolution of the $\boldsymbol{\alpha}$-moment satisfies

$$\begin{aligned} \frac{d\mathbb{E}[\phi_{\boldsymbol{\alpha}}(\mathbf{x})]}{dt} = & -\sum_{s=1}^{k} \delta_{i_s} \mathbb{E}[\phi_{\mathbf{1}}(\mathbf{x})] \\ & + \sum_{s=1}^{k} \sum_{\ell \in \mathcal{N}_{i_s}^-} \beta_{i_s \ell} \big(\mathbb{E}[\phi_{\mathbf{1}}^{-s}(\mathbf{x}) x_\ell] - \mathbb{E}[\phi_{\mathbf{1}}(\mathbf{x}) x_\ell] \big). \end{aligned} \tag{13}$$

Proof: See Appendix A. ■

Remark 2: As an example, consider $k = 2$, $i_1 = 1, i_2 = 2$ and $\boldsymbol{\alpha} = [1, 1]$, we have that $\phi_{\boldsymbol{\alpha}}(\mathbf{x}) = x_1x_2$, $\phi_{\boldsymbol{\alpha}}^{-i_1}(\mathbf{x}) = x_2$ and $\phi_{\boldsymbol{\alpha}}^{-i_2}(\mathbf{x}) = x_1$. Subsequently, from (13), we can obtain the time derivative of $\mathbb{E}[\phi_{\boldsymbol{\alpha}}(\mathbf{x})] = \mathbb{E}[x_1x_2]$, which follows similar expression in (12).

Theorem 3 shows that the time derivative of the $\boldsymbol{\alpha}$-moment $\mathbb{E}[\phi_{\boldsymbol{\alpha}}(\mathbf{x})]$ depends on $\mathbb{E}[\phi_{\mathbf{1}}(\mathbf{x}) x_\ell]$, which is a moment of higher order.

Remark 3: From (13), we have that the time derivatives of n first order moment $\mathbb{E}[x_i]$ depends on second-order moments of the form $\mathbb{E}[x_ix_j]$, which requires us to characterize the time derivatives of $\binom{n}{2}$ second-order moments. Inductively, the time derivatives of $\binom{n}{k}$ k-th-order moment is dependent on moments of oder $(k + 1)$; hence, a complete characterization of the dynamics requires $\sum_{k=1}^{n} \binom{n}{k} = 2^n$ number of ODEs, which is exponential on n.

In order to close the differential equation in (13), we propose to approximate $\mathbb{E}[\phi_{\mathbf{1}}(\mathbf{x}) x_\ell]$ using $\mathbb{E}[\phi_{\boldsymbol{\alpha}}(\mathbf{x})]$ for $|\boldsymbol{\alpha}| \leq k$. In the next section, we achieve this goal by upper and lower bound the term $\mathbb{E}[\phi_{\mathbf{1}}(\mathbf{x}) x_\ell]$.

B. SDP-Based Moment Closure

In this subsection, we will develop a framework to obtain both upper and lower bounds on the dynamics of the $\boldsymbol{\alpha}$-moment, $\mathbb{E}[\phi_{\boldsymbol{\alpha}}]$. In this direction, we will bound the higher-order term $\mathbb{E}[\phi_{\mathbf{1}}(\mathbf{x}) x_\ell]$ using lower-order moments $\mathbb{E}[\phi_{\boldsymbol{\beta}}(\mathbf{x})]$ for $|\boldsymbol{\beta}| \leq k$. For example, the widely used mean-field approximation [16] is an approach to close the first-order mean-dynamics $\mathbb{E}[x_i]$ by approximating the second-order term $\mathbb{E}[x_ix_j]$ using the following product of two first-order terms $\mathbb{E}[x_i]\mathbb{E}[x_j]$.

In what follows, we develop a framework to find two systems of differential equations whose solutions are guaranteed to upper and lower bound the dynamics of any $\boldsymbol{\alpha}$-moment. Our approach utilizes Putinar's Positivstellensatz to derive bounds on an $\boldsymbol{\alpha}$-moment in terms of lower-order moments. Before we present this approach, we first introduce several definitions. Given a set $\mathcal{I} \subseteq [n]$, we define $\mu_{\mathcal{I}} = \mathbb{E}[\Pi_{i \in \mathcal{I}} x_i]$. Furthermore, given a set of k *distinct* indices $\mathcal{I}_k = \{i_1, \ldots, i_k\} \subseteq [n]$, we define the (finite) multi-sequence of moments $\mathbf{y}(\mathcal{I}_k) = \{\mathbb{E}[\Pi_{s \in S} x_s]\}_{S \subseteq \mathcal{I}_k, |S| < k}$.

In what follows, we bound the moment $\mathbb{E}[\phi_{\boldsymbol{\alpha}}] = \mu_{\mathcal{I}_k}$ using lower-order moments contained in the set $\mathbf{y}(\mathcal{I}_k)$. To achieve this goal, we notice that at each time $t > 0$, $\mathbf{x}(t)$ is a $\{0, 1\}^n$-valued random variable. Subsequently, for a given time t, $[x_{i_1}(t), \ldots, x_{i_k}(t)]^\top$ follows a distribution supported on $\{0, 1\}^k$. In particular, the $\boldsymbol{\alpha}$-moment of $[x_{i_1}(t), \ldots, x_{i_k}(t)]^\top$ for $\boldsymbol{\alpha} = \mathbf{1}_k$ is equal to $\mu_{\mathcal{I}_k}$. Therefore, the sequence of moments $\hat{\mathbf{y}}(\mathcal{I}_k) = \mathbf{y}(\mathcal{I}_k) \cup \{\mu_{\mathcal{I}_k}\}$ must be $\{0, 1\}^k$-feasible (see Definition 1). Consequently, an upper bound (respectively, lower bound) on $\mu_{\mathcal{I}_k}$ can be obtained by finding the largest (respectively, smallest) value of $\mu_{\mathcal{I}_k}$ such that $\hat{\mathbf{y}}(\mathcal{I}_k)$ is $\{0, 1\}^k$-feasible.

To achieve the above objective, we propose to exploit the semidefinite inequalities in Theorem 1, regarding the moment and localizing matrices of $\mathbf{y}(\mathcal{I}_k)$. However, (8) provides conditions for an *infinite* sequence to be K-feasible whereas the sequence $\mathbf{y}(\mathcal{I}_k)$ is finite. To circumvent this issue, we will extend the finite multi-sequence of moments $\mathbf{y}(\mathcal{I}_k)$ into an infinite sequence such that the results in Theorem 1 are applicable. As we discuss below, this extension is possible due to the binary nature of the random variables x_i. More specifically, although $\mathbf{y}(\mathcal{I}_k)$ contains a finite sequence of moments, we can extend this sequence using the following observation: Given a set of q disjoint indices $\{i_1,\ldots,i_q\}\subseteq[n]$, since x_i are binary random variables, we have that:

$$\mathbb{E}[\Pi_{s=1}^q x_{i_s}]=\mathbb{E}[\Pi_{s=1}^q x_{i_s}^{\alpha_s}], \tag{14}$$

for all $q\le k$, where $\alpha_s>0$ for all $s\in[q]$. Subsequently, $\mathbf{y}(\mathcal{I}_k)$ can be extended uniquely into an infinite sequence, as follows: Given $\hat{\mathbf{y}}(\mathcal{I}_k)$, we construct its associated *infinite extension* as $\mathbf{y}_\infty(\mathcal{I}_k)=\{y_{\boldsymbol{\alpha}}\}_{\boldsymbol{\alpha}\in\mathbb{N}^k}$, with $y_{\boldsymbol{\alpha}}$ satisfying (14). Consequently, given a compact semi-algebraic set K, the K-feasibility of $\mathbf{y}(\mathcal{I}_k)$ is equivalent to the K-feasibility of $\mathbf{y}_\infty(\mathcal{I}_k)=\{y_{\boldsymbol{\alpha}}\}_{\boldsymbol{\alpha}\in\mathbb{N}^k}$.

Secondly, in order to apply Theorem 1, we show below (in Lemma 4) that the infinite-dimensional matrices in (8) are positive semidefinite, if and only if, certain finite-dimensional matrices are positive semidefinite. Before rigorously stating this claim, we need to introduce several additional notions. Given $k\le n$, we let $\kappa=\lceil k/2\rceil$, and $N_k=\binom{k+\kappa}{k}$. Finally, given $s\in[k]$, we let e_s denote the s-th standard basis vector of $\mathbb{R}^k$. With the help of these notions, we define the following finite-dimensional matrices. Let $M_\kappa(\hat{\mathbf{y}}(\mathcal{I}_k))\in\mathbb{R}^{N_k\times N_k}$ be defined entry-wise by

$$[M_\kappa(\hat{\mathbf{y}}(\mathcal{I}_k))]_{\boldsymbol{\alpha},\boldsymbol{\beta}}=y_{\boldsymbol{\alpha}+\boldsymbol{\beta}}, \tag{15}$$

for all $\boldsymbol{\alpha},\boldsymbol{\beta}\in\mathbb{N}^k_\kappa$. Essentially, if $y_{\boldsymbol{\gamma}}=\mathbb{E}[\mathbf{x}^{\boldsymbol{\gamma}}]$ for all $\boldsymbol{\gamma}\in\mathbb{N}^k$, then M_κ is the principal sub-matrix of size N_k of the infinite moment matrix in Theorem 1. In addition to M_κ, we now construct a collection of finite-dimensional matrices to replace the infinite-dimensional localizing matrices in Theorem 1. To achieve this goal, we first notice that the measure of the random vector $[x_{i_1},\ldots,x_{i_k}]^\top$ is supported on $\tilde{S}_k=[0,1]^k\supset\{0,1\}^k$, which is both compact and semi-algebraic. By defining $g_s^1(\mathbf{x})=1-x_{i_s}$ and $g_s^0(\mathbf{x})=x_{i_s}$ for all $s\in[k]$, the hypercube $\tilde{S}_k$ can be represented as

$$\tilde{S}_k=\{\mathbf{x}\in\mathbb{R}^k: g_s^1(\mathbf{x})\ge 0, g_s^0(\mathbf{x})\ge 0,\forall s\in[k]\}. \tag{16}$$

Next, for each $s\in[k]$, we define two matrices $[L_\kappa^1(\hat{\mathbf{y}}(\mathcal{I}_k),s)]$ and $[L_\kappa^0(\hat{\mathbf{y}}(\mathcal{I}_k),s)]$, as follows:

$$[L_\kappa^1(\hat{\mathbf{y}}(\mathcal{I}_k),s)]_{\boldsymbol{\alpha},\boldsymbol{\beta}}=y_{\boldsymbol{\alpha}+\boldsymbol{\beta}}-y_{\boldsymbol{\alpha}+\boldsymbol{\beta}+e_s}, \tag{17}$$

and

$$[L_\kappa^0(\hat{\mathbf{y}}(\mathcal{I}_k),s)]_{\boldsymbol{\alpha},\boldsymbol{\beta}}=y_{\boldsymbol{\alpha}+\boldsymbol{\beta}+e_s}. \tag{18}$$

As we will prove in Lemma 4, the matrices in (17) and (18) can be used as finite-dimensional localizing matrices for $g_s^1(\mathbf{x})$ and $g_s^0(\mathbf{x})$, respectively.

Remark 4: From (14), we see that whenever $y_{\boldsymbol{\alpha}+\boldsymbol{\beta}}=y_{\boldsymbol{\alpha}+\boldsymbol{\beta}+e_s}$, the corresponding entry in (17) is zero. More specifically, given an integer vector $\boldsymbol{\alpha}\in\mathbb{N}^k$, we define $\mathcal{A}=\{i:[\boldsymbol{\alpha}]_i\ne 0\}$ and $\mathcal{B}=\{i:[\boldsymbol{\beta}]_i\ne 0\}$. Consequently, $[L_\kappa^1(\hat{\mathbf{y}}(\mathcal{I}_k),s)]_{\boldsymbol{\alpha},\boldsymbol{\beta}}=0$, if and only if,

$$\mathcal{A}\cup\mathcal{B}=\mathcal{A}\cup\mathcal{B}\cup\{i_s\}. \tag{19}$$

Next, we show that the positive semidefiniteness of the infinite-dimensional matrices in Theorem 1 is equivalent to the positive semidefiniteness of the finite-dimensional matrices in (15), (17), and (18).

Lemma 4: Let $\{x_{i_s}\}_{s\in[k]}$ be a collection of binary random variables such that $\hat{\mathbf{y}}(\mathcal{I}_k)=\{\mathbb{E}[\Pi_{s\in S}x_s]\}_{S\subseteq\mathcal{I}_k}$, and denote its associated infinite extension by $\mathbf{y}_\infty(\mathcal{I}_k)$. Then, the sequence $\mathbf{y}_\infty(\mathcal{I}_k)$ is $\tilde{S}_k$-feasible, if and only if,

$$\begin{gathered}M_\kappa(\hat{\mathbf{y}}(\mathcal{I}_k))\succeq 0,\ \text{and}\\ L_\kappa^1(\hat{\mathbf{y}}(\mathcal{I}_k),s)\succeq 0, L_\kappa^0(\hat{\mathbf{y}}(\mathcal{I}_k),s)\succeq 0,\forall s\in[k].\end{gathered} \tag{20}$$

Proof: See Appendix A. ■

Remark 5: From (15), (17), and (18), we have that $M_\kappa(\hat{\mathbf{y}}(\mathcal{I}_k))=L_\kappa^1(\hat{\mathbf{y}}(\mathcal{I}_k),s)+L_\kappa^0(\hat{\mathbf{y}}(\mathcal{I}_k),s)$. Subsequently, positive semidefiniteness of $L_\kappa^1(\hat{\mathbf{y}}(\mathcal{I}_k),s)$ and $L_\kappa^0(\hat{\mathbf{y}}(\mathcal{I}_k),s)$ implies that $M_\kappa(\hat{\mathbf{y}}(\mathcal{I}_k))\succeq 0$.

Based on the above lemma, we can derive upper and lower bounds on the moment $\mathbb{E}[\boldsymbol{\phi}_{\mathbf{1}}(\mathbf{x})x_\ell]=\mu_{\mathcal{I}_k\cup\{\ell\}}$ in (13) by solving, respectively, the following semidefinite programs:

$$\overline{\mu}_{\mathcal{I}_k\cup\{\ell\}}=\max_{\mu_{\mathcal{I}_k\cup\{\ell\}}}\mu_{\mathcal{I}_k\cup\{\ell\}}\ \text{s.t. (20) holds.} \tag{21}$$

$$\underline{\mu}_{\mathcal{I}_k\cup\{\ell\}}=\min_{\mu_{\mathcal{I}_k\cup\{\ell\}}}\mu_{\mathcal{I}_k\cup\{\ell\}}\ \text{s.t. (20) holds.} \tag{22}$$

Hence, given a set $\mathcal{I}_k$, we have that $\mu_{\mathcal{I}_k\cup\{\ell\}}\in[\underline{\mu}_{\mathcal{I}_k\cup\{\ell\}},\overline{\mu}_{\mathcal{I}_k\cup\{\ell\}}]$. Based on this, one could be tempted to obtain an upper (respectively, a lower) bound on the evolution of $\mathbb{E}[\phi_{\boldsymbol{\alpha}}]$ by solving the ODE in (13) after replacing the higher-order term $\mathbb{E}[\boldsymbol{\phi}_{\mathbf{1}}(\mathbf{x})x_\ell]$ by $\underline{\mu}_{\mathcal{I}_k\cup\{\ell\}}$ (respectively, $\overline{\mu}_{\mathcal{I}_k\cup\{\ell\}}$). However, this is not true, since a monotone relationship between derivatives does not preserve the monotonicity between the solutions of the ODEs, as discussed in [27].

To address this issue, we propose to make slight modifications on the entries of the localizing matrices to invoke a multidimensional version of Grönwall's comparison lemma [27]. More specifically, for a given set $\mathcal{J}\subseteq\mathcal{I}_k$, let $\hat{\mu}_{\mathcal{J}}$ and $\check{\mu}_{\mathcal{J}}$ be upper and lower bounds on the moment $\mu_{\mathcal{J}}$, i.e., $\mu_{\mathcal{J}}\in[\check{\mu}_{\mathcal{J}},\hat{\mu}_{\mathcal{J}}]$. For a given $\boldsymbol{\gamma}\in\mathbb{N}^k_\kappa$, let us define $\mathcal{J}=\{i\in[n]:[\boldsymbol{\gamma}]_i\ne 0\}$, as well as $\hat{y}_{\boldsymbol{\gamma}}=\hat{\mu}_{\mathcal{J}}$ and $\check{y}_{\boldsymbol{\gamma}}=\check{\mu}_{\mathcal{J}}$. Let us also define the following modifications on the localizing matrices described in (17) and (18):

$$[\tilde{L}_\kappa^1(\hat{\mathbf{y}}(\mathcal{I}_k),s)]_{\boldsymbol{\alpha},\boldsymbol{\beta}}=\begin{cases}0, & \text{if (19) holds,}\\ \hat{y}_{\boldsymbol{\alpha}+\boldsymbol{\beta}}-\check{y}_{\boldsymbol{\alpha}+\boldsymbol{\beta}+e_s}, & \text{otherwise.}\end{cases}, \tag{23}$$

and

$$[\tilde{L}_\kappa^0(\mathbf{y}(\mathcal{I}_k),s)]_{\boldsymbol{\alpha},\boldsymbol{\beta}}=\hat{y}_{\boldsymbol{\alpha}+\boldsymbol{\beta}+e_s}. \tag{24}$$

In the next theorem, we formally show how to obtain upper and lower bounds on the evolution of $\mathbb{E}[\phi_{\boldsymbol{\alpha}}]$ using a modification of the ODE in (13) involving (23) and (24).

Theorem 5: Given a directed graph $G=(\mathcal{V},\mathcal{E})$, let us define a sequence of functions $\{\hat{\mu}_{\mathcal{I}}(t),\check{\mu}_{\mathcal{I}}(t)\}_{\mathcal{I}\subseteq[n],|\mathcal{I}|\le k}$ satisfying the following ODEs:

$$\frac{d\hat{\mu}_{\mathcal{I}}(t)}{dt} = -\sum_{s=1}^{|\mathcal{I}|}\delta_{i_s}\hat{\mu}_{\mathcal{I}}(t) + \sum_{s=1}^{|\mathcal{I}|}\sum_{\ell\in\mathcal{N}_{i_s}^{-}}\beta_{i_s\ell}\Big(\hat{\mu}_{\mathcal{I}\cup\{\ell\}\setminus\{i_s\}}(t) - \underline{\mu}_{\mathcal{I}\cup\{\ell\}}(t)\Big)$$

and

$$\frac{d\check{\mu}_{\mathcal{I}}(t)}{dt} = -\sum_{s=1}^{|\mathcal{I}|}\delta_{i_s}\check{\mu}_{\mathcal{I}}(t) + \sum_{s=1}^{|\mathcal{I}|}\sum_{\ell\in\mathcal{N}_{i_s}^{-}}\beta_{i_s\ell}\big(\check{\mu}_{\mathcal{I}\cup\{\ell\}\setminus\{i_s\}}(t) - \overline{\mu}_{\mathcal{I}\cup\{\ell\}}(t)\big),$$

for all $\mathcal{I}\subseteq[n], |\mathcal{I}|\le k$, where

$$\underline{\mu}_{\mathcal{I}\cup\{\ell\}} = \begin{cases}\check{\mu}_{\mathcal{I}\cup\{\ell\}}, & \text{if } |\mathcal{I}\cup\{\ell\}|\le k,\\ \check{\mu}^{\star}_{\mathcal{I}\cup\{\ell\}}, & \text{otherwise,}\end{cases} \tag{25}$$

and

$$\overline{\mu}_{\mathcal{I}\cup\{\ell\}} = \begin{cases}\hat{\mu}_{\mathcal{I}\cup\{\ell\}}, & \text{if } |\mathcal{I}\cup\{\ell\}|\le k,\\ \hat{\mu}^{\star}_{\mathcal{I}\cup\{\ell\}}, & \text{otherwise.}\end{cases} \tag{26}$$

In particular, $\check{\mu}^{\star}_{\mathcal{I}\cup\{\ell\}}$ and $\hat{\mu}^{\star}_{\mathcal{I}\cup\{\ell\}}$ are, respectively, the solutions that minimize/maximize the following SDPs:

$$\begin{aligned}\min_{\mu_{\mathcal{I}\cup\{\ell\}}} / \max_{\mu_{\mathcal{I}\cup\{\ell\}}}\ & \mu_{\mathcal{I}\cup\{\ell\}}\\ \text{subject to } & \tilde{L}^1_{\kappa}(\mathbf{y}(\mathcal{I}\cup\{\ell\}),s)\succeq 0, \forall s\in[k],\\ & \tilde{L}^0_{\kappa}(\mathbf{y}(\mathcal{I}\cup\{\ell\}),s)\succeq 0, \forall s\in[k].\end{aligned} \tag{27}$$

Let $\mu_{\mathcal{I}}(t) = \mathbb{E}[\phi_{\boldsymbol{\alpha}}]$ be the solution of the ODE in (13). Then, if $\hat{\mu}_{\mathcal{I}}(0)\ge\mu_{\mathcal{I}}(0)\ge\check{\mu}_{\mathcal{I}}(0)$, we have that $\hat{\mu}_{\mathcal{I}}(t)\ge\mu_{\mathcal{I}}(t)\ge\check{\mu}_{\mathcal{I}}(t)$, for all $t\ge 0$ and $\mathcal{I}\in[n], |\mathcal{I}|\le k$. ⋄

Proof: See Appendix A. ∎

In the above theorem, we have provided an SDP-based moment-closure procedure for SIS spreading process. More specifically, when $|\mathcal{I}|<k$, the ODEs in the statement of the above theorem resemble the ODE in (13). Nonetheless, when $|\mathcal{I}|=k$, the term $\mathbb{E}[\phi_{\mathbf{1}}(\mathbf{x})x_\ell]$ in (13) may be of order $k+1$; hence, the resulting system of ODEs cannot be solved. In the above theorem, we derive bounds for moments that are of order $k+1$ by solving the finite-dimensional SDPs in (27). Notice that these SDPs involve, solely, moments of order up to k. Consequently, all the moments in the ODEs in Theorem 5 are of order less or equal to k, resulting in a closed system of differential equations. The theorem also states that, when the ODEs in Theorem 5 share the same initial conditions as the ODE in (13), the solutions are upper and lower bounds on the exact dynamics of $\mathbb{E}[\phi_{\boldsymbol{\alpha}}(\mathbf{x})]$. In the next section, we illustrate the proposed approach to perform a first-order moment-closure of the mean dynamics of the SIS spreading process.

C. First-Order Moment Closure

In theory, we can upper and lower bound the dynamics of the mean spreading process in (2) by solving the ODEs in Theorem 5. In practice, these ODEs are solved via numerical methods using a discretized time interval. Notice that, according to Theorem 5, we need to solve the SDPs in (25) and (26) in each time step, which can be computationally very challenging in large-scale applications. To undertake this issue, we will develop a simplified procedure by finding a closed-form solution of the SDPs for the first-order mean dynamics. Our approach towards deriving a closed-form expression consists of two steps: First, we explicitly write the moment and localizing matrices in (20) for the first-order mean dynamics; then, we use a generalized version of Sylvester's criterion [28] to find a closed-form solution of the resulting SDPs.

When considering the first-order mean-dynamics in (2), we aim to derive upper and lower bounds on the second-order moments $\mu_{ij}=\mathbb{E}[x_ix_j]$ for $i\ne j$, in terms of first-order moments $\mu_i=\mathbb{E}[x_i]$. In this case, since $k=2$, we have that $\mathcal{I}_2=\{i,j\}$ (i.e., $i_1=i$ and $i_2=j$). Subsequently, the multi-sequence of interest $\mathbf{y}(\mathcal{I}_2)$ is given by $\mathbf{y}(\mathcal{I}_2)=\{1,\mu_i,\mu_j\}$, and its associated infinite extension is equal to $\mathbf{y}_\infty(\mathcal{I}_2)=\{1,\mu_i,\mu_j,\mu_{ij},\ldots\}$. More specifically, in the multisequence $\mathbf{y}_\infty(\mathcal{I}_2)$, the entries are indexed as follows: $\mu_{\boldsymbol{\alpha}}=\mu_i$ if $\alpha_2=0$, $\mu_{\boldsymbol{\alpha}}=\mu_j$ if $\alpha_1=0$, and $\mu_{\boldsymbol{\alpha}}=\mu_{ij}$ otherwise. Subsequently, according to (15), we have that

$$M_1(\hat{\mathbf{y}}(\mathcal{I}_2)) = \begin{bmatrix}1 & \mu_i & \mu_j\\ \mu_i & \mu_i & \mu_{ij}\\ \mu_j & \mu_{ij} & \mu_j\end{bmatrix}. \tag{28}$$

Moreover, from (17) and (18), we construct the following four matrices:

$$\begin{aligned}L_1^0(\hat{\mathbf{y}}(\mathcal{I}_2),i) &= \begin{bmatrix}\mu_i & \mu_i & \mu_{ij}\\ \mu_i & \mu_i & \mu_{ij}\\ \mu_{ij} & \mu_{ij} & \mu_{ij}\end{bmatrix},\\ L_1^0(\hat{\mathbf{y}}(\mathcal{I}_2),j) &= \begin{bmatrix}\mu_j & \mu_{ij} & \mu_j\\ \mu_{ij} & \mu_{ij} & \mu_{ij}\\ \mu_j & \mu_{ij} & \mu_j\end{bmatrix}.\end{aligned} \tag{29}$$

$$L_1^1(\hat{\mathbf{y}}(\mathcal{I}_2),i) = \begin{bmatrix}1-\mu_i & 0 & \mu_j-\mu_{ij}\\ 0 & 0 & 0\\ \mu_j-\mu_{ij} & 0 & \mu_j-\mu_{ij}\end{bmatrix}, \tag{30}$$

and

$$L_1^1(\hat{\mathbf{y}}(\mathcal{I}_2),j) = \begin{bmatrix}1-\mu_j & \mu_i-\mu_{ij} & 0\\ \mu_i-\mu_{ij} & \mu_i-\mu_{ij} & 0\\ 0 & 0 & 0\end{bmatrix}. \tag{31}$$

Hence, according to Lemma 4, the sequence $\mathbf{y}_\infty(\mathcal{I}_2)$ has an $\tilde{S}_2$-representing measure with $\tilde{S}_2=\{\mathbf{x}\in\mathbb{R}^2 : x_i,x_j\in[0,1]\}$, if and only if, (28)–(31) are all positive semidefinite. The main idea of our approach is to use a generalized version of Sylvester's criterion to replace the linear matrix inequalities in (27) by polynomial inequalities, as shown in the theorem below.

Theorem 6: Consider a directed graph $G=(\mathcal{V},\mathcal{E})$ and a set of n initial values $\{\mu_i(0)\}_{i=1}^n$. Let us define two sequences of

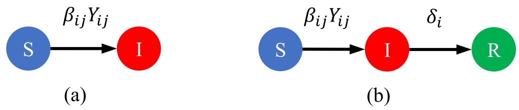


Fig. 2. In (a) and (b), we show the transition between states in the SI-spreading process and SIR-spreading process, respectively.

functions $\{\hat{\mu}_i(t)\}_{i=1}^n$ and $\{\check{\mu}_i(t)\}_{i=1}^n$ satisfying the following ODEs:

$$\frac{d\hat{\mu}_i}{dt} = -\delta_i\hat{\mu}_i + \sum_j \beta_{ij}\hat{\mu}_j - \sum_j \beta_{ij}\overline{\mu}_{ij},$$

$$\frac{d\check{\mu}_i}{dt} = -\delta_i\check{\mu}_i + \sum_j \beta_{ij}\check{\mu}_j - \sum_j \beta_{ij}\underline{\mu}_{ij},$$

with $\hat{\mu}_i(0) = \check{\mu}_i(0) = \mu_i(0)$, where

$$\overline{\mu}_{ij} = \max\{\hat{\mu}_i + \hat{\mu}_j - 1, 0\}, \quad (32)$$
$$\underline{\mu}_{ij} = \min\{\check{\mu}_i, \check{\mu}_j\}. \quad (33)$$

Then $\hat{\mu}_i(t) \geq \mu_i(t) \geq \check{\mu}_i(t)$, for all $t \geq 0$ and $i \in [n]$. $\diamond$

Proof: See Appendix A. ■

Several remarks are in order. First, note that we do not need to solve a semidefinite program to numerically find the upper and lower bounds stated in the above theorem. Instead, we need to solve a system of $2n$ piece-wise affine differential equations, where the piece-wise nonlinearities are described in (32) and (33). Furthermore, it is, in principle, possible to use the proposed approach to obtain a whole hierarchy of moment closures by considering higher-order moments. For example, we could derive a system of $n + m$ differential equations, where m is the number of edges in the graph, using both n first-order and m second-order moments. Finally, it is worth noting that the proposed technique can be generalized to analyze the mean dynamics of other spreading processes, as we illustrate in the next section.

IV. Moment Closure of Other Popular Spreading Models

In this section, we will apply the SDP-based moment closure framework herein proposed to find upper and lower bounds on the stochastic dynamics of two other networked epidemic models, namely, the SI and the SIR models.

A. Susceptible-Infected (SI) Epidemic Model

In the SI networked epidemic model [6], a susceptible node can be infected by its infected in-neighbors; however, once the node is infected, it remains infectious forever (see Fig. 2-(a) for a detailed transition diagram). Let x_i be a binary random variable representing the state of node i, where $x_i(t) = 0$ if node i is susceptible at time t and $x_i(t) = 1$ if it is infected. The stochastic dynamics of the networked SI process can be modeled using the following jump-process:

$$dx_i = (1 - x_i) \sum_{j\in\mathcal{N}_i^-} x_j dP_{\beta_{ij}}. \quad (34)$$

Notice that this SDE is similar to (9), after removing the term describing the recovery process. Consequently, using the techniques used to prove Theorem 3, we can readily obtain the following ODE describing the evolution of any moment $\mu_{\mathcal{I}}(t)$, for any choice of $\mathcal{I} \subseteq [n]$:

$$\frac{d\mu_{\mathcal{I}}(t)}{dt} = \sum_{s=1}^{|\mathcal{I}|} \sum_{\ell\in\mathcal{N}_{i_s}^-} \beta_{i_s\ell}\big(\mu_{\mathcal{I}\cup\{\ell\}\setminus\{i_s\}} - \mu_{\mathcal{I}\cup\{\ell\}}\big). \quad (35)$$

Since the random variables defining the states of nodes in the network are binary, we can use the techniques used in the analysis of the SIS model to find upper and lower bounds in the moment dynamics. In particular, the finite-dimensional moment and localizing matrices proposed in Section III-B can be directly used in here. Thus, we can obtain the following corollary from Theorem 5.

Corollary 1: Given a directed graph $G = (\mathcal{V}, \mathcal{E})$, let us define two sequences of functions $\{\hat{\mu}_{\mathcal{I}}(t)\}_{\mathcal{I}\subseteq[n],|\mathcal{I}|\leq k}$ and $\{\check{\mu}_{\mathcal{I}}(t)\}_{\mathcal{I}\subseteq[n],|\mathcal{I}|\leq k}$ satisfying the following ODE's:

$$\frac{d\hat{\mu}_{\mathcal{I}}(t)}{dt} = \sum_{s=1}^{|\mathcal{I}|} \sum_{\ell\in\mathcal{N}_{i_s}^-} \beta_{i_s\ell}\Big(\hat{\mu}_{\mathcal{I}\cup\{\ell\}\setminus\{i_s\}}(t) - \underline{\mu}_{\mathcal{I}\cup\{\ell\}}(t)\Big)$$

$$\frac{d\check{\mu}_{\mathcal{I}}(t)}{dt} = \sum_{s=1}^{|\mathcal{I}|} \sum_{\ell\in\mathcal{N}_{i_s}^-} \beta_{i_s\ell}\big(\check{\mu}_{\mathcal{I}\cup\{\ell\}\setminus\{i_s\}}(t) - \overline{\mu}_{\mathcal{I}\cup\{\ell\}}(t)\big),$$

for all $\mathcal{I} \subseteq [n], |\mathcal{I}| \leq k$, where $\overline{\mu}_{\mathcal{I}\cup\{\ell\}}$ and $\underline{\mu}_{\mathcal{I}\cup\{\ell\}}$ are defined as in (25) and (26), respectively. If $\hat{\mu}_{\mathcal{I}}(0) \geq \mu_{\mathcal{I}}(0) \geq \check{\mu}_{\mathcal{I}}(0)$, then $\hat{\mu}_{\mathcal{I}}(t) \geq \mu_{\mathcal{I}}(t) \geq \check{\mu}_{\mathcal{I}}(t)$, for all $t \geq 0$ and $\mathcal{I} \in [n], |\mathcal{I}| \leq k$. $\diamond$

B. Susceptible-Infected-Removed (SIR) Spreading Process

In the case of SIR spreading process, nodes in G can be in one out of three states: *susceptible*, *infected*, or *removed*, at any time instance. A node is in the *removed* state when it has been infected in the past, it has recovered from the infection, and has developed permanent immunity to the disease (see Fig. 2-(b) for a detailed transition diagram); hence, it cannot be infected again in the future. In what follows, we use $\{0, 1\}$-valued random variables $x_{i,S}(t), x_{i,I}(t)$, and $x_{i,R}(t)$ to indicate whether node i is susceptible, infected, or removed at time t, respectively. Since node i can only be in exactly one compartment at every time instance, we have that $x_{i,S}(t) + x_{i,I}(t) + x_{i,R}(t) = 1$ for all $t \geq 0$. With these definitions, the two transition probabilities among states are characterized by:

$$\mathbb{P}(x_{i,I}(t+h) = 1 \,|\, x_{i,S}(t) = 1) = h \sum_{j\in\mathcal{N}_i^-} \beta_{ij}x_{j,I}(t) + o(h),$$
$$\mathbb{P}(x_{i,R}(t+h) = 1 \,|\, x_{i,I}(t) = 1) = \delta_i h + o(h). \quad (36)$$

We assume that a node is either susceptible or infected at time $t = 0$. The evolution of the network states are characterized by the following set of SDEs:

$$d\begin{bmatrix} x_{i,S} \\ x_{i,I} \\ x_{i,R} \end{bmatrix} = \begin{bmatrix} 0 \\ -x_{i,I} \\ x_{i,I} \end{bmatrix} dP_{\delta_i} + \sum_{j\in\mathcal{N}_i^-} \begin{bmatrix} -x_{i,S}x_{j,I} \\ x_{i,S}x_{j,I} \\ 0 \end{bmatrix} dP_{\beta_{ij}}, \quad (37)$$

for all $i \in [n]$, where all the Poisson jump-processes are independent. From (37), the expectation of the random variable $x_{i,S}$ satisfies

$$\frac{d\mathbb{E}[x_{i,S}(t)]}{dt} = \sum_{j\in\mathcal{N}_i^-} \beta_{ij}\mathbb{E}[x_{i,S}(t)x_{j,I}(t)]. \quad (38)$$

Therefore, in order to solve for $\mathbb{E}[x_{i,S}(t)]$, it is necessary to characterize the evolution of $\mathbb{E}[x_{i,S}x_{j,I}]$ over time.

In what follows, we apply the framework herein proposed to derive a closed system of ODEs bounding the mean SIR spreading process. We start by computing the mean dynamics of the SIR spreading process via Ito's formula, as follows.

Theorem 7: Consider the networked SIR process described in (37). Given the vectors $\boldsymbol{\alpha}, \boldsymbol{\beta}, \boldsymbol{\gamma} \in \mathbb{N}^n$, define the monomial $\phi_{\boldsymbol{\alpha},\boldsymbol{\beta},\boldsymbol{\gamma}}(\mathbf{x}) = \Pi_{i=1}^n x_{i,S}^{\alpha_i} x_{i,I}^{\beta_i} x_{i,R}^{\gamma_i}$. Then,

$$\begin{aligned} \frac{d\mathbb{E}[\phi_{\boldsymbol{\alpha},\boldsymbol{\beta},\boldsymbol{\gamma}}(\mathbf{x})]}{dt} &= -\sum_{s=1}^n \delta_s \mathbf{1}_{\beta_s\neq 0}\mathbb{E}[\phi_{\boldsymbol{\alpha},\boldsymbol{\beta},\boldsymbol{\gamma}}(\mathbf{x})] \\ &+ \sum_{s=1}^n \delta_s \mathbf{1}_{\beta_s=0\cap\gamma_s\neq 0}\mathbb{E}[\Pi_{k\in[n],k\neq\ell} x_{k,S}^{\alpha_k} x_{k,I}^{\beta_k} x_{k,R}^{\gamma_k} x_{\ell,S}^{\alpha_\ell} x_{\ell,I}] \\ &- \sum_{s=1}^n \sum_{\ell\in\mathcal{N}_s^-} \beta_{s\ell}\mathbf{1}_{\alpha_s\neq 0}\mathbb{E}[\phi_{\boldsymbol{\alpha},\boldsymbol{\beta},\boldsymbol{\gamma}}(\mathbf{x})x_{\ell,I}] \\ &+ \sum_{s=1}^n \sum_{\ell\in\mathcal{N}_s^-} \beta_{s\ell}\mathbf{1}_{\alpha_s=0\cap\beta_s\neq 0}\mathbb{E}[\phi_{\boldsymbol{\alpha},\boldsymbol{\beta},\boldsymbol{\gamma}}(\mathbf{x})x_{\ell,I}] \quad (39) \end{aligned}$$

Proof: See Appendix A. ■

Hereafter, given two sets of indices $\mathcal{I}, \mathcal{J} \subseteq n]$, we define

$$\mu_{\mathcal{I},\mathcal{J}} = \mathbb{E}[\Pi_{i\in\mathcal{I}} x_{i,S}\Pi_{j\in\mathcal{J}} x_{j,I}]. \quad (40)$$

In particular, when $\mathcal{I}$ (resp. $\mathcal{J}$) is a singleton, i.e., $\mathcal{I} = \{i\}$ (resp., $\mathcal{J} = \{j\}$), we also write $\mu_{\mathcal{I},\mathcal{J}} = \mu_{i,\mathcal{J}}$ (resp., $\mu_{\mathcal{I},\mathcal{J}} = \mu_{\mathcal{I},j}$). Letting $\boldsymbol{\gamma} = \mathbf{0}_n$ in (39), we obtain

$$\begin{aligned} \frac{d\mu_{\mathcal{I},\mathcal{J}}(t)}{dt} &= -\sum_{s\in\mathcal{J}} \delta_s\mu_{\mathcal{I},\mathcal{J}}(t) - \sum_{s\in\mathcal{I}}\sum_{\ell\in\mathcal{N}_s^-}\beta_{s\ell}\mu_{\mathcal{I},\mathcal{J}\cup\{\ell\}} \\ &+ \sum_{s\in\mathcal{J}\setminus\mathcal{I}}\sum_{\ell\in\mathcal{N}_s^-}\beta_{s\ell}\mu_{\mathcal{I},\mathcal{J}\cup\{\ell\}}, \quad (41) \end{aligned}$$

which depends only on moments of $x_{i,S}$ and $x_{i,I}$ for $i \in [n]$. In order to solve the above system of ODEs, we need to provide bounds on $\mu_{\mathcal{I},\mathcal{J}\cup\{\ell\}}$ using lower-order moments. To achieve this goal, similar to the case of SIS spreading process, we aim to construct moment and localizing matrices, as listed in (20). We consider a slight abuse of notations by replacing k in $\mathbf{y}(\mathcal{I}_k)$ with $|\mathcal{I}| + |\mathcal{J}|$, and $\mathbf{y}(\mathcal{I}_k)$ with

$$\mathbf{y}(\mathcal{I},\mathcal{J}) = \{\mu_{\mathcal{I}',\mathcal{J}'}\}_{|\mathcal{I}'|+|\mathcal{J}'|<k}.$$

With this definition, we construct finite-dimensional matrices analogous to the ones in (15), (17), and (18) using elements in $\mathbf{y}_{\mathcal{I},\mathcal{J}}$ accordingly. For example, to close the first-order mean dynamics of the SIR spreading process, the moment matrix defined in (15) becomes:

$$M_1(\mathbf{y}(\{i\},\{j\})) = \begin{bmatrix} 1 & \mu_{i,\emptyset} & \mu_{\emptyset,j} \\ \mu_{i,\emptyset} & \mu_{i,\emptyset} & \mu_{\emptyset,j} \\ \mu_{\emptyset,j} & \mu_{i,j} & \mu_{\emptyset,j} \end{bmatrix}. \quad (42)$$

Since $x_{i,S}$ and $x_{i,I}$ are binary random variables for all $i \in \mathcal{V}$, Lemma 4 can be applied without loss of generality.

To provide upper and lower bounds for the moment $\mu_{\mathcal{I},\mathcal{J}\cup\{\ell\}}$, we build $2k+1$ matrices using elements in $\mathbf{y}(\mathcal{I},\mathcal{J}\cup\{\ell\})$ and solve for the maximum and minimum value $\mu_{\mathcal{I},\mathcal{J}\cup\{\ell\}}$ such that those matrices are positive semidefinite. Denoting those extreme values by $\underline{\mu}_{\mathcal{I},\mathcal{J}\cup\{\ell\}}$ and $\overline{\mu}_{\mathcal{I},\mathcal{J}\cup\{\ell\}}$, we have that $\mu_{\mathcal{I},\mathcal{J}\cup\{\ell\}} \in [\underline{\mu}_{\mathcal{I},\mathcal{J}\cup\{\ell\}}, \overline{\mu}_{\mathcal{I},\mathcal{J}\cup\{\ell\}}]$. Finally, we adopt a similar treatment to the localizing matrices as in (23) and (24), i.e., replacing the entries within localizing matrices by upper and lower estimates $\hat{\mu}_{\mathcal{I},\mathcal{J}}$ and $\check{\mu}_{\mathcal{I},\mathcal{J}}$. We use $\check{\mu}^\star_{\mathcal{I},\mathcal{J}\cup\{\ell\}}$ and $\hat{\mu}^\star_{\mathcal{I},\mathcal{J}\cup\{\ell\}}$ to denote the lower and upper estimates of $\mu_{\mathcal{I},\mathcal{J}}$ obtained by solving SDPs using modified localizing matrices. As a result, we obtain the following theorem for the networked SIR epidemic model:

Theorem 8: Consider the networked SIR process described in (37). Let us define a sequence of functions $\{\hat{\mu}_{\mathcal{I},\mathcal{J}}(t), \check{\mu}_{\mathcal{I},\mathcal{J}}(t)\}_{\mathcal{I},\mathcal{J}\subseteq n],|\mathcal{I}|\leq k}$ satisfying the following ODEs:

$$\begin{aligned} \frac{d\hat{\mu}_{\mathcal{I},\mathcal{J}}(t)}{dt} &= -\sum_{s\in\mathcal{J}}\delta_s\hat{\mu}_{\mathcal{I},\mathcal{J}}(t) - \sum_{s\in\mathcal{I}}\sum_{\ell\in\mathcal{N}_s^-}\beta_{s\ell}\underline{\mu}_{\mathcal{I},\mathcal{J}\cup\{\ell\}}, \\ &+ \sum_{s\in\mathcal{J}\setminus\mathcal{I}}\sum_{\ell\in\mathcal{N}_s^-}\beta_{s\ell}\overline{\mu}_{\mathcal{I},\mathcal{J}\cup\{\ell\}}, \end{aligned}$$

$$\begin{aligned} \frac{d\check{\mu}_{\mathcal{I},\mathcal{J}}(t)}{dt} &= -\sum_{s\in\mathcal{J}}\delta_s\check{\mu}_{\mathcal{I},\mathcal{J}}(t) - \sum_{s\in\mathcal{I}}\sum_{\ell\in\mathcal{N}_s^-}\beta_{s\ell}\overline{\mu}_{\mathcal{I},\mathcal{J}\cup\{\ell\}}, \\ &+ \sum_{s\in\mathcal{J}\setminus\mathcal{I}}\sum_{\ell\in\mathcal{N}_s^-}\beta_{s\ell}\underline{\mu}_{\mathcal{I},\mathcal{J}\cup\{\ell\}}, \end{aligned}$$

where

$$\overline{\mu}_{\mathcal{I},\mathcal{J}\cup\{\ell\}} = \begin{cases} \hat{\mu}_{\mathcal{I}\cup\{\ell\}}, & \text{if } |\mathcal{I}| + |\mathcal{J}\cup\{\ell\}| \leq k, \\ \hat{\mu}^\star_{\mathcal{I},\mathcal{J}\cup\{\ell\}}, & \text{otherwise}, \end{cases} \quad (43)$$

and

$$\underline{\mu}_{\mathcal{I},\mathcal{J}\cup\{\ell\}} = \begin{cases} \check{\mu}_{\mathcal{I},\mathcal{J}\cup\{\ell\}}, & \text{if } |\mathcal{I}| + |\mathcal{J}\cup\{\ell\}| \leq k, \\ \check{\mu}^\star_{\mathcal{I},\mathcal{J}\cup\{\ell\}}, & \text{otherwise}, \end{cases} \quad (44)$$

for all $\mathcal{I}, \mathcal{J} \subseteq n]$. If $\hat{\mu}_{\mathcal{I},\mathcal{J}}(0) \geq \mu_{\mathcal{I},\mathcal{J}}(0) \geq \check{\mu}_{\mathcal{I},\mathcal{J}}(0)$, then $\hat{\mu}_{\mathcal{I},\mathcal{J}}(t) \geq \mu_{\mathcal{I},\mathcal{J}}(t) \geq \check{\mu}_{\mathcal{I},\mathcal{J}}(t)$, for all $\mathcal{I}, \mathcal{J} \subseteq n]$ and $t \geq 0$.

Proof: See Appendix A. ■

In the next section, we demonstrate the performance of the moment-closure framework herein proposed on both the SIS and SIR epidemic processes taking place in a real social network.

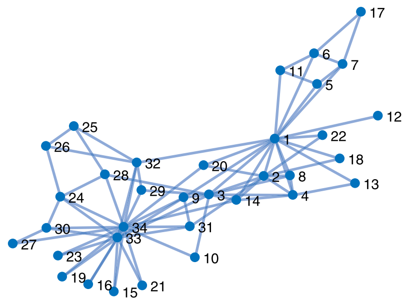


Fig. 3. Topology of the Zachary's Karate Club, representing friendships among 34 individuals.

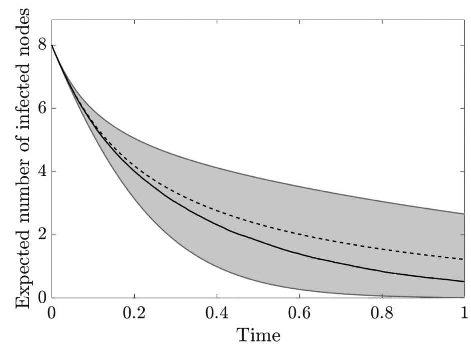


Fig. 4. This figure depicts upper and lower bounds on the expected number of infected nodes. The solid black line represent the empirical average over 10000 realizations of the number of infected nodes over time. The dashed line and the shaded region represent the expected number of infected nodes calculated via the mean-field approximation and the SDP-based moment-closure technique, respectively.

V. Simulation

In this section, we demonstrate the SDP-based moment-closure framework by finding upper and lower bounds on the probabilities of infection of all nodes in a real social network. In our first set of simulations (Section V-A), we implement the exact stochastic SIS spreading process, as described in [7]. We simulate 10,000 realizations of the stochastic process using the same initial conditions and compute the evolutions of the empirical average of the probabilities of infection, which is an approximation of the mean SIS dynamics. We then execute our SDP-based moment-closure technique, using Theorem 6, in order to obtain the upper and lower bounds on the mean SIS dynamics, $\hat{\mu}_i(t)$ and $\check{\mu}_i(t)$. Furthermore, we compare the time evolution of these bounds with the widely used mean-field approximation (3). In our second set of simulations (Section V-B), we apply similar analysis to the SIR spreading process. In our third set of simulations (Section V-C), we compare the time evolution of first-order moment closure and second-order moment closure obtained using Theorem 6 and 5, respectively.

A. Moment-Closure of the SIS Epidemic Process

In this subsection, we run the stochastic SIS dynamics over the Zachary's Karate Club [29], plotted in Fig. 3. In our experiments, we choose the individuals with labels $\mathcal{S} = \{3, 5, 6, 14, 16, 17, 20, 23\}$ to be initially infected. The infection rates satisfy $\beta_{ij} = \beta = 1$ for all $(i, j) \in \mathcal{E}$ and the recovery rates are $\delta_i = \delta = 7.4$ for all nodes. According to [15], the expected number of infected individuals converges towards zero exponentially fast if $\tau = \frac{\beta}{\delta} < \frac{1}{\lambda_1(A)}$, where $\lambda_1(A)$ is the largest absolute eigenvalue of the adjacency matrix. In our case, the largest eigenvalue of the Zachary's network equals to $\lambda_1(A) = 6.7257$; hence, the condition $\tau < \frac{1}{\lambda_1(A)}$ is satisfied. As illustrated in Fig. 4, the empirical average of number of infected nodes decreases exponentially over time. In Fig. 5, we plot the evolution of the mean SIS dynamics of each node in the Zachary's network.Our simulations show the validity of the bounds obtained by our moment-closure framework.

B. Moment-Closure of the SIR Epidemic Process

We proceed to demonstrate our SDP-based moment-closure scheme on the SIR model. In these experiments, we use again Zachary's network. In our simulations, we have selected the following set of initially infected nodes: $\mathcal{D} = \{5, 22, 28, 31, 32\}$; all remaining nodes are initially in the susceptible state. We set the infection rates to be $\beta_{ij} = \beta = 10$, whereas the recovery rates are $\delta_i = \delta = 6.7257$ for all nodes. Due to space limitations, we show in Fig. 6 the evolution of $\{\hat{\mu}_{i,S}(t), \hat{\mu}_{i,I}(t), \hat{\mu}_{i,R}(t)\}$ and $\{\check{\mu}_{i,S}(t), \check{\mu}_{i,I}(t), \check{\mu}_{i,R}(t)\}$ for the nodes in the subset $\{2, 7, 22, 29\}$. Notice that the proposed moment-closure technique does indeed upper and lower bounds the true mean dynamics of the SIR model. Nonetheless, the performance of these bounds varies. For example, in Fig. 6-(c), both bounds remain close to the true mean dynamics. However, in Fig. 6-(a), the upper estimate $\hat{\mu}_{2,I}$ fails to keep track of the true evolution of $\mu_{2,I}(t)$. There are several possible reasons for this to happen. For example, as shown in (37), at every time instance, we have $\mu_{i,S}(t) + \mu_{i,I}(t) + \mu_{i,R}(t) = 1$; however, the proposed upper and lower estimates fail to preserve this property.

C. First-Order and Second-Order Moment-Closure of the SIS Epidemic Process

In this subsection, we run the stochastic SIS dynamics over a directed graph with six nodes, whose topology is depicted in Fig. 7. In our experiments, we choose node 1 and node 4 to be the infected nodes at time $t = 0$, as colored in red in Fig. 7. The infection rates satisfy $\beta_{ij} = \beta = 1$ for all $(i, j) \in \mathcal{E}$ and the recovery rates are $\delta_i = \delta = 7.4$ for all nodes. Similar to Section V-A, we choose these two parameters such that $\tau = \frac{\beta}{\delta} < \frac{1}{\lambda_1(A)}$; hence the expected number of infected individuals converges towards zero exponentially fast. In Fig. 8, we plot the evolution of the mean SIS dynamics of all six nodes in the graph. Furthermore, we plot the evolution of upper and lower bounds using first- and second-order moment-closure framework, as described in Theorem 5. From this experiment, we observe that second-order moment closure outperforms first-order moment closure on certain nodes of the network. However, such a conclusion is non-general.

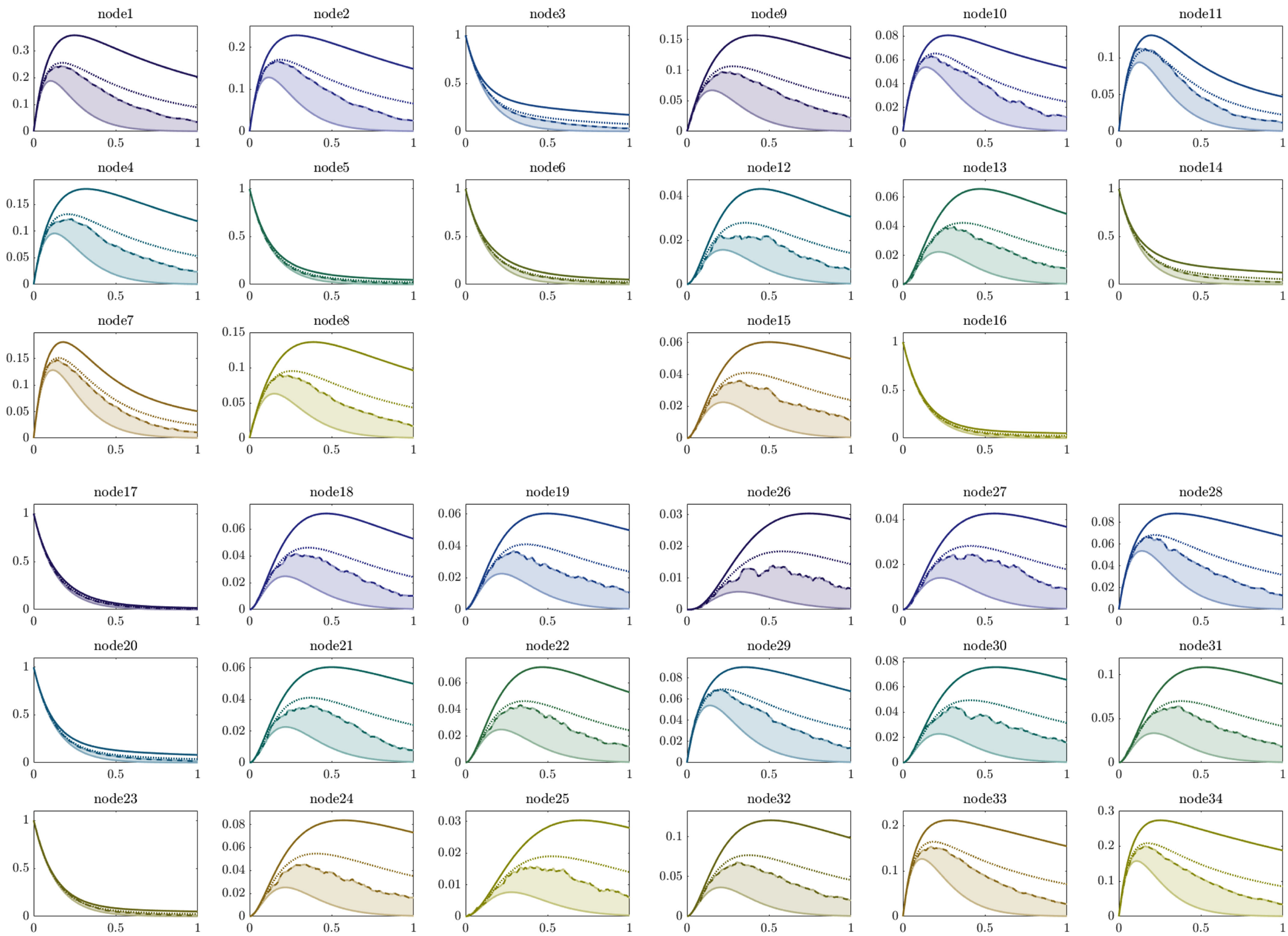


Fig. 5. Dashed lines represent the empirical averages of 10,000 realizations of the stochastic SIS dynamics for each node i. The dotted lines represent the trajectories obtained from the mean-field approximation for each node. The solid lines represent $\hat{\mu}_i(t)$ and $\check{\mu}_i(t)$ for each node i. Finally, the shaded areas are filling the gap between the empirical average and $\check{\mu}_i(t)$ for each node i.

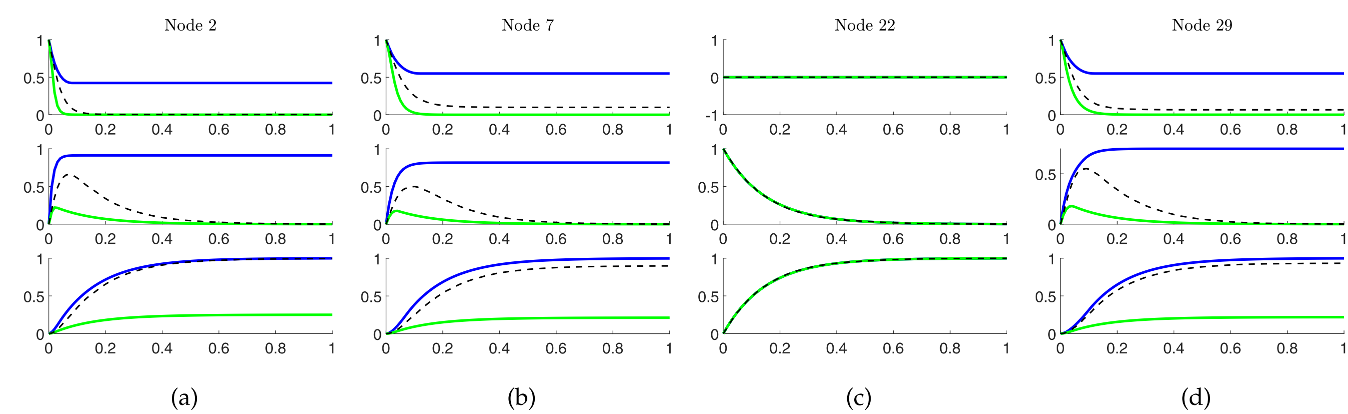


Fig. 6. Dashed lines represent the average of 10,000 realizations of stochastic SIR dynamics for each node i. In subfigures (a)– (d), we show the evolution of $\hat{\mu}_{i,C}(t)$ and $\check{\mu}_{i,C}(t)$, where $C \in \{S, I, R\}$, for the nodes $i = 2, 7, 22, 29$, respectively. For instance, in (a), blue and green lines in each of the subplots (from up to down) show the evolution of $\{\hat{\mu}_{2,S}(t), \check{\mu}_{2,S}(t)\}$, $\{\hat{\mu}_{2,I}(t), \check{\mu}_{2,I}(t)\}$, and $\{\hat{\mu}_{2,R}(t), \check{\mu}_{2,R}(t)\}$, respectively.

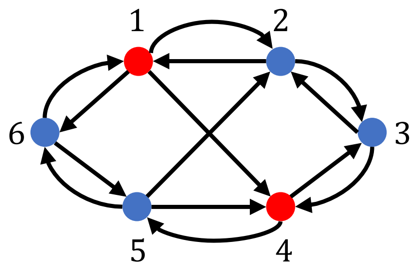


Fig. 7. Topology of directed graph with six nodes used in the third set off simulations in Section V-C. Nodes that are in infected state at $t = 0$ is colored in red, whereas nodes that are in susceptible state are colored in blue.

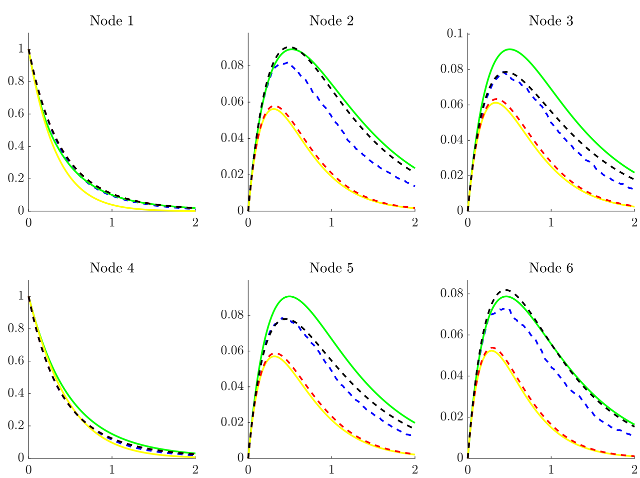


Fig. 8. The dashed black and red lines represent the upper and lower estimates $\hat{\mu}_i(t)$ and $\check{\mu}_i(t)$ for each node i obtained using second-order moment closure (i.e., $k = 2$ in Theorem 5), whereas the other dashed line in the middle represent the empirical averages of 20,000 realizations of the stochastic SIS dynamics for each node i. The solid (green and yellow) lines represent $\hat{\mu}_i(t)$ and $\check{\mu}_i(t)$ for each node i solved using Theorem 6.

VI. Conclusion

In this paper, we have analyzed the (exact) stochastic dynamics of the networked SIS, SI, and SIR epidemic models with heterogeneous spreading and recovery rates. The analysis of these models are, in general, very challenging since their state space grows exponentially with the number of nodes in the network. A common approach to overcome this challenge is to apply moment-closure techniques to approximate the exact stochastic dynamics via ordinary differential equations. However, most existing moment-closure techniques do not provide quantitative guarantees on the quality of the approximation, limiting the applicability of these techniques. To overcome this limitation, we have proposed a novel moment-closure framework which allows us to derive explicit quality guarantees. This framework is based on recent results from real algebraic geometry relating the multidimensional moment problem with semidefinite programming. We have illustrate how this technique can be used to derive upper and lower bounds on the exact (stochastic) dynamics of the SIS, SI, and SIR models. Moreover, we have provided a simplified version of our moment-closure technique to approximate the mean dynamics of the SIS model using a linear number of piecewise-affine differential equations. Finally, we have illustrated the validity of our results via numerical simulations in the Zachary's Karate Club network as well as an artificial digraph with six nodes.

Appendix

In this appendix, we provide proofs for the lemmas and theorems in this paper. In particular, our central idea of proving Lemma 9, Theorem 3 and Theorem 7 relies on carefully applying Theorem 2 to the jump-processes defined by the (stochastic) SIS and SIR processes, respectively. We then simplify the expressions by using the fact that the random variables representing the states of nodes in the network are binary.

Lemma 9: Given a directed graph $G = (\mathcal{V}, \mathcal{E})$, $i_1, \ldots, i_k \in [n]$, and $\boldsymbol{\alpha} \in \mathbb{N}^k$. If

$$dx_i = -x_i dP_{\delta_i} + (1 - x_i) \sum_{j \in \mathcal{N}_i^-} x_j dP_{\beta_{ij}}, \tag{45}$$

for all $i \in \mathcal{V}$, then

$$\frac{d\mathbb{E}[\phi_{\boldsymbol{\alpha}}(\mathbf{x})]}{dt} = \frac{d\mathbb{E}[\phi_{\mathbf{1}}(\mathbf{x})]}{dt}. \tag{46}$$

Proof of Lemma 9: To show (46), we first write (45) in the form of (10). Notice that there are $|\mathcal{V}| + |\mathcal{E}|$ Poisson counters in total, thus we define $h_\ell : \mathbb{R}^n \to \mathbb{R}^n$, for $\ell \in [|\mathcal{V}| + |\mathcal{E}|]$. Each h_ℓ is defined as follows: (i) when $\ell \in [n]$, we let $h_\ell(\mathbf{x}) = [0, \ldots, -x_\ell, \ldots, 0]^\top$, and (ii) when $\ell > n$, we order the edges $(j, i) \in \mathcal{E}$ and assign then with a label ℓ; hence, $h_\ell(\mathbf{x}) = [0, \ldots, (1 - x_i)x_j, \ldots, 0]^\top$, i.e., each $(j, i) \in \mathcal{E}$ is associated with a function h_ℓ.

With these definitions, it follows from (11) that

$$d\phi_{\mathbf{1}}(\mathbf{x}) = -\Pi_{s=1}^k x_{i_s} \left(\sum_{s=1}^k dP_{\delta_{i_s}} \right) + \sum_{s=1}^k \sum_{\ell \in \mathcal{N}_s^-} x_{i_1} \cdot\cdot(\;\; 1 - x_{i_s}) x_\ell \cdots x_{i_k} dP_{\beta_{s\ell}}. \tag{47}$$

Notice that the random variables x_i's are supported on $[0, 1]$, for all $i \in [n]$, therefore $\mathbb{E}[\phi_{\boldsymbol{\alpha}}(\mathbf{x})]$ exists and is finite for every $\boldsymbol{\alpha} \in \mathbb{N}^n$. Subsequently, from (47), we have that:

$$\frac{d\mathbb{E}[\phi_{\mathbf{1}}(\mathbf{x})]}{dt} = -\sum_{s=1}^k \delta_{i_\ell} \mathbb{E}[\phi_{\mathbf{1}}(\mathbf{x})] + \sum_{s=1}^k \sum_{\ell \in \mathcal{N}_{i_s}^-} \beta_{i_s \ell} \mathbb{E}[x_{i_1} \cdot\cdot(\;\; 1 - x_{i_s}) x_\ell \cdots x_{i_k}]. \tag{48}$$

On the other hand,

$$d\phi_{\boldsymbol{\alpha}}(\mathbf{x}) = -\phi_{\boldsymbol{\alpha}}(\mathbf{x}) \left(\sum_{s=1}^k dP_{\delta_{i_s}} \right) + \sum_{s \in [k]} \sum_{\ell \in \mathcal{N}_s^-} x_{i_1}^{\alpha_1} \cdots \left((x_{i_s} + (1 - x_{i_s}) x_\ell)^{\alpha_s} - x_{i_s}^{\alpha_s} \right) \cdots x_{i_k}^{\alpha_k} dP_{\beta_{s\ell}}. \tag{49}$$

Since $x_i \in \{0,1\}$ for all $i \in [n]$, the term $(x_{i_s} + (1-x_{i_s})x_\ell)^{\alpha_s} - x_{i_s}^{\alpha_s}$ equals to

$$\sum_{\kappa=0}^{\alpha_s-1} \binom{\alpha_s}{\kappa} x_{i_s}^{\kappa} ((1-x_{i_s})x_\ell)^{\alpha_s-\kappa}.$$

However, since $x_{i_s} \in \{0,1\}$, the above term can be further simplified to $((1-x_{i_s})x_\ell)^{\alpha_s}$. Therefore, taking expectation of (49) leads to:

$$\frac{d\mathbb{E}[\phi_{\boldsymbol{\alpha}}(\mathbf{x})]}{dt} = -\sum_{\ell=1}^{k} \delta_{i_\ell} \mathbb{E}[\phi_{\boldsymbol{\alpha}}(\mathbf{x})] + \sum_{s=1}^{k} \sum_{\ell \in \mathcal{N}_{i_s}^-} \beta_{i_s\ell} \mathbb{E}[x_{i_1}^{\alpha_1} \cdots (1-x_{i_s})^{\alpha_s} x_\ell^{\alpha_s} \cdots x_{i_k}^{\alpha_k}], \quad (50)$$

where the second equality is due to x_i are binary random variables. The proof finishes by noticing that the right-hand-side of the above equation is equal to $\frac{d\mathbb{E}[\phi_{\mathbf{1}}(\mathbf{x})]}{dt}$. ■

With the above lemma, we proceed to prove Theorem 3.

Proof of Theorem 3: From Lemma 9, we have that

$$\frac{d\mathbb{E}[\phi_{\boldsymbol{\alpha}}(\mathbf{x})]}{dt} = -\sum_{\ell=1}^{k} \delta_{i_\ell} \mathbb{E}[\phi_{\boldsymbol{\alpha}}(\mathbf{x})] + \sum_{s=1}^{k} \sum_{\ell \in \mathcal{N}_{i_s}^-} \beta_{i_s\ell} \mathbb{E}[x_{i_1}^{\alpha_1} \cdots (1-x_{i_s})^{\alpha_s} x_\ell^{\alpha_s} \cdots x_{i_k}^{\alpha_k}]. \quad (51)$$

Meanwhile, $\frac{d\mathbb{E}[\phi_{\boldsymbol{\alpha}}(\mathbf{x})]}{dt} = \frac{d\mathbb{E}[\phi_{\mathbf{1}}(\mathbf{x})]}{dt}$ holds for all $\boldsymbol{\alpha}$, thus rearranging the term $\mathbb{E}[x_{i_1} \cdots (1-x_{i_s})x_\ell \cdots x_{i_k}]$ leads us to (13). ■

In order to show Lemma 4, we introduce the following lemma, which establishes an equivalence relationship on the positive-semidefinteness of two matrices provided that one can be constructed from another in a special way.

Lemma 10: Consider a matrix $A = [a_{ij}] \in \mathbb{R}^{n\times n}$, a sequence of integers $d_1, \ldots, d_n \in \mathbb{N}$, and a mapping $f : \mathbb{R}^{n\times n} \to \mathbb{R}^{\sum_{i=1}^n d_i \times \sum_{i=1}^n d_i}$ defined as

$$f(A) = \begin{bmatrix} a_{11}J_{d_1,d_1} & a_{11}J_{d_1,d_2} & \cdots & a_{1n}J_{d_1,d_n} \\ * & a_{22}J_{d_2,d_2} & \cdots & a_{2n}J_{d_2,d_n} \\ \vdots & \vdots & \ddots & \vdots \\ * & * & * & a_{nn}J_{d_n,d_n} \end{bmatrix},$$

where J_{pq} is the $p \times q$ matrix of all ones. Then, if $A \succeq 0$, we have that $f(A) \succeq 0$. ⋄

Proof of Lemma 10: To proof the Lemma, let us define $T_{d_1,\ldots,d_n} \in \mathbb{R}^{\sum_{i=1}^n d_i \times n}$ as

$$T_{d_1,\ldots,d_n} = \begin{bmatrix} \mathbf{1}_{d_1} & \cdots & 0 \\ \vdots & \ddots & \vdots \\ * & \cdots & \mathbf{1}_{d_n} \end{bmatrix},$$

i.e., a block-diagonal matrix with its diagonal blocks specified by $\mathbf{1}_{d_1}, \cdots, \mathbf{1}_{d_n}$.

Next, we notices that given a matrix $A = [a_{ij}] \in \mathbb{R}^{n\times n}$,

$$f(A) = \begin{bmatrix} a_{11}J_{d_1,d_1} & a_{11}J_{d_1,d_2} & \cdots & a_{1n}J_{d_1,d_n} \\ * & a_{22}J_{d_2,d_2} & \cdots & a_{2n}J_{d_2,d_n} \\ \vdots & \vdots & \ddots & \vdots \\ * & * & * & a_{nn}J_{d_n,d_n} \end{bmatrix} = T_{d_1,\ldots,d_n} A T_{d_1,\ldots,d_n}^\top.$$

Consequently, $A \succeq 0$ implies that $f(A) = T_{d_1,\ldots,d_n} AT_{d_1,\ldots,d_n}^\top \succeq 0$.

Suppose that there exists $\mathbf{v}$ such that $\mathbf{v}^\top A\mathbf{v} < 0$, then we construct $\mathbf{w} \in \mathbb{R}^{\sum_{i=1}^n d_i}$ as follows iteratively. The first d_1 entries of $\mathbf{w}$ are all equals to $\mathbf{v}_1/d_1$, the i-th d_i entries of $\mathbf{w}$ are all equals to $\mathbf{v}_i/d_i$. Thus, $T_{d_1,\ldots,d_n}^\top \mathbf{w} = \mathbf{v}$. Thus, $f(A)$ is negative definite. Consequently, we have shown that $A \succeq 0$ if and only if $f(A) \succeq 0$. ■

With the help of Lemma 10, we are able to prove Lemma 4. The main idea of proving Lemma 4 lies in constructing matrices that exploits the special structures of moment and localizing matrices, in the form described by Lemma 10. This is possible since the random variables $\{x_i(t) : i \in \mathcal{V}\}$ describing the states of nodes i are binary.

Proof of Lemma 4: Notice that $\tilde{S}_k = [0,1]^k$ is a compact, semi-algebraic set, and it satisfies the Putinar's condition, it follows that $\mathbf{y}_\infty(\mathcal{I}_k)$ is $\tilde{S}_k$-feasible if and only if the conditions in (8) are satisfied. Subsequently, it suffices to show that the matrices in (20) implies the positive semi-definiteness of the moment and localizing matrices specified according to Theorem 1 and vice versa.

Consider $r \in \mathbb{N}$ and $r \geq \bar{k}$, the construction of () together with the definition of $\mathbf{y}_\infty(\mathcal{I}_k)$ implies that $M_{\bar{k}}(\mathbf{y}(\mathcal{I}_k))$ is the $\bar{k}$-th order principal submatrix of $M_r(\mathbf{y}_\infty(\mathcal{I}_k))$. Since $M_r(\mathbf{y}_\infty(\mathcal{I}_k)) \succeq 0$, we have that $M_{\bar{k}}(\mathbf{y}(\mathcal{I}_k)) \succeq 0$ as well. Similarly, the positive semi-definiteness of localizing matrices of $\mathbf{y}_\infty(\mathcal{I}_k)$ implies that both $L_{\bar{k}}^1(\mathbf{y}(\mathcal{I}_k), s)$ and $L_{\bar{k}}^0(\mathbf{y}(\mathcal{I}_k), s)$ are positive semi-definite for all $s \in [k]$.

Conversely, if $M_{\bar{k}}(\mathbf{y}(\mathcal{I}_k))$ is positive semi-definite, we aim to show that $M_r(\mathbf{y}_\infty(\mathcal{I}_k))$ is also positive semi-definite for all $r \in \mathbb{N}$. Notice that it suffices to show the above relationship holds for $r > k$. To achieve this goal, we proceed by permuting the entries in $M_r(\mathbf{y}_\infty(\mathcal{I}_k))$. Without loss of generality, we assume that $\mathcal{I}_k = \{1, \ldots, k\}$. Given a a set $\mathcal{W} \subseteq \mathcal{I}_k$, we use $x_{\mathcal{W}} = \Pi_{i\in\mathcal{W}} x_i$. Next, we consider

$$\mathbf{v}_r(\mathbf{x}_{\mathcal{I}_k}) = [1, x_1, \ldots, x_k, x_1^2, \ldots x_k^2, \ldots, x_1^r, \ldots, x_k^r],$$

and

$$\mathbf{v}_r'(\mathbf{x}_{\mathcal{I}_k}) = \Big[1, \overbrace{x_1, \ldots, x_1^r,}^{\text{monomials involving only}} \quad x_1 \ldots, \quad \underbrace{x_{\mathcal{W}}, \ldots}_{\text{monomials involving only}} \quad x_w \text{ for } w \in \mathcal{W}, \ldots, x_{\mathcal{I}_k}\Big].$$

Thus, $\mathbf{v}_r'(\mathbf{x}_{\mathcal{I}_k})$ is a permutation on the entries in $\mathbf{v}_r(\mathbf{x}_{\mathcal{I}_k})$. Let $N_r = \binom{k+r}{r}$, and S_r be the permutation group on the set $[N_r]$, there exists $\pi : S_r \to S_r$ such that for all $i \in [N_r]$ we have $\mathbf{v}_r'(\mathbf{x}_{\mathcal{I}_k})_i = \pi(\mathbf{v}_r(\mathbf{x}_{\mathcal{I}_k})_j$ for some $j \in [N_r]$.

Consider the following N_r-dimensional matrix $\hat{M}_r$ whose entries are defined by:

$$[\hat{M}_r]_{\boldsymbol{\alpha},\boldsymbol{\beta}} = y_\infty(\mathcal{I}_k)_{\pi^{-1}(\boldsymbol{\alpha}+\boldsymbol{\beta})} \tag{52}$$

for all $\boldsymbol{\alpha}, \boldsymbol{\beta} \in \mathbb{N}^k_{r/2}$. Thus, there exists a permutation matrix $P \in \{0,1\}^{N_r \times N_r}$ such that $\hat{M}_r = PM_r(\mathbf{y}_\infty(\mathcal{I}_k))P^{-1}$. Moreover, $\hat{M}_r$ is in the following form:

$$\begin{bmatrix} 1 & \mu_1 \mathbf{1}^\top_{r/2} & \cdots & * \\ * & \mu_1 J_{r/2,r/2} & \cdots & \vdots \\ * & * & \ddots & \vdots \\ * & * & * & \mu_{\mathcal{I}_k} J_{q,q} \end{bmatrix}, \tag{53}$$

for some $q \in \mathbb{N}$. Similarly, we can permute the matrix $M_{\bar{k}}(\mathbf{y}(\mathcal{I}_k))$ into the above form. Thus, there exists M such that $T_{d_1,\ldots,d_{\bar{k}}} M T^\top_{d_1,\ldots,d_{\bar{k}}} = \hat{P} M_{\bar{k}}(\mathbf{y}(\mathcal{I}_k))\hat{P}^{-1}$ for some sequence of $d_1, \ldots, d_{\bar{k}}$. Furthermore, there exists another sequence $\{d_1, \ldots, d_{r/2}\}$ such that $T_{d_1,\ldots,d_{r/2}} M T^\top_{d_1,\ldots,d_{r/2}} = \hat{M}_r$. Consequently, applying Lemma V twice, we have that if $M_{\bar{k}}(\mathbf{y}(\mathcal{I}_k)) \succeq 0$ then $M_r(\mathbf{y}_\infty(\mathcal{I}_k)) \succeq 0$. The above claim holds for arbitrary r, thus the result follows. The relationship between finite and infinite dimensional localizing matrices can be shown using the above procedure. ∎

Proof of Theorem 5: To show the monotone relationship $\hat{\mu}_\mathcal{I}(t) \geq \mu_\mathcal{I}(t) \geq \check{\mu}_\mathcal{I}(t)$ holds for all $\mathcal{I}$ and $t \geq 0$, we apply the multi-variate comparison lemma, i.e., Theorem 1.2 from [30]. More specifically, we aim to show that when $\hat{\mu}_\mathcal{I} = \mu_\mathcal{I}$, $\dot{\hat{\mu}}_\mathcal{I} \geq \dot{\mu}_\mathcal{I}$ for all $\hat{\mu}_\mathcal{J} \geq \mu_\mathcal{J}, \mathcal{J} \neq \mathcal{I}$ and $\check{\mu}_\mathcal{J} \leq \mu_\mathcal{J}, \forall |\mathcal{J}| \leq k$.

On one hand, when $|\mathcal{I} \cup \{\ell\}\}| \leq k$, we have that all the terms with positive coefficients are bounded above by upper estimates $\hat{\mu}_{\mathcal{I}\cup\ell\setminus\{i_s\}}$, whereas the terms with negative coefficients are bounded below by $\check{\mu}_{\mathcal{I}\cup\{\ell\}}$; thence $\dot{\hat{\mu}}_\mathcal{I} \geq \dot{\mu}_\mathcal{I}$.

On the other hand, when $|\mathcal{I} \cup \{\ell\}\}| = k+1$, it suffices to show that $\mu_{\mathcal{I}\cup\{\ell\}}$ is feasible in the SDPs (27). Consider the random variable $\mathbf{x}_{\mathcal{I}\cup\{\ell\}} = [x_{i_1}, \ldots, x_\ell]^\top$, its underlying measure at time t is supported on $\tilde{S}_{|\mathcal{I}|+1}$, which is a compact and semi-algebraic set. Let $\mathbf{y}$ be the infinite multi-sequence consisting of all the moments of $\mathbf{x}_{\mathcal{I}\cup\{\ell\}}$. From $\mathbf{y}$, we can readily construct moment matrix $M_r(\mathbf{y})$, and localizing matrices $L_r(g_j\mathbf{y})$, for any given $r \in \mathbb{N}$. Consequently, according to Theorem 1, these matrices are positive semi-definite. Moreover, according to Lemma 4, the positive semi-definiteness of these matrices are equivalent to positive semi-definiteness of $M_{\bar{k}}(\mathbf{y}(\mathcal{I} \cup \{\ell\}))$, $L^1_{\bar{k}}(\mathbf{y}(\mathcal{I} \cup \{\ell\}), s)$, and $L^0_{\bar{k}}(\mathbf{y}(\mathcal{I} \cup \{\ell\}), s)$ for all $s \in [k]$. Consequently, $\{\mu_\mathcal{J}\}_{\mathcal{J}\subseteq\mathcal{I}\cup\{\ell\}}$ is a feasible solution to both 25 and (26). Meanwhile, the eigenvalues of (23) and (24) are monotonic in terms of their entries, which implies that $\{\mu_\mathcal{J}\}_{\mathcal{J}\subseteq\mathcal{I}\cup\{\ell\}}$ is also feasible with respect to both the minimization and maximization problems (27). Furthermore, this holds for all $\hat{\mu}_\mathcal{J} \geq \mu_\mathcal{J}, \mathcal{J} \neq \mathcal{I}$ and $\check{\mu}_\mathcal{J} \leq \mu_\mathcal{J}, \forall |\mathcal{J}| \leq k$. Summarizing the above claims, we have that if $\hat{\mu}_\mathcal{I}(0) \geq \mu_\mathcal{I}(0)$, then $\hat{\mu}_\mathcal{I}(t) \geq \mu_\mathcal{I}(t)$ holds for all $\mathcal{I}$ and $t \geq 0$. The above argument readily applies to the comparison between $\mu_\mathcal{I}(t)$ and $\check{\mu}_\mathcal{I}(t)$. ∎

Proof of Theorem 6: To show the results, we use the fact that a matrix is positive semi-definite, if and only if, all its principal minors are non-negative [31] and apply it to (28)–(31), respectively. In particular, in the view of Remark 5, it suffices for us to characterize the condition for positive semidefniteness of matrices (29)–(31). More specifically, we compute all principal minors of localizing matrices and require them to be non-negative. Therefore, from $L^0_1(\mathbf{y}(\mathcal{I}_2), i), L^0_1(\mathbf{y}(\mathcal{I}_2), j) \succeq 0$, we obtain $\mu_{ij} \leq \min\{\mu_i, \mu_j\}$. From $L^1_1(\mathbf{y}(\mathcal{I}_2), i), L^1_1(\mathbf{y}(\mathcal{I}_2), j) \succeq 0$, we have $\mu_{ij} \geq \mu_i + \mu_j - 1$. Therefore, $\{\mu_i, \mu_j, \mu_{ij}\}$ is a feasible moment sequence provided that all the above constraints on μ_{ij} are satisfied simultaneously. This is equivalent to $\mu_{ij} \in [l_{ij}, u_{ij}]$, where

$$u_{ij} = \min\{\mu_i, \mu_j\}, \tag{54}$$

and

$$l_{ij} = \max\{\mu_i + \mu_j - 1, 0\}. \tag{55}$$

Since $\beta_{ij} > 0$, maximizing over $-\sum_{j=1}^n \beta_{ij}\mu_{ij}$ is equivalent to minimize μ_{ij}. In particular, the minimum of μ_{ij} is attained at l_{ij}. Thus, the upper bound is obtained. Similarly, to obtain the lower bound, we maximize μ_{ij} for all $i, j \in [n]$. The optimal solution of these two problems are $\overline{\mu}^\star_{ij}$ and $\underline{\mu}^\star_{ij}$, respectively.

Next, we aim to show that $\hat{\mu}_i(t) \geq \mu_i(t) \geq \check{\mu}_i(t)$. To achieve this goal, we utilize Proposition 1.4 from [30]. It suffice to show that when $\mu_i(t) = \hat{\mu}_i(t)$, $\dot{\hat{\mu}}_i \geq \dot{\mu}_i$ for all $\hat{\mu}_j \geq \mu_j$. Consider the difference between $\dot{\hat{\mu}}_i$ and $\dot{\mu}_i$

$$\begin{aligned} \frac{d\hat{\mu}_i}{dt} - \frac{d\mu_i}{dt} &= \sum_j \beta_{ij}\big[\hat{\mu}_j - \mu_j + \mu_{ij} - \overline{\mu}_{ij}\big] \\ &= \sum_j \big[\hat{\mu}_j - \mu_j + \mu_{ij} - \max\{0, \hat{\mu}_j + \mu_i - 1\}\big] \end{aligned}$$

The above equality is due to the assumption that $\hat{\mu}_i = \mu_i$. Consider the following cases: (i) $\hat{\mu}_j = \mu_j$, then the right-hand-side is larger than zero according to (26); (ii) $\hat{\mu}_j > \mu_j$ and $\hat{\mu}_j + \mu_i - 1 \leq 0$, the RHS is non-negative trivially; and (iii) $\hat{\mu}_j > \mu_j$ and $\hat{\mu}_j + \mu_i - 1 > 0$, it follows that:

$$\begin{aligned} \frac{d\hat{\mu}_i}{dt} - \frac{d\mu_i}{dt} &= \sum_j \big[\hat{\mu}_j - \mu_j + \mu_{ij} - \hat{\mu}_j - \mu_i + 1\big] \\ &\geq \sum_j \big[\mu_{ij} - \max\{0, \mu_i + \mu_j - 1\}\big] \geq 0. \end{aligned}$$

As a consequence, $\hat{\mu}_i(t) \geq \mu_i(t)$ according to [30]. Similar analysis holds for comparing $\dot{\check{\mu}}_i$ and $\dot{\mu}_i$. ∎

Proof of Theorem 7: Similar to the treatment in Lemma V, we define $h_\ell : \mathbb{R}^{3n} \to \mathbb{R}^{3n}$, for $\ell \in [|\mathcal{V}| + |\mathcal{E}|]$. Each h_ℓ is defined as follows: (i) when $\ell \in [n]$, we let

$$h_\ell(\mathbf{x}) = [0, 0, 0, \ldots, 0, -x_{\ell,I}, x_{\ell,I}, \ldots, 0, 0, 0]^\top,$$

and (ii) for every $(i, j) \in \mathcal{E}$, we let

$$h_{(i,j)}(\mathbf{x}) = [0, 0, 0, \ldots, -x_{i,S}x_{j,I}, x_{i,S}x_{j,I}, \ldots, 0, 0, 0]^\top.$$

With these definitions, according to (11), when $\ell \in [n]$, we have that $\phi_{\boldsymbol{\alpha},\boldsymbol{\beta},\boldsymbol{\gamma}}(\mathbf{x} + h_\ell(\mathbf{x})) - \phi_{\boldsymbol{\alpha},\boldsymbol{\beta},\boldsymbol{\gamma}}(\mathbf{x})$ is equal to

$$\Pi_{k\in[n],k\neq\ell} x_{k,S}^{\alpha_k} x_{k,I}^{\beta_k} x_{k,R}^{\gamma_k} \left[x_{\ell,S}^{\alpha_\ell}(x_{\ell,R} + x_{\ell,I})^{\gamma_\ell} - x_{\ell,S}^{\alpha_\ell} x_{\ell,I}^{\beta_\ell} x_{\ell,R}^{\gamma_\ell} \right].$$

Consequently, when $\beta_\ell \neq 0$, the term above equals to $\phi_{\boldsymbol{\alpha},\boldsymbol{\beta},\boldsymbol{\gamma}}(\mathbf{x})$. Meanwhile, $x_{\ell,I}$ and $x_{\ell,R}$ are binary variables. Moreover, $x_{\ell,R} + x_{\ell,I} \leq 1$ for all $t \geq 0$. Thus, $(x_{\ell,I} + x_{\ell,R})^{\gamma_\ell} - x_{\ell,I}^{\gamma_\ell} = x_{\ell,I}$. Subsequently, we have

$$\phi_{\boldsymbol{\alpha},\boldsymbol{\beta},\boldsymbol{\gamma}}(\mathbf{x} + h_\ell(\mathbf{x})) - \phi_{\boldsymbol{\alpha},\boldsymbol{\beta},\boldsymbol{\gamma}}(\mathbf{x}) = \begin{cases} -\phi_{\boldsymbol{\alpha},\boldsymbol{\beta},\boldsymbol{\gamma}}(\mathbf{x}), \text{ if } \beta_\ell \neq 0, \\ \Pi_{k\in[n],k\neq\ell} x_{k,S}^{\alpha_k} x_{k,I}^{\beta_k} x_{k,R}^{\gamma_k} x_{\ell,S}^{\alpha_\ell} x_{\ell,I}, \text{ if } \beta_\ell = 0 \text{ and } \gamma_\ell \neq 0, \\ 0, \text{ otherwise.} \end{cases} \tag{56}$$

Similarly, for a given $(j, i) \in \mathcal{E}$, we have

$$\begin{aligned} &\phi_{\boldsymbol{\alpha},\boldsymbol{\beta},\boldsymbol{\gamma}}(\mathbf{x} + h_{(j,i)}(\mathbf{x})) - \phi_{\boldsymbol{\alpha},\boldsymbol{\beta},\boldsymbol{\gamma}}(\mathbf{x}) \\ &= \Pi_{k\in[n],k\neq i} x_{k,S}^{\alpha_k} x_{k,I}^{\beta_k} x_{k,R}^{\gamma_k} x_{i,S}^{\alpha_i} x_{i,R}^{\gamma_i} \\ &\times (\Big[1 - x_{j,I})^{\alpha_i} (x_{i,I} + x_{i,S}x_{j,I})^{\beta_i} - x_{i,I}^{\beta_i} \Big]. \end{aligned} \tag{57}$$

On one hand, when $\alpha_i \neq 0$, we observe that if $x_{j,I} = 0$, then $(1 - x_{j,I})^{\alpha_i}(x_{i,I} - x_{i,S}x_{j,I})^{\beta_i} - x_{i,I}^{\beta_i}$ equals to zero, whereas if $x_{j,I} = 1$, the term equals to $-x_{i,I}^{\beta_i}$.

On the other hand, when $\alpha_i = 0$ and $\beta_i = 0$, (18) equals to 0. Finally, we consider the case when $\alpha_i = 0$ and $\beta_i \neq 0$. In this context, it suffices to examine $(x_{i,I} + x_{i,S}x_{j,I})^{\beta_i} - x_{i,I}^{\beta_i}$. Notice that when $x_{i,S} = 0$ the sum equals to 0, and $x_{j,I}$ otherwise. Thus, it can be simplified into $x_{i,S}x_{j,I}$. Summarizing the above cases, let

$$Q = \phi_{\boldsymbol{\alpha},\boldsymbol{\beta},\boldsymbol{\gamma}} / x_{i,I}^{\beta_i}$$

we obtain that:

$$\phi_{\boldsymbol{\alpha},\boldsymbol{\beta},\boldsymbol{\gamma}}(\mathbf{x} + h_{(j,i)}(\mathbf{x})) - \phi_{\boldsymbol{\alpha},\boldsymbol{\beta},\boldsymbol{\gamma}}(\mathbf{x}) = \begin{cases} -\phi_{\boldsymbol{\alpha},\boldsymbol{\beta},\boldsymbol{\gamma}}(\mathbf{x})x_{j,I}, \text{ if } \alpha_i \neq 0, \\ Qx_{i,S}x_{j,I}, \text{ if } \alpha_i = 0 \text{ and } \beta_i \neq 0, \\ 0, \text{ otherwise.} \end{cases} \tag{58}$$

Finally, (39) is obtained by taking expectation on the sum of $|\mathcal{V}| + |\mathcal{E}|$ equations (56) and (58). ∎

References

[1] J. O. Kephart and S. R. White, "Directed-graph epidemiological models of computer viruses," in *Proc. IEEE Comput. Soc. Symp. Res. Secur. Privacy*, 1991, pp. 343–359.

[2] Z. Liu, Y. Lai, and N. Ye, "Propagation and immunization of infection on general networks with both homogeneous and heterogeneous components," *Phys. Rev. E*, vol. 67, no. 3, 2003, Art. no. 031911.

[3] A. Vespignani, "Modelling dynamical processes in complex socio-technical systems," *Nature Phys.*, vol. 8, no. 1, pp. 32–39, 2012.

[4] R. Pastor-Satorras, C. Castellano, P. V. Mieghem, and A. Vespignani, "Epidemic processes in complex networks," *Rev. Modern Phys.*, vol. 87, no. 3, p. 925–987, 2015.

[5] J. Ahn and B. Hassibi, "Global dynamics of epidemic spread over complex networks," in *Proc. IEEE Conf. Decis. Control*, 2013, pp. 4579–4585.

[6] C. Nowzari, V. M. Preciado, and G. J. Pappas, "Analysis and control of epidemics: A survey of spreading processes on complex networks," *IEEE Control Syst.*, vol. 36, no. 1, pp. 26–46, Feb. 2016.

[7] P. V. Mieghem, J. Omic, and R. Kooij, "Virus spread in networks," *IEEE/ACM Trans. Netw.*, vol. 17, no. 1, pp. 1–14, Feb. 2009.

[8] E. Volz and L. A. Meyers, "Susceptible–infected–recovered epidemics in dynamic contact networks," *Proc. Roy. Soc. London B: Biological Sci.*, vol. 274, no. 1628, pp. 2925–2934, 2007.

[9] F. D. Sahneh and C. Scoglio, "Epidemic spread in human networks," in *Proc. IEEE Conf. Decis. Control Eur. Control Conf.*, 2011, pp. 3008–3013.

[10] C. Nowzari, V. M. Preciado, and G. J. Pappas, "Stability analysis of generalized epidemic models over directed networks," in *Proc. IEEE Conf. Decision Control*, 2014, pp. 6197–6202.

[11] V. M. Preciado, M. Zargham, C. Enyioha, A. Jadbabaie, and G. J. Pappas, "Optimal vaccine allocation to control epidemic outbreaks in arbitrary networks," in *Proc. IEEE Conf. Decision Control*, 2013, pp. 7486–7491.

[12] V. M. Preciado, M. Zargham, C. Enyioha, A. Jadbabaie, and G. J. Pappas, "Optimal resource allocation for network protection against spreading processes," *IEEE Trans. Control Netw. Syst.*, vol. 1, no. 1, pp. 99–108, Mar. 2014.

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

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

[15] Y. Wang, D. Chakrabarti, C. Wang, and C. Faloutsos, "Epidemic spreading in real networks: An eigenvalue viewpoint," in *Proc. Int. Symp. Reliable Distrib. Syst.*, 2003, pp. 25–34.

[16] E. Cator and P. V. Mieghem, "Second-order mean-field susceptible-infected-susceptible epidemic threshold," *Phys. Rev. E*, vol. 85, no. 5, 2012, Art. no. 056111.

[17] F. D. Sahneh, C. Scoglio, and P. V. Mieghem, "Generalized epidemic mean-field model for spreading processes over multilayer complex networks," *IEEE/ACM Trans. Netw.*, vol. 21, no. 5, pp. 1609–1620, Oct. 2013.

[18] A. Lamperski, K. R. Ghusinga, and A. Singh, "Analysis and control of stochastic systems using semidefinite programming over moments," *IEEE Trans. Autom. Control*, vol. 64, no. 4, pp. 1726–1731, 2018.

[19] N. J. Watkins, C. Nowzari, and G. J. Pappas, "A robust moment closure for general continuous-time epidemic processes," in *Proc. IEEE Conf. Decision sControl*. IEEE, 2018, pp. 244–249.

[20] A. S. Mata and S. C. Ferreira, "Pair quenched mean-field theory for the susceptible-infected-susceptible model on complex networks," *Europhysics Lett.*, vol. 103, no. 4, 2013, Art. no. 48003.

[21] X. Chen, M. Ogura, K. R. Ghusinga, A. Singh, and V. M. Preciado, "Semidefinite bounds for moment dynamics: Application to epidemics on networks," in *Proc. IEEE Conf. Decis. Control*, 2017, pp. 2448–2454.

[22] J. B. Lasserre, *Moments, Positive Polynomials and Their Applications*, vol. 1. Singapore: World Scientific, 2009.

[23] E. Cator and P. V. Mieghem, "Nodal infection in markovian susceptible-infected-susceptible and susceptible-infected-removed epidemics on networks are non-negatively correlated," *Phys. Rev. E*, vol. 89, no. 5, 2014, Art. no. 052802.

[24] W. Feller, *An Introduction to Probability Theory and Its Applications*, vol. 2. Hoboken, NJ, USA: Wiley, 2008.

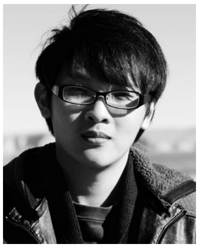
[25] M. Putinar, "Positive polynomials on compact semi-algebraic sets," *Indiana Univ. Math. J.*, vol. 42, no. 3, pp. 969–984, 1993.

[26] F. B. Hanson, *Applied Stochastic Processes and Control for Jump-Diffusions: Modeling, Analysis, and Computation*, vol. 13. Philadelphia, PA, USA: SIAM, 2007.

[27] M. W. Hirsch and H. Smith, "Monotone dynamical systems," in *Handbook Differential Equations: Ordinary Differential Equations*, vol. 2. New York, NY, USA: Elsevier, 2006, pp. 239–357.

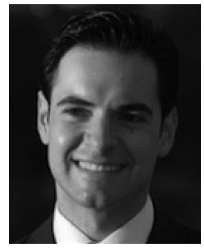
[28] R. A. Horn and C. B. Johnson, *Matrix Analysis*. Cambridge, U.K.: Cambridge Univ. Press, 2012.

[29] W. W. Zachary, "An information flow model for conflict and fission in small groups," *J. Anthropological Res.*, vol. 33, no. 4, pp. 452–473, 1977.



Ximing Chen received his Ph.D. in the electrical and systems engineering program at the University of Pennsylvania in 2019. He obtained his bachelor's degree in electronic engineering with honors research option from the Hong Kong University of Science and Technology in 2013. His research interests include control and optimization of structural systems and dynamic processes in complex networks using graph and measure-theoretic approaches.

Masaki Ogura is an Associate Professor in the Graduate School of Information Science and Technology at the Osaka University, Japan. He received his M.Sc. degree in informatics from Kyoto University in 2009, and his Ph.D. in mathematics from Texas Tech University in 2014. He was a Postdoctoral Researcher at the University of Pennsylvania from 2014 to 2017, and an Assistant Professor at the Nara Institute of Science and Technology, Japan from 2017 to 2019. His research interests include network science, dynamical systems, stochastic processes with applications in networked epidemiology, design engineering, and biological physics. He was a runner-up of the 2019 Best Paper Award by the IEEE TRANSACTIONS ON NETWORK SCIENCE AND ENGINEERING and a recipient of the 2012 SICE Best Paper Award.



Victor M. Preciado received the Ph.D. degree in electrical engineering and computer science from the Massachusetts Institute of Technology, in 2008. He is currently an Associate Professor of electrical and systems engineering at the University of Pennsylvania, where he is a member of the Networked and Social Systems Engineering (NETS) program, the Warren Center for Network and Data Sciences, and the Applied Math and Computational Science (AMCS) program. He was a recipient of the 2017 National Science Foundation CAREER Award, the 2018 Outstanding Paper Award by the *IEEE Control Systems Magazine*, and a runner-up of the 2019 Best Paper Award by the IEEE TRANSACTIONS ON NETWORK SCIENCE AND ENGINEERING. His main research interests lie at the intersection of big data and network science; in particular, in using innovative mathematical and computational approaches to capture the essence of complex, high-dimensional dynamical systems. Relevant applications of this line of research can be found in the context of socio-technical networks, brain dynamical networks, healthcare operations, biological systems, and critical technological infrastructure. He is a Senior Member of IEEE and Associate Editor of the IEEE TRANSACTIONS ON NETWORK SCIENCE AND ENGINEERING.