IMA Journal of Complex Networks (2017) Page 1 of 31
doi:10.1093/comnet/xxx000

An analytical framework for the study of epidemic models on activity driven networks

LORENZO ZINO

"Lagrange" Department of Mathematical Sciences, Politecnico di Torino, Corso Duca degli Abruzzi 24, 10129 Turin, Italy and "Peano" Department of Mathematics, University of Turin, Via Carlo Alberto 10, 10123 Turin, Italy.
lorenzo.zino@unito.it

ALESSANDRO RIZZO†

Department of Electronics and Telecommunications, Politecnico di Torino, Corso Duca degli Abruzzi 24, 10129 Turin, Italy and Office of Innovation, New York University Tandon School of Engineering, Brooklyn, New York 11201, USA.

AND

MAURIZIO PORFIRI†

Department of Mechanical and Aerospace Engineering, New York University Tandon School of Engineering, Brooklyn, New York 11201, USA.
†Corresponding authors. Emails: alessandro.rizzo@polito.it, mporfiri@nyu.edu

[Received on 26 September 2017]

Network theory has greatly contributed to an improved understanding of epidemic processes, offering an empowering framework for the analysis of real-world data, prediction of disease outbreaks, and formulation of containment strategies. However, the current state of knowledge largely relies on time-invariant networks, which are not adequate to capture several key features of a number of infectious diseases. Activity driven networks constitute a promising modeling framework to describe epidemic spreading over time varying networks, but a number of technical and theoretical gaps remain open. Here, we lay the foundations for a novel theory to model general epidemic spreading processes over time-varying, activity driven networks. Our theory derives a continuous-time model, based on ordinary differential equations (ODEs), which can reproduce the dynamics of any discrete-time epidemic model evolving over an activity driven network. A rigorous, formal framework is developed, so that a general epidemic process can be systematically mapped, at first, on a Markov jump process, and then, in the thermodynamic limit, on a system of ODEs. The obtained ODEs can be integrated to simulate the system dynamics, instead of using computationally intensive Monte Carlo simulations. An array of mathematical tools for the analysis of the proposed model is offered, together with techniques to approximate and predict the dynamics of the epidemic spreading, from its inception to the endemic equilibrium. The theoretical framework is illustrated step-by-step through the analysis of a susceptible-infected-susceptible process. Once the framework is established, applications to more complex epidemic models are presented, along with numerical results that corroborate the validity of our approach. Our framework is expected to find application in the study of a number of critical phenomena, including behavioral changes due to the infection, unconscious spread of the disease by exposed individuals, or the removal of nodes from the network of contacts.

Keywords: data-driven predictions; epidemic curve; Markov process; ordinary differential inclusions; temporal; time-varying.

1. Introduction

Time-varying interactions are ubiquitous in natural and technological networks associated with the interactions of dynamic agents [1–7]. In our hyper-connected and fast-changing world, the time-varying nature of contacts between individuals plays a fundamental role on the spreading of epidemics [2, 8, 9]. The mathematical analysis of time-varying networks has contributed to an improved understanding of epidemic spreading, far beyond the limitations imposed by time-invariant networks of contacts [1, 10–17].

In the last few years, discrete-time activity driven networks (ADNs) have emerged as a powerful paradigm to study epidemic spreading over realistic time-varying networks. ADNs have been successfully proposed to study time-varying networks, overcoming simplifying assumptions on the separation of the time scales of the nodes' and links' dynamics, which are typical of time-invariant models. This has enabled the formalization of more realistic models where the nodes' and links' dynamics can evolve concurrently [2]. ADNs can account for inherent heterogeneities in populations [2, 18–20], such that individuals could differ in their ability to form contacts and spread the disease. Initially formulated for modeling simple epidemic models [2, 21], such as the susceptible-infected-susceptible (SIS) and the susceptible-infected-recovered (SIR) [22], ADNs have then supported the development of realistic epidemic models [3, 8, 23–28]. Models have been progressively enriched by including the effect of timely sanitary intervention, for which simulation-based analyses have been carried out [21, 24]. ADNs-based models have also been adopted to study real outbreaks, yielding satisfactory predictions, such as in the case of the 2014-2015 outbreak of Ebola Virus Disease in Liberia [24].

The overarching idea of ADNs lies in the definition of an activity potential associated with each individual, sampled from a continuous distribution and used to represent the propensity of individuals to establish contacts with others. However, many simplifying assumptions have been retained in the very first incarnations of ADNs. A memoryless feature in the formation of links, both spatially and temporally, was assumed. In contrast, real networks of contacts often present a complex structure, induced by memory phenomena and geographic locality [29], which manifest through strong and weak ties [30]. Heterogeneity is indeed present not only in the individual activity, but also in the attractiveness [31]. The ADN paradigm has been rapidly extended over the last few years, and we have now access to a wide range of modeling tools for studying them with: i) links that are formed with memory [5, 32], ii) community structures [33, 34], and iii) heterogeneity in the individual attractiveness [35, 36]. Despite these compelling improvements in the modeling of realistic outbreaks on complex networks, most of the studies are based on extensive Monte Carlo simulations, and analytical results are only limited to linearized mean-field approximations and asymptotic master equations for basic epidemic models [37].

In our recent work [38], we formulated a continuous-time ADN model with a discrete activity potential distribution to afford new analytical insight into epidemic spreading. Changing the classical perspective of a discrete-time dynamics for continuous distributions to a continuous-time dynamics for discrete distributions, we were successful in linking ADN modeling to established mathematical techniques for the study of complex systems. Our method could be interpreted as a reversed procedure of the Gillespie algorithm [39, 40], which allows for simulating trajectories of stochastic equations of complex systems in the presence of multiple reactions.

For the SIS model, this change of perspective leads to a rigorous analytical treatment of the epidemics, which affords the accurate prediction of the epidemic outbreak. Moreover, our approach is not affected by the issues that naturally arise when selecting the time step of discrete-time dynamical processes, which could influence the system dynamics [41, 42], and it relies on a reduced number of parameters with respect to traditional ADNs [2, 3, 8, 21, 23–27, 34]. The dependence on a limited

number of model parameters is critical for robust model identification from real-world data [24, 43–46].

Motivated by the potential of this novel perspective for the study of ADNs, here, we seek to generalize the framework to more realistic epidemic models, through the development of a new theory for the analytical study of the spread of epidemics on heterogeneous time-varying networks. Our approach has the potential of improving our comprehension of the mechanics of epidemic spreading, toward new techniques to control and stop the spread [47–50]. The generality of our framework allows for the inclusion of key real-world features that have been recently proposed for ADNs [5, 32–36], thereby offering a comprehensive and versatile basis for analytically exploring new modeling concepts.

The main contributions of this paper are: i) the formulation of a general stochastic epidemic model on ADNs as a continuous-time dynamical system; ii) the derivation of a deterministic approximation valid in the thermodynamic limit of large populations; and iii) the development of an array of empowering mathematical techniques for the analysis of such models.

In the general formulation of the epidemic model, we include several realistic features from the literature, such as the presence of different health states [51], the possibility of immunization [22], and behavioral changes due to infection [8]. Our formulation consists of a Markov jump process, constructed according the temporal evolution of the ADN and the specific dynamics of the epidemics. Then, using Kurtz theorem [52], we derive a nonlinear system of ordinary differential equations (ODEs) that approximates the stochastic epidemic model in the thermodynamic limit. With respect to mathematical techniques, we demonstrate the possibility of using ordinary differential inclusions (ODIs) to i) estimate the fraction of individuals per health states when the epidemic becomes endemic, and ii) accurately predict the evolution of the epidemic curve. Also, we propose a data-driven approach to generate very accurate short- and medium-time-horizon predictions, using few epidemics data that are sporadically sampled at the population level.

The rest of the paper is organized as follows. In Section 2, we introduce the general formulation of a stochastic epidemic model on ADNs and we compute its large-scale deterministic approximation. In Section 3, we establish a range of techniques to practically tackle the analysis of the obtained epidemic models. Section 4 is devoted to the application of these techniques to well known epidemic models, exemplifying a number of realistic features of epidemic spreading. Finally, Section 5 summarizes the main conclusions of this work.

2. Deterministic approximation in large-scale populations

Here, we introduce our general formulation of an epidemic model on ADNs and derive its deterministic approximation in the thermodynamic limit. To facilitate the comprehension of our approach we provide, step-by-step, its illustration on the simple SIS model.

2.1 *From the epidemic dynamics to the stochastic process*

We consider a population of n individuals, each one associated with a node of a time-varying, undirected graph $\mathcal{G}(t) = (\mathcal{V}, \mathcal{E}(t))$, with $t \in \mathbb{R}_+$, where $\mathbb{R}_+$ is the set of non-negative real numbers. $\mathcal{V} = \{1, \dots, n\}$ is the time-invariant node set, and $\mathcal{E}(t) \subseteq \mathcal{V} \times \mathcal{V}$ is the time-varying edge set associated with the network of contacts.

Let $\mathcal{A}$ be the set of all compartments of the model, defined as all the admissible health states for the individuals in the population. We denote with $Y_v(t) \in \mathcal{A}$ the state of node $v \in \mathcal{V}$ at time t. In the simplest case, $\mathcal{A}$ comprises only two states, denoting whether the node is susceptible to the infection or infected. This is the case of the susceptible-infected (SI) or SIS models [12]. More realistic models include the

presence of other compartments to identify individuals who already recovered from the infection and are now immunized [22], or include different stages of the disease in an infected individual [51]. We denote compartments with italic capital letters.

Nodes are divided into k activation classes, depending on their propensity to generate connections. Each class $i \in \mathcal{K} := \{1,\dots,k\}$ is characterized by a nominal activity rate a_i, where $a_1 < \cdots < a_k$ are equidistant non-negative real numbers, representing the nominal (average) number of contacts an individual in the ith class generates during each time unit, which can be, for example, one day, or one hour, depending on the unit of measurement used. We use the notation $a(v)$, $v \in \mathcal{V}$, to identify the nominal activity rate of the vth node. The quantity n_i denotes the number of nodes in the ith class. Real-world data suggest that the actual activity rate of an individual can be reduced based on his/her health state [8, 53, 54]. Therefore, we introduce an $|\mathcal{A}|$-dimensional vector $\rho \in [0,1]^{|\mathcal{A}|}$, whose A-th component ρ_A represents the fraction of the reduced activity rate for an individual who has health state A (if no reduction occurs for the health state $A \in \mathcal{A}$, then $\rho_A = 1$). We name the vector ρ *activity reduction vector*. The effective activity rate of node v at time t is $\tilde{a}(v) = \rho_{Y_v(t)} a(v)$. In the absence of activity reduction, $\rho_A = 1$, $\forall A \in \mathcal{A}$, as in the classical SIS model.

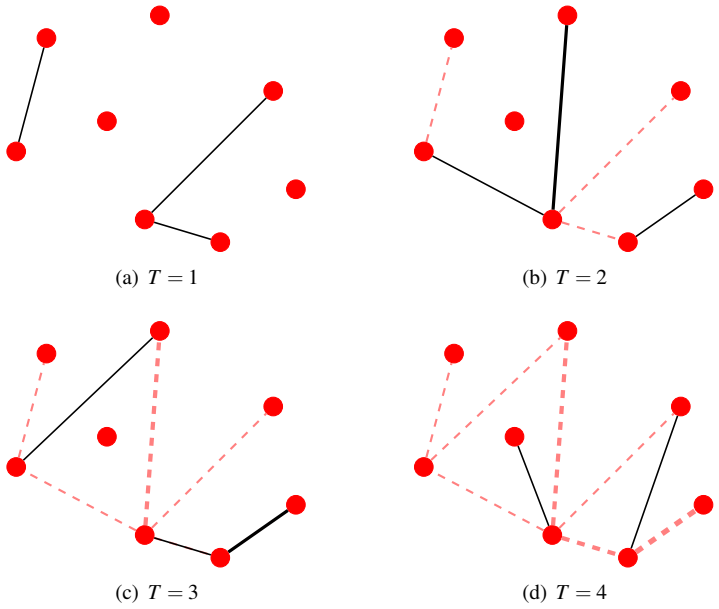
Starting from $t = 0$, node $v \in \mathcal{V}$ becomes active after a time that is sampled from an exponentially distributed random variable with parameter $\tilde{a}(v)$. In other words, the activations of node $v \in \mathcal{V}$ are distributed according to the realizations of a Poisson process with rate $\tilde{a}(v)$. When node $v \in \mathcal{V}$ activates, it contacts one node w, chosen uniformly at random over $\mathcal{V}$, generating an undirected edge (v,w). We hypothesize that the duration of this contact is instantaneous and that the epidemic propagates through these ephemeral contacts.

By integrating such continuous-time ADNs over a time window of duration T, we obtain weighted graphs, whose weighted edges identify the number of times the ADN mechanism activates that particular connection. Figure 1 illustrates this integration mechanism, depicting how the connectivity of the weighted graph increases with the duration of the time window, whereby more connections will be activated. For short time windows, the weighted graph will generally be disconnected, with only a limited number of contacts having occurred.

To model the propagation mechanism, we define a stochastic matrix $\Lambda^{(A)} \in \mathbb{R}_+^{|\mathcal{A}|\times|\mathcal{A}|}$, for each $A \in \mathcal{A}$. The propagation mechanism of the epidemics is as follows. At time t, an instantaneous edge (v,w) enters the edge set. If $Y_v(t) = A$ and $Y_w(t) = B$, then node v updates its state in a stochastic fashion, such that its new state is $C \in \mathcal{A}$ with probability $\Lambda_{AC}^{(B)}$. Similarly, node w updates its state to state $D \in \mathcal{A}$ with a probability $\Lambda_{BD}^{(A)}$. In many real-world situations, most of the entries of the matrices Λ are zeros, and only few transitions will actually occur. The edge is immediately removed and may activate again, according to the same stochastic rule.

Simultaneously, individuals may update their state and change compartments spontaneously, without any interaction with others. This is the case of the transitions between different stages of the illness, or of the final recovery from it. In order to model these dynamics, we introduce a matrix $\Theta \in \mathbb{R}_+^{|\mathcal{A}|\times|\mathcal{A}|}$, whose non-null entries θ_{AB}, with $A,B \in \mathcal{A}$, represent the rate of the spontaneous transition of each individual from state A to state B. The quantity θ_{AB}^{-1} corresponds to the expected time needed for the transition to occur.

Hence, if at time t node $v \in \mathcal{V}$ has state $Y_v(t) = A$, then, for each B such that θ_{AB} is non-null, an independent exponentially distributed random variable with parameter θ_{AB} associated with the transition from A to B is initialized. If B is the compartment associated with the variable that realizes the minimum among these random variables, node v updates its state to B after a time equal to the realization of such random variable.



(a) $T = 1$ (b) $T = 2$ (c) $T = 3$ (d) $T = 4$

FIG. 1. Integrated continuous-time ADNs. Black edges are connections generated during each time step of duration 1, red dashed edges are connections that are activated up to time $T = 1, 2, 3,$ and 4, respectively. The thickness of each edge represents its weight.

Summing up, an epidemic model on an ADN is defined through the following variables:

- a set of individuals $\mathcal{V} = \{1, \dots, n\}$, partitioned into k activation classes, each one characterized by having an *activity rate* $a_i \in \mathbb{R}_+$, $i = 1, \dots, k$;
- a set of *compartments* $\mathcal{A}$;
- a vector $\rho \in [0,1]^{|\mathcal{A}|}$, called *activity reduction vector*;
- a set of stochastic matrices $(\Lambda^{(A)})_{A \in \mathcal{A}}$, with $\Lambda^{(A)} \in \mathbb{R}_+^{|\mathcal{A}| \times |\mathcal{A}|}$, called *pairwise interaction kernels*; and
- a non-negative (entry-wise) *spontaneous transition rates matrix* $\Theta \in \mathbb{R}_+^{|\mathcal{A}| \times |\mathcal{A}|}$.

The above mentioned dynamics constitutes a Markov jump process on the space $\mathcal{A}^n$ of all configurations, where the admissible transitions are those from a state $y = (y_1, \dots, y_n)$ to state y' that differs from y in a single component $v \in V$. When $y_v = A$ and $y'_v = B$, we denote the corresponding transition rate with $q_v(A, B|y)$, such that

$$q_v(A,B|y) = \delta_A(y_v) \left[\theta_{AB} + a(v)\rho_{y_v} \frac{1}{n} \sum_{w \in \mathcal{V}} \Lambda_{AB}^{(y_w)} + \frac{1}{n} \sum_{w \in \mathcal{V}} a(w)\rho_{y_w} \Lambda_{AB}^{(y_w)} \right]. \tag{2.1}$$

Here, we have used the indicator function $\delta_S(\cdot)$ over a set S, defined as follows $\delta_S(\cdot)$, where

$$\delta_S(x) := \begin{cases} 1 & \text{if } x \in S \\ 0 & \text{if } x \notin S. \end{cases} \tag{2.2}$$

With a minor abuse of notation, when the set is the singleton $S = \{i\}$, we write $\delta_i(\cdot) := \delta_S(\cdot)$.

In order to clarify this general framework, we illustrate its application for the SIS model [22], already studied in [38], as a simple, but explanatory instance.

EXAMPLE 2.1 (The SIS model) In the standard SIS model, only two compartments are present: the individuals susceptible to the infection, denoted by S, and those infected, denoted by I. Therefore, $\mathcal{A} = \{S, I\}$. Two dynamics are possible: i) the propagation of the infection, which occurs with a fixed probability $\lambda \in [0,1]$ when a susceptible individual contacts an infected one; and ii) the recovery and transition to the susceptible state, which occurs spontaneously with a rate μ. The pairwise interaction kernel comprises the matrices $\Lambda^{(S)}$ and $\Lambda^{(I)}$, while the spontaneous transition rates matrix is Θ; these matrices are defined as follows:

$$\Lambda^{(S)} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, \quad \Lambda^{(I)} = \begin{pmatrix} 1-\lambda & \lambda \\ 0 & 1 \end{pmatrix}, \quad \Theta = \begin{pmatrix} 0 & 0 \\ \mu & 0 \end{pmatrix}. \tag{2.3}$$

Here, the first row/column of the matrices refers to susceptible individuals and the second to infected ones. Hence, entries $\Lambda_{12}^{(S)}$, $\Lambda_{12}^{(I)}$, and Θ_{12} of the matrices define transitions from compartment S to I, while entries $\Lambda_{21}^{(S)}$, $\Lambda_{21}^{(I)}$, and Θ_{21} transitions from compartment I to S.

In the classical SIS model, no activity reduction is contemplated. Hence, by using in (2.1) the explicit expressions for the matrices in (2.3), the non-zero transition rates of the Markov jump process $Y(t)$ over $\{S,I\}^n$ from state $Y(t) = y$ are

$$\begin{aligned} q_v(S,I|y) &= \delta_S(x_v)\left[\frac{1}{n}a(v)\lambda \sum_{w\in\mathcal{V}} \delta_I(y_w) + \frac{1}{n}\lambda \sum_{w\in\mathcal{V}} \delta_I(y_w)a(w)\right], \\ q_v(I,S|y) &= \delta_I(x_v)\mu. \end{aligned} \tag{2.4}$$

The state space of the Markov jump process $Y(t)$ is the space of all possible state configurations over all the nodes of the graph, that is, $\mathcal{A}^n$, whose size grows exponentially as the number of nodes in the graph increases. Therefore, a brute force analysis of the process $Y(t)$ is not feasible for the case of large graphs, that is, when $n \to \infty$. To tackle this issue, we project the stochastic process on a lower dimensional state space in a thoughtful way, such that the Markovianity of the stochastic process is preserved by the low-dimensional process.

To this aim, we consider the $|\mathcal{A}| \times k$-dimensional stochastic process $Z(t) = Z(Y(t))$, whose generic component $Z_{Ai}(t)$, $A \in \mathcal{A}$, $i \in \mathcal{K}$ corresponds to the fraction of individuals belonging to compartment $A \in \mathcal{A}$ within the ith activation class. More specifically, we write

$$Z_{Ai}(t) = \frac{|\{v : a(v) = a_i, Y_v(t) = A\}|}{n_i}. \tag{2.5}$$

Then, the following lemma holds true:

LEMMA 2.1 Process $Z(t)$ is a Markov jump process and its admissible transitions are those from state $Z(t) = z$ to a state z' that differs from z in exactly two components that belong to the same activation class $i \in \mathcal{K}$: for some $A, B \in \mathcal{A}$, the component z_{Bi} is increased by $1/n_i$, while the component z_{Ai} is reduced by the same quantity. Its transition rate is

$$p_i(A,B|z) = n_i z_{Ai}\left[\theta_{AB} + \sum_{C\in\mathcal{A}} \Lambda_{AB}^{(C)}\left(a_i\rho_A \sum_{j\in\mathcal{K}} \frac{n_j}{n} z_{Cj} + \rho_C \sum_{j\in\mathcal{K}} \frac{n_j}{n} a_j z_{Cj}\right)\right]. \tag{2.6}$$

Proof. After each transition of the process $Y(t)$, a single node v of the network has updated its state from A to B, for some $A,B \in \mathcal{A}$. Let us suppose that v belongs to the ith activation class. Then, the effect of the transition of the process $Y(t)$ reflects on the process $Z(t)$ as an increase of $1/n_i$ of the component Z_{Bi}, and a decrease of $1/n_i$ of the component Z_{Ai}. We can compute the rate of this transition by noticing from (2.1) that if $a(v) = a(w)$, then $q_v(A,B|y) = q_w(A,B|y)$, $\forall A,B \in \mathcal{A}$ and $\forall y \in \mathcal{A}^n$, such that $y_v = y_w$. Therefore, when $Y(t) = y$ and consecutively $Z(t) = Z(Y(t)) = z$, we compute

$$\begin{aligned}
p_i(A,B|y) &= \sum_{v:a(v)=a_i,y_v=A} q_v(A,B|y) \\
&= \sum_{v:a(v)=a_i,y_v=A} \left[\theta_{AB} + a_i\rho_A \frac{1}{n} \sum_{w\in\mathcal{V}} \Lambda_{AB}^{(y_w)} + \frac{1}{n} \sum_{w\in\mathcal{V}} a(w)\rho_{y_w}\Lambda_{AB}^{(y_w)} \right] \\
&= \sum_{v:a(v)=a_i,y_v=A} \left[\theta_{AB} + a_i\rho_A \frac{1}{n} \sum_{j\in\mathcal{K}} \sum_{C\in A} \Lambda_{AB}^{(C)} + \frac{1}{n} \sum_{j\in\mathcal{K}} \sum_{C\in A} a_j\rho_C\Lambda_{AB}^{(C)} \right] \\
&= \sum_{v:a(v)=a_i,y_v=A} \left[\theta_{AB} + a_i\rho_A \frac{1}{n} \sum_{j\in\mathcal{K}} n_j \sum_{C\in\mathcal{A}} z_{Cj}\Lambda_{AB}^{(C)} + \frac{1}{n} \sum_{j\in\mathcal{K}} a_j n_j \sum_{C\in\mathcal{A}} z_{Cj}\rho_C\Lambda_{AB}^{(C)} \right] \\
&= n_i z_{Ai} \left[\theta_{AB} + a_i\rho_A \sum_{j\in\mathcal{K}} \frac{n_j}{n} \sum_{C\in\mathcal{A}} z_{Cj}\Lambda_{AB}^{(C)} + \sum_{j\in\mathcal{K}} \frac{n_j}{n} a_j \sum_{C\in\mathcal{A}} \rho_C z_{Cj}\Lambda_{AB}^{(C)} \right] \\
&= n_i z_{Ai} \left[\theta_{AB} + a_i\rho_A \sum_{C\in\mathcal{A}} \Lambda_{AB}^{(C)} \sum_{j\in\mathcal{K}} \frac{n_j}{n} z_{Cj} + \sum_{C\in\mathcal{A}} \rho_C\Lambda_{AB}^{(C)} \sum_{j\in\mathcal{K}} \frac{n_j}{n} a_j z_{Cj} \right],
\end{aligned}$$

which only depends on z. Thus, we conclude that $Z(t)$ is a Markov jump process. □

REMARK 2.1 By construction, the components of (2.6) satisfy the following equation:

$$\sum_{A\in\mathcal{A}} Z_{Ai}(t) = 1, \quad \forall i \in \mathcal{K}, \forall t \in \mathbb{R}_+. \tag{2.7}$$

Thus, only the first $|\mathcal{A}|-1$ components of the vector $Z(t)$ for each class are independent and they are sufficient to reconstruct the entire state configuration (up to a relabeling of the nodes). Therefore, the state variable of the system is a $(|\mathcal{A}|-1)\times k$-dimensional vector.

EXAMPLE 2.2 (The SIS model (cont'd)) In the SIS model, already analyzed in Ex. 2.1, the state variable of the system is a k-dimensional vector with $|\mathcal{A}| = 2$. In this case, the notation can be simplified to $Z_i := Z_{Ii}$ with $Z_{Si} = 1 - Z_i$. Therefore, when $Z(t) = z$,

$$\begin{aligned}
p_i(S,I|z) &= n_i(1-z_i)\lambda \left[a_i \sum_{j\in\mathcal{K}} \frac{n_j}{n} z_j + \sum_{j\in\mathcal{K}} \frac{n_j}{n} a_j z_j \right], \\
p_i(I,S|z) &= n_i z_i \mu.
\end{aligned} \tag{2.8}$$

2.2 *From the stochastic process to a set of differential equations*

In this paper, we consider the thermodynamic limit of large populations, that is, $n \to \infty$. In this case, we suppose that the fraction of nodes in each class n_i/n is always bounded away from 0 and converges to a fixed ratio $\eta_i > 0$, when $n \to \infty$. Before presenting our approach to construct a system of ODEs, whose solution approximates the stochastic process as the number of nodes n grows, we define the following quantities.

DEFINITION 2.3 We define the *asymptotic transition rates* as

$$\tilde{p}_i(A,B|z) := \lim_{n\to\infty} \frac{p_i(A,B|z)}{n_i}, \quad \forall i \in \mathcal{K}, A,B \in \mathcal{A}, \tag{2.9}$$

when the limit exists. From (2.6), we have that

$$\tilde{p}_i(A,B|z) = z_{Ai}\left[\theta_{AB} + \sum_{C\in\mathcal{A}} \Lambda_{AB}^{(C)}\left(\rho_A a_i \sum_{j\in\mathcal{K}} \eta_j z_{Ci} + \rho_C \sum_{j\in\mathcal{K}} \eta_j a_j z_{Cj}\right)\right]. \tag{2.10}$$

LEMMA 2.2 (Thermodynamic limit) Let $\zeta(t)$ be the solution of the following system of ODEs:

$$\dot{\zeta}_{Ai} = \sum_{B\in\mathcal{A}} \tilde{p}_i(B,A|\zeta) - \sum_{C\in\mathcal{A}} \tilde{p}_i(A,C|\zeta), \quad A \in \mathcal{A}, i \in \mathcal{K}, \tag{2.11}$$

with initial conditions $\zeta_{Ai}(0) = Z_{Ai}(0)$. Then, $\forall T > 0, \exists C > 0$ such that

$$\mathbb{P}\left[\sup_{t\in[0,T]} ||Z(t) - \zeta(t)|| > \varepsilon\right] \leqslant 2k(|\mathcal{A}|-1)\exp(-Cn\varepsilon^2), \tag{2.12}$$

where $\mathbb{P}[\cdot]$ is the probability.

Proof. The estimation in (2.12) is a straightforward consequence of the application of Kurtz's theorem [52] to the Markov jump process $Z(t)$. We note that regularity conditions for the application of Kurtz's theorem, that is, Lipschitz-continuity of the terms in (2.10), are always verified. Remark 2.1 allows for considering only the first $|\mathcal{A}|-1$ components of each class, when computing the thermodynamic limit within Kurtz's theorem, thereby obtaining the multiplicative coefficient in (2.12). □

The interpretation of Lemma 2.2 for large scale networks is intuitive. In fact, formula in (2.12) guarantees that the solution of the system of ODEs in (2.11) is arbitrarily close to the stochastic process $Z(t)$, in the sense that during any finite time-horizon, the probability that the distance between the stochastic and deterministic trajectories is larger than a fixed constant approaches 0 exponentially fast in n.

As an example, we present the system of ODEs associated with the SIS model.

EXAMPLE 2.4 (The SIS model (cont'd)) For the SIS model, introduced in Ex. 2.1, with the simplified notation presented in Ex. 2.2, the asymptotic transition rates are

$$\begin{aligned} \tilde{p}_i(S,I|z) &= (1-z_i)\lambda\left[a_i \sum_{j\in\mathcal{K}} \eta_j z_j + \sum_{j\in\mathcal{K}} \eta_j a_j z_j\right], \\ \tilde{p}_i(I,S|z) &= z_i\mu. \end{aligned} \tag{2.13}$$

Therefore, the system of ODEs in (2.11) reads

$$\dot{z}_i = (1-z_i)\lambda\left[a_i \sum_{j\in\mathcal{K}} \eta_j z_j + \sum_{j\in\mathcal{K}} \eta_j a_j z_j\right] - z_i\mu, \quad i \in \mathcal{K}. \tag{2.14}$$

3. Techniques for the analysis of the system of ODEs

Although the system of ODEs in (2.11), obtained in the thermodynamic limit through Lemma 2.2, offers an elegant representation of the epidemic spreading, its analytical treatment and implementation to real-world scenarios poses a number of technical and practical challenges. With respect to the analytical

treatment, the right-hand-side of (2.11) is, in general, a nonlinear function of all the other variables, hampering the computation of an exact solution or tight approximations.

With respect to real-world applications, this system cannot be readily employed for predictions of infectious disease outbreaks, since its numerical implementation requires the knowledge of detailed initial conditions in the form of the fraction of nodes for each different state within each class. Unfortunately, in real-world scenarios, it is only possible to measure, or at least estimate, global variables such as the fraction of individuals per compartment (overall), and information about classes is in general not available.

3.1 *The macroscopic variables*

A thoughtful change of variables may help in both the analytical treatment and real-world applications. Specifically, given $\zeta(t)$ as in Lemma 2.2, we introduce the set of macroscopic variables

$$M(t) = \{M_{Aj}(t)\}_{A\in\mathcal{A},\, j\in\mathcal{K}}, \tag{3.1}$$

whose generic element is

$$M_{Aj}(t) = \sum_{i\in\mathcal{K}} \eta_i a_i^{j-1} \zeta_{Ai}(t), \qquad j \in \mathcal{K}, A \in \mathcal{A}. \tag{3.2}$$

Similar to $Z(t)$ and $\zeta(t)$, the set $M(t)$ can be naturally sorted in a $|\mathcal{A}| \times k$-dimensional vector. We refer to the subscript j as the order of the macroscopic variables, such that, given a compartment $A \in \mathcal{A}$, the corresponding first-order macroscopic variable is M_{A1}, the second-order variable is M_{A2}, and so on.

The first-order macroscopic variables have a straightforward interpretation as the overall fraction of individuals for each compartment. The second-order macroscopic variables can be interpreted as the average activation rates of individuals for each compartment. In general, the jth-order macroscopic variables will account for the $(j-1)$th statistical moment of the activity rate distribution for each compartment. Even though variables in (3.2) may be defined for any integer order $j \geqslant 1$, we will prove in the following that those with $j > k$ are linearly dependent on the first k variables, where we recall that k is the number of activation classes. However, in the remainder of this paper we will keep using variables of order $k+1$, defined as $M_{A(k+1)}(t) = \sum_{i\in\mathcal{K}} \eta_i a_i^k \zeta_{Ai}(t)$, $\forall A \in \mathcal{A}$, to retain a compact notation.

LEMMA 3.1 (Change of variables) The set of macroscopic variables $M(t)$ represents the whole state of the system without any missing degree of freedom, and its components evolve according to the following system of ODEs:

$$\begin{aligned}\dot{M}_{Aj}(t) \;=\; & \sum_{B\in\mathcal{A}} M_{Bj}\theta_{BA} - M_{Aj}\sum_{B\in\mathcal{A}}\theta_{AB} + \sum_{B\in\mathcal{A}} \rho_B M_{B(j+1)} \sum_{C\in\mathcal{A}} \Lambda_{BA}^{(C)} M_{C1} + \sum_{B\in\mathcal{A}} M_{Bj} \sum_{C\in\mathcal{A}} \Lambda_{BA}^{(C)} \rho_C M_{C2} \\ & + \rho_A M_{A(j+1)} \sum_{B\in\mathcal{A}} M_{Bj} \sum_{C\in\mathcal{A}} \Lambda_{AC}^{(B)} - M_{Aj} \sum_{B\in\mathcal{A}} \rho_B M_{B(j+1)} \sum_{C\in\mathcal{A}} \Lambda_{AC}^{(B)},\end{aligned} \tag{3.3}$$

for all $A \in \mathcal{A}$ and $j \in \mathcal{K}$.

Proof. First, we need to show that the whole state of the system is represented by $M(t)$. The change of variables can be written in a vectorial form as $M(t) = Q\zeta(t)$, which takes a simplified form due to the fact that $M_{A1}, \ldots, M_{Ak}$ only depend on the components related to compartment A of the vector $\zeta(t)$, for any $A \in \mathcal{A}$, yielding a block-diagonal structure for Q. Thus, each block of Q encapsulates the change of

variables for a single compartment A and its corresponding k equations, as

$$\begin{bmatrix} M_{A1}(t) \\ M_{A2}(t) \\ M_{A3}(t) \\ \dots \\ M_{Ak}(t) \end{bmatrix} = \begin{bmatrix} 1 & 1 & 1 & \dots & 1 \\ a_1 & a_2 & a_3 & \dots & a_k \\ a_1^2 & a_2^2 & a_3^2 & \dots & a_k^2 \\ \dots & \dots & \dots & \dots & \dots \\ a_1^{k-1} & a_2^{k-1} & a_3^{k-1} & \dots & a_k^{k-1} \end{bmatrix} \begin{bmatrix} \eta_1 & 0 & 0 & 0 & 0 \\ 0 & \eta_2 & 0 & 0 & 0 \\ 0 & 0 & \eta_3 & 0 & 0 \\ 0 & 0 & 0 & \dots & 0 \\ 0 & 0 & 0 & 0 & \eta_k \end{bmatrix} \begin{bmatrix} \zeta_{A1}(t) \\ \zeta_{A2}(t) \\ \zeta_{A3}(t) \\ \dots \\ \zeta_{Ak}(t) \end{bmatrix}. \tag{3.4}$$

To avoid the loss of degrees of freedom in this change of variables, each block of matrix Q should be full rank. From (3.4), we notice that the first matrix is the transpose of a Vandermonde matrix, which is full rank, and the second matrix is a diagonal full-rank matrix. Therefore, the product between these two matrices is non-singular.

Next, we consider that, $\forall A \in \mathcal{A}$ and $\forall\, j \in \mathcal{K}$, the following equation holds:

$$\dot{M}_{Aj}(t) = \sum_{i \in \mathcal{K}} \eta_i a_i^{j-1} \dot{\zeta}_{Ai}(t). \tag{3.5}$$

A direct substitution of the explicit expressions of the terms $\dot{\zeta}_{Ai}(t)$ from (2.11) into (3.5) yields the claim. □

We introduce the following notation for the jth statistical moments of the activity rates distribution for large scale networks:

$$\alpha_j := \lim_{n \to \infty} \sum_{i \in \mathcal{K}} \frac{n_i}{n} a_i^j = \sum_{i \in \mathcal{K}} \eta_i a_i^j, \tag{3.6}$$

where $\alpha_0 = 1$.

REMARK 3.1 Using the definition in (3.2) and the claim in Remark 2.1, we find

$$\sum_{A \in \mathcal{A}} M_{Aj}(t) = \sum_{A \in \mathcal{A}} \sum_{i \in \mathcal{K}} \eta_i a_i^{j-1} \zeta_{Ai}(t) = \sum_{i \in \mathcal{K}} \eta_i a_i^{j-1} \sum_{A \in \mathcal{A}} \zeta_{Ai}(t) = \sum_{i \in \mathcal{K}} \eta_i a_i^{j-1} = \alpha_{j-1}. \tag{3.7}$$

Similarly to the case of Remark 2.1, equation (3.7) allows for reducing the number of linearly independent variables in $M(t)$. Selecting a compartment $A \in \mathcal{A}$, its corresponding macroscopic variables can be written as a function of the remaining macroscopic variables, as

$$M_{Aj}(t) = \alpha_{j-1} - \sum_{B \in \mathcal{A} \setminus \{A\}} M_{Bj}(t) \qquad \forall\, j \in \mathcal{K}. \tag{3.8}$$

Then, system (3.3) can be reduced by eliminating the equations corresponding to the compartment A and rewriting the remaining right-hand-side of (3.3) using (3.8) when needed.

When considering the system of ODEs obtained in Lemma 3.1 for the epidemic spreading in terms of the macroscopic variables, we note that several of the challenges posed by the formulation in Lemma 2.2 are resolved. We gather these observations in the following remark.

REMARK 3.2 Equation (3.3) indicates that the evolution of $M_{Aj}(t)$ depends only on few other variables, corresponding to orders 1, 2, j, and $j+1$. Moreover, in practical epidemic models many transitions are not possible; thus, many of the elements of matrices Θ and Λ are null. These considerations yield a distinctive block structure for (3.3), which facilitates the computation of the epidemic threshold of the model, that is, the conditions for which the disease-free equilibrium is globally asymptotically stable. An example of such a computation is presented in the following.

EXAMPLE 3.1 (The SIS model (cont'd)) For the SIS model, using Remark 3.1, we can reduce the system obtained in (3.3) to a k-dimensional system of ODEs. Consistently with [38], we opt for reducing the system using only the macroscopic variables related to the compartment I, recalling that $M_{Sj} = \alpha_{j-1} - M_{Ij}$. Hence, (3.3) reads

$$\begin{aligned} \dot{M}_{I1} &= (\lambda\alpha_1 - \mu)M_{I1} + \lambda M_{I2} - 2\lambda M_{I1}M_{I2}, \\ \dot{M}_{I2} &= \lambda\alpha_2 M_{I1} + (\lambda\alpha_1 - \mu)M_{I2} - \lambda M_{I1}M_{I3} - \lambda M_{I2}^2, \\ \dot{M}_{Ij} &= \lambda\alpha_j M_{I1} + \lambda\alpha_{j-1}M_{I2} - \mu M_{Ij} - \lambda M_{I1}M_{I(j+1)} - \lambda M_{I2}M_{Ij}, \end{aligned} \tag{3.9}$$

with $j = 3, \dots, k$.

The epidemic threshold of the SIS model can be computed from (3.9) by linearizing in the vicinity of the origin and studying the stability of the resulting linear system. Due to Remark 3.2, the linearized system is block-triangular, such that stability is controlled by the 2×2 block associated with the first two macroscopic variables M_{I1} and M_{I2}, namely,

$$J = \begin{bmatrix} \lambda\alpha_1 - \mu & \lambda \\ \lambda\alpha_2 & \lambda\alpha_1 - \mu \end{bmatrix}. \tag{3.10}$$

This leads to the epidemic threshold

$$\frac{\lambda}{\mu} < \frac{1}{\alpha_1 + \sqrt{\alpha_2}}, \tag{3.11}$$

which is in agreement with [2, 8, 21].

With respect to real-world applications, the initial conditions for the first-order macroscopic variables (corresponding to the initial fraction of individuals in each compartment) is often available, or at least it is possible to estimate it with a reasonable precision. However, some issues have still to be addressed in order to provide useful estimations of the evolution of the epidemics. In fact, even though the low-order macroscopic variables can be measured or estimated from available epidemics data, it is not straightforward to estimate high-order statistical moments of activation rates from raw epidemics data. According to several studies [2, 55], the distribution of the activity rates follows a power-law distribution, possibly with some cutoff. High-order statistical moments of power-law distributions tend to blow up, and, while beneficial, cutoffs could still cause numerical instability. Besides problems in the estimation of the initial condition, the direct application of (3.3) could introduce numerical issues for the integration of high-order macroscopic variables, thereby propagating initial errors in the estimation of the initial condition.

In order to tackle these issues, we propose two possible approaches. On the one hand, as discussed in Section 3.2, the system of equations can be reduced by using bounds that eliminate the high-order macroscopic variables, which could cause numerical issues and whose initial conditions are difficult to measure. Following this approach, we estimate the distribution of the health states among the individuals in the endemic state and in the transient phase. On the other hand, Section 3.3 presents a data-driven approach, which uses epidemic data sampled at the population level during the evolution of the epidemics to perform a prediction of the epidemic dynamics.

3.2 *A low-dimensional system of differential inclusions*

We propose a reduced system based on a subset of equations of the system of macroscopic variables (3.3), obtained from its block structure. By extending the technique developed in [38] for the

SIS model, we consider only the equations corresponding to macroscopic variables up to a select order k^*. According to Remark 3.2, the ODE for a general jth order macroscopic variable with $j < k^*$ does not depend on any of the eliminated macroscopic variables. On the other hand, the ODEs corresponding to the k^*th order variables may depend on the (k^*+1)th order ones, that is, $M_{A(k^*+1)}$, with $A \in \mathcal{A}$. Thus, system (3.3) cannot be simply truncated to order k^*. To overcome this issue, we put forward some bounds relating $M_{A(k^*+1)}$ to lower-order variables and to the statistical moments of the activity rate distribution. Similar to [38], some elementary bounds are derived as

$$M_{A(k^*+1)}(t) \in \mathcal{I}[M_{A(k^*+1)}(t)] := [a_1 M_{Ak^*}, \min\{a_k M_{Ak^*}(t), \alpha_{k^*}\}], \qquad \forall A \in \mathcal{A}, \forall t \in \mathbb{R}_+. \tag{3.12}$$

We observe that bounds in (3.12) are only based on structural properties of the network and the activity distribution. Tighter bounds on $M_{A(k^*+1)}$ may be established if more information about the activity rate distribution is available, or some observations from epidemic data can be accessed.

Using bounds established in (3.12), or tighter bounds if available, in the system of ODEs in (3.3), a system of ODIs can be obtained. The system is composed of $|\mathcal{A}|(k^*-1)$ ODEs, referring to the macroscopic variables of the first k^*-1 orders, and $|\mathcal{A}|$ ODIs for the k^*th order macroscopic variables. These ODIs have the following general formulation:

$$\begin{aligned} \dot{M}_{Ak^*}(t) \quad \in \quad & \sum_{B\in\mathcal{A}} M_{Bk^*}\theta_{BA} - M_{Ak^*}\sum_{B\in\mathcal{A}}\theta_{AB} + \\ & + \sum_{B\in\mathcal{A}} \rho_B \mathcal{I}[M_{B(k^*+1)}(t)] \sum_{C\in\mathcal{A}} \Lambda^{(C)}_{BA} M_{C1} + \sum_{B\in\mathcal{A}} M_{Bk^*} \sum_{C\in\mathcal{A}} \Lambda^{(C)}_{BA}\rho_C M_{C2} + \\ & - \rho_A \mathcal{I}[M_{A(k^*+1)}(t)] \sum_{B\in\mathcal{A}} M_{Bk^*} \sum_{C\in\mathcal{A}} \Lambda^{(B)}_{AC} - M_{Ak^*} \sum_{B\in\mathcal{A}} \rho_B \mathcal{I}[M_{B(k^*+1)}(t)] \sum_{C\in\mathcal{A}} \Lambda^{(B)}_{AC}. \end{aligned} \tag{3.13}$$

REMARK 3.3 Recalling the variable reduction in (3.8) based on Remark 3.1, the system of ODIs can be reduced to a $(|\mathcal{A}|-1)k^*$-dimensional system composed of $(|\mathcal{A}|-1)(k^*-1)$ ODEs and $|\mathcal{A}|-1$ ODIs.

REMARK 3.4 In general, the ODE for $M_{A(k^*+1)}(t)$ may depend on macroscopic variables up to the (k^*+2)th order. However, based on the features of the model, such a dependence may not be present for some $A \in \mathcal{A}$. In this case, instead of using the bounds in (3.12) for $M_{A(k^*+1)}(t)$, one may add the following ODI to (3.13):

$$\dot{M}_{A(k^*+1)}(t) \in \sum_{B\in\mathcal{A}} M_{B(k^*+1)}\theta_{BA} - M_{A(k^*+1)}\sum_{B\in\mathcal{A}}\theta_{AB} + \sum_{B\in\mathcal{A}} M_{B(k^*+1)} \sum_{C\in\mathcal{A}} \Lambda^{(C)}_{BA}\rho_C M_{C2}, \tag{3.14}$$

with the understanding that any occurrence of $\mathcal{I}[M_{A(k^*+1)}(t)]$ in the system is replaced by $M_{A(k^*+1)}(t)$.

This reduced system of ODIs has a relevant use in the special cases $k^* = 1$ and $k^* = 2$. In the first case, one can bound the values of the zeros on the right-hand sides of (3.3), which coincides with the proportion of individuals per health state at the endemic equilibria. Most of the epidemic models used in epidemiology have a unique endemic equilibrium [56, 57], but in general there could exist more than one. In order to construct such bounds, one should take into account the dependence of the growth of the first-order macroscopic variables on the variables we are seeking to bound (in this case, second-order variables) through (3.12). For a choice of a first-order macroscopic variable and a higher-order variable, we aim at understanding whether the upper-bound for the higher-order quantity translates into an upper- or a lower-bound on the first-order variable. This analysis, as shown later in this paper, is performed by studying the right-hand-side of (3.3).

In the case $k^* = 2$, one can bound the evolution of the fraction of individuals per each compartment, that is, the first-order macroscopic variables, during the transient phase, as the epidemic approaches the endemic state. In order to construct such bounds, one should take into account the dependence of the growth of the first-order macroscopic variables on the third-order variables, similar to the case $k^* = 1$, discussed above. By following this procedure, we can substitute either the maximum or the minimum value (depending on the exhibited dependence) of the third-order macroscopic variables from (3.12) to establish an upper-bound and a lower-bound for the evolution of the first-order variables in the transient phase.

EXAMPLE 3.2 (The SIS model(cont'd)) For the SIS model, setting $k^* = 1$ we bound

$$M_{I2} \in [a_1 M_{I1}, \min\{\alpha_1, a_k M_{I1}\}], \tag{3.15}$$

thereby obtaining

$$\dot{M}_{I1} \in (\lambda\alpha_1 - \mu) M_{I1} + \lambda(1 - 2M_{I1})\,[a_1 M_{I1}, \min\{\alpha_1, a_k M_{I1}\}]\,. \tag{3.16}$$

The fraction of infected individuals in the endemic state can be bounded by setting the right-hand-side of (3.16) to 0 and finding an interval for the solution, that is,

$$\begin{aligned}
\bar{M}_{I1} &\in \left[\frac{\lambda(a_1+\alpha_1)-\mu}{2\lambda a_1}, \min\left\{\frac{\lambda\alpha_1}{\lambda\alpha_1+\mu}, \frac{\lambda(a_k+\alpha_1)-\mu}{2\lambda a_1}\right\}\right] && \text{if } \lambda\alpha_1 < \mu; \\
\bar{M}_{I1} &= \frac{1}{2} && \text{if } \lambda\alpha_1 = \mu; \\
\bar{M}_{I1} &\in \left[\max\left\{\frac{\lambda\alpha_1}{\lambda\alpha_1+\mu}, \frac{\lambda(a_k+\alpha_1)-\mu}{2\lambda a_1}\right\}, \frac{\lambda(a_1+\alpha_1)-\mu}{2\lambda a_1}\right] && \text{if } \lambda\alpha_1 > \mu.
\end{aligned} \tag{3.17}$$

The analysis of the dependence on M_{I2} has been carried out by considering the right-hand-side of the ODE for M_{I1} in (3.9). We observe that the multiplicative coefficient of the term M_{I2} is greater than 0 if and only if $M_{I1} < 1/2$. Thus, we conclude that M_{I1} has more marked increase when M_{I2} is larger, if $M_{I1} < 1/2$; the opposite behavior is exhibited when $M_{I1} > 1/2$. Similar arguments are used in other models throughout the paper. Explicit computations are detailed in [38].

In the case $k^* = 2$, we observe from the second equation in (3.9) that M_{I2} has more marked increase when M_{I3} is smaller. Recalling how M_{I1} depends on M_{I2} for $k^* = 1$, we conclude that M_{I1} has more marked increase when M_{I3} is smaller, if $M_{I1} < 1/2$, and it exhibits the opposite behavior otherwise. Thus, two ancillary ODEs for the evolution of the macroscopic variable M_{I2} can be constructed that combine the upper- and the lower-bound in (3.12) based on the dependency discussed above, in order to act as upper- and lower-bound on M_{I1}. The coupling of each of these equations with the first equation in (3.9) produces an upper- and a lower-bound on the evolution of the variable M_{I1}, as the epidemic becomes endemic (explicit computations are detailed in [38]).

3.3 *Short- and medium-time-horizon predictions*

An alternative approach entails the use of epidemic data, sampled at the population level at every time period T, where T is the forecast horizon of the predictions. The goal of these predictions is to approximate the epidemic curve, described by the evolution of the first-order macroscopic variables. Considering the accuracy obtained in the estimation of the transient evolution of the process obtained in

Section 3.2 using $k^* = 2$, we construct a reduced system of $2 \times |\mathcal{A}|$ ODEs for the first- and second-order macroscopic variables, and we hypothesize the third-order variables to be linearly dependent on the first-order ones as

$$M_{A3}(t) = K_A M_{A1}(t), \qquad \forall A \in \mathcal{A}. \tag{3.18}$$

where the proportionality constant K_A is estimated from the available epidemic data. Therefore, the system of equations reduces to

$$\begin{aligned}
\dot{M}_{A1}(t) &= \sum_{B\in\mathcal{A}} M_{B1}\theta_{BA} - M_{A1}\sum_{B\in\mathcal{A}}\theta_{AB} + \\
&\quad + \sum_{B\in\mathcal{A}} \rho_B M_{B2} \sum_{C\in\mathcal{A}} \Lambda_{BA}^{(C)} M_{C1} + \sum_{B\in\mathcal{A}} M_{B1} \sum_{C\in\mathcal{A}} \Lambda_{BA}^{(C)} \rho_C M_{C2} + \\
&\quad - \rho_A M_{A2} \sum_{B\in\mathcal{A}} M_{B1} \sum_{C\in\mathcal{A}} \Lambda_{AC}^{(B)} - M_{A1} \sum_{B\in\mathcal{A}} \rho_B M_{B2} \sum_{C\in\mathcal{A}} \Lambda_{AC}^{(B)} \\
\dot{M}_{A2}(t) &= \sum_{B\in\mathcal{A}} M_{B2}\theta_{BA} - M_{A2}\sum_{B\in\mathcal{A}}\theta_{AB} + \\
&\quad + \sum_{B\in\mathcal{A}} \rho_B K_B M_{B1} \sum_{C\in\mathcal{A}} \Lambda_{BA}^{(C)} M_{C1} + \sum_{B\in\mathcal{A}} M_{B2} \sum_{C\in\mathcal{A}} \Lambda_{BA}^{(C)} \rho_C M_{C2} + \\
&\quad - \rho_A K_A M_{A1} \sum_{B\in\mathcal{A}} M_{B2} \sum_{C\in\mathcal{A}} \Lambda_{AC}^{(B)} - M_{A2} \sum_{B\in\mathcal{A}} \rho_B K_B M_{B1} \sum_{C\in\mathcal{A}} \Lambda_{AC}^{(B)}.
\end{aligned} \tag{3.19}$$

Here, we note that the number of equations in (3.19) can be reduced to $2(|\mathcal{A}|-1)$, following Remark 3.1.

The algorithm generates piece-wise predictions, named $M_{Aj}^{(h)}(t)$, with $A \in \mathcal{A}$, $j = \{1,2\}$, and $h \in \mathbb{N}$, where $\mathbb{N}$ is the set of non-negative integers and $M_{Aj}^{(h)}(t)$ is the prediction of the evolution of the macroscopic variable $M_{Aj}(t)$ in the time interval $t \in [hT, (h+1)T)$, defined as follows.

We suppose that at time $t = 0$ some epidemic data are available. Specifically, we assume to have access to the fraction of individuals per compartment $M_{A1}(0)$, $\forall A \in \mathcal{A}$. Then, the algorithm is implemented as follows.

Algorithm 3.3 (Prediction algorithm) We initialize the algorithm by setting $M_{A1}^{(0)}(0) = M_{A1}(0)$, $\forall A \in \mathcal{A}$. If information about the value of the second-order macroscopic variables is available, then $M_{A2}^{(0)}(0)$ can also be initialized. Otherwise, we set $M_{A2}^{(0)}(0) = \alpha_1 M_{A1}(0)$, where α_1 in (3.6) is obtained from the power-law distribution of the activity rate distribution. We initialize $K_A^{(0)} = \alpha_2$ and $K_A = K_A^{(0)}$, $\forall A \in \mathcal{A}$. Then, we set $h = 0$ and the algorithm loops through the following steps:

1. the system of ODEs in (3.19) is numerically integrated from time hT to time $(h+1)T$, producing the solutions $M_{Aj}^{(h)}(t)$, $A \in \mathcal{A}$, $j = \{1,2\}$;

2. at time $t = (h+1)T$, we sample epidemic data at the population level. Specifically we measure $M_{A1}(t)$, $\forall A \in \mathcal{A}$, and, if data is accessible, we measure also $M_{A2}(t)$, $\forall A \in \mathcal{A}$;

3. we set the initial conditions $M_{A1}^{(h+1)}(t) = M_{A1}(t)$, $\forall A \in \mathcal{A}$ and, if available, $M_{A2}^{(h+1)}(t) = M_{A2}(t)$, $\forall A \in \mathcal{A}$;

4. using the sampled epidemic data, we update the proportionality constants

$$K_A^{(h+1)} = f(K_A^{(h)}, M_{Aj}^{(h+1)}(t), M_{Aj}(t)). \tag{3.20}$$

If $M_{A2}(t)$ is not accessible, we set $M_{A2}^{(h+1)}(t) = f(M_{Aj}^{(h+1)}(t), M_{A1}(t))$, $\forall A \in \mathcal{A}$; and

5. h is incremented by 1, and the process resumes from step 1.

EXAMPLE 3.4 (The SIS model (cont'd)) In the case of the SIS model, predictions can be generated by integrating the system

$$\begin{aligned} \dot{M}_{I1} &= -\mu M_{I1} + \lambda\alpha_1 M_{I1} + \lambda M_{I2} - 2\lambda M_{I1} M_{I2}, \\ \dot{M}_{I2} &= -\mu M_{I2} + \lambda\alpha_2 M_{I1} + \lambda\alpha_1 M_{I2} - \lambda K M_{I1}^2 - \lambda M_{I2}^2, \end{aligned} \tag{3.21}$$

where the proportionality constant, initially set to $K^{(0)} = \alpha_2$, updated by choosing a parameter $\beta > 0$, according to [38], follows:

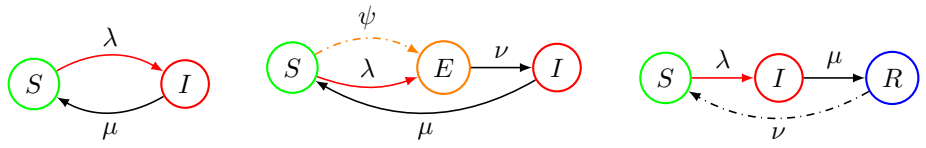
$$K^{(h+1)} = K^{(h)} \left[1 + \beta \frac{M_{I1}((h+1)T) - M_{I1}^{(h)}((h+1)T)}{1 - 2M_{I1}((h+1)T)} \right]. \tag{3.22}$$

4. Application to representative epidemic models

Here, we focus on some representative epidemic models, each one exemplifying some unique features with respect to the standard SIS model. Each model is analyzed within the continuous-time discrete-distribution paradigm of ADNs developed in Section 2, by using the mathematical techniques presented in Section 3.

At first, in Section 4.1, we consider the presence of behavioral changes due to infection [8]. Thereafter, in Section 4.2, we investigate the presence of an intermediate phase between the susceptible and the infected one [51], in which an individual has already been exposed to the pathogen, but he/she is not aware of this fact. This model is typically referred to as susceptible-exposed-infected-susceptible (SEIS), where exposed individuals may be latent (not contagious [58]) or contagious [51]. Finally, in Section 4.3, we include the mechanism of immunization [22], which can be temporary in the case of the susceptible-infected-recovered-susceptible model (SIRS), used for influenza [59], or permanent in the case of the susceptible-infected-recovered model (SIR), used for smallpox [60]. The possible state transitions characterizing these three models are illustrated in Fig. 2. All of these features can be combined, generating more involved and realistic epidemic models.

In each section, along with the mathematical derivation of the specific model, we test our approach against Monte Carlo simulations (over 200 trials) of the stochastic process, generated by means of a Gillespie algorithm [39, 40]. In all these examples, we set the model parameters based on the flu case study in a university campus proposed in [38]. Specifically, the model parameters are: $n = 30896$, $k = 59$ ($a_1 = 0.12$, $a_{59} = 6$), $\gamma = 2.09$, $\lambda = 0.43$, $\mu = 0.138$, $\alpha_1 = 0.317$, $\alpha_2 = 0.381$, and the time unit is a day. Details on the derivation of these parameters are in [38], along with a comparison of the epidemic curve and of the accuracy of our prediction techniques using a set of parameter based on a different case study. Only model-specific, new parameters will be mentioned in the figure captions and their effect will be analyzed in the simulations.



(a) SIS (b) SEIS (c) SIR(S)

FIG. 2. State transitions characterizing the epidemic models analyzed in this paper. Black lines are spontaneous transitions, colored lines are transitions taking place after a contact with a node of a specific compartment (color of the line). Dashed lines are transitions that may not take place, depending on the specific model implementation.

4.1 *SIS model with behavioral changes due to infection*

Here, we consider a model where individuals reduce their activity because of the infection. The addition of this new feature is based on real-world evidence, which suggests that infected individuals tend to reduce their activity, due to the epidemics itself (sick people may not go to work, for example) and control policies, such as quarantine [61, 62]. For the sake of simplicity, we focus on the SIS model, which we extend by including the activity reduction vector $\rho = [1, \sigma]$, where the parameter $\sigma \in [0,1]$ represents the reduction of activity of the infected individuals. Notably, $\rho = 0$ means that infected individuals do not generate links, whereas $\sigma = 1$ indicates the absence of behavioral changes. In contrast with [8], where modeling behavioral changes requires the use of two different parameters, the present continuous-time discrete-distribution framework allows for the use of a single parameter σ.

Following the same approach sketched for the standard SIS model in Ex. 2.2 and Ex. 2.4, the system of ODEs in the thermodynamic limit (Lemma 2.2) reads

$$\dot{z}_i = (1-z_i)\lambda \left[a_i \sum_{j\in\mathcal{K}} \eta_j z_j + \sigma \sum_{j\in\mathcal{K}} \eta_j a_j z_j \right] - z_i \mu, \quad i \in \mathcal{K}. \tag{4.1}$$

A comparison between the numerical integration of (4.1) and the simulation of the stochastic Markov process (both single trajectories and Monte Carlo average estimation) for different values of n is presented in Fig. 3, demonstrating the accuracy of the deterministic approximation as n grows.

By using the change of variables in Lemma 3.1, we obtain the following k-dimensional system of ODEs:

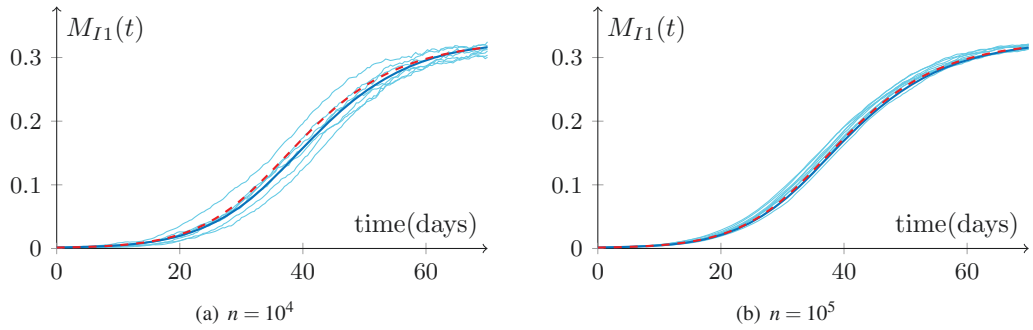
$$\begin{aligned} \dot{M}_{I1} &= (\lambda\alpha_1 - \mu)M_{I1} + \sigma\lambda M_{I2} - (1+\sigma)\lambda M_{I1}M_{I2}, \\ \dot{M}_{I2} &= \lambda\alpha_2 M_{I1} + (\sigma\lambda\alpha_1 - \mu)M_{I2} - \lambda M_{I1}M_{I3} - \sigma\lambda M_{I2}^2, \\ \dot{M}_{Ij} &= \lambda\alpha_j M_{I1} + \sigma\lambda\alpha_{j-1}M_{I2} - \mu M_{Ij} - \lambda M_{I1}M_{I(j+1)} - \sigma\lambda M_{I2}M_{Ij}, \end{aligned} \tag{4.2}$$

with $j = 3, \ldots, k$.

In order to analyze the stability of the disease-free equilibrium, we can take advantage of the particular block structure of the system in (4.2). Similar to [8], we derive the epidemic threshold

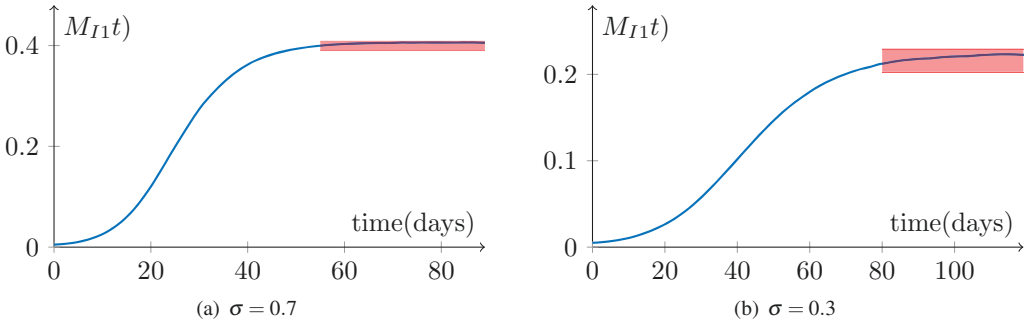
$$\frac{\lambda}{\mu} < \frac{2}{(1+\sigma)\alpha_1 + \sqrt{(1-\sigma)^2\alpha_1^2 + 4\sigma\alpha_2}}. \tag{4.3}$$

REMARK 4.1 The expression in (4.3) simplifies to $(\alpha_1 + \sqrt{\alpha_2})^{-1}$ in the case of the standard SIS model when $\sigma = 1$, in agreement with Ex. 3.1.



(a) $n = 10^4$ (b) $n = 10^5$

FIG. 3. Representative subset of Monte Carlo simulation of the evolution of the fraction of infected individuals in the stochastic process (thin cyan) and average estimation over 200 trials (solid blue) compared with the solution of the deterministic system of ODEs from (4.1) (dashed red) for the flu case study [38] with $\sigma = 0.5$, for two different population sizes n.



(a) $\sigma = 0.7$ (b) $\sigma = 0.3$

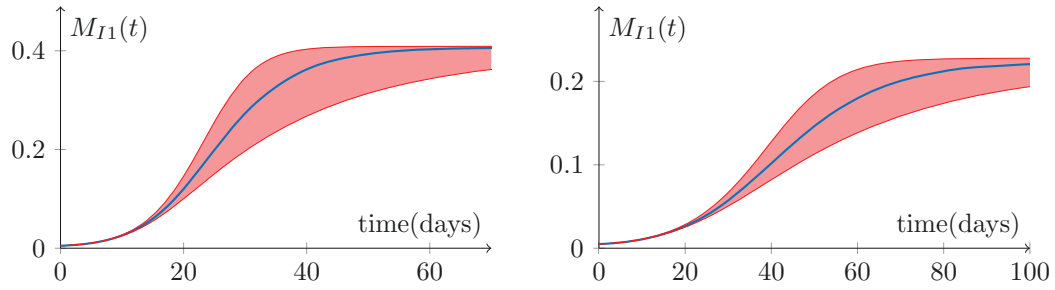
FIG. 4. Monte Carlo simulations (averaged over 200 trials) of the evolution of the fraction of infected individuals in the stochastic process (blue) for the flu case study [38], and bounds (red) obtained with $k^* = 1$, for different values of the behavioral parameter σ.

Next, we consider the k^*-dimensional system of ODIs obtained by using the bounds on $M_{I(k^*+1)}$ in the k^*th equation of (4.2), presented in Section 3. Similar to Ex. 3.2, we use the case $k^* = 1$ to estimate the fraction of infected individuals at the endemic equilibrium $\bar{M}_{I1}$, such that

$$\begin{aligned}
\bar{M}_{I1} &\in \left[\frac{\lambda(\sigma a_1+\alpha_1)-\mu}{(1+\sigma)\lambda a_1}, \min\left\{\frac{\lambda\alpha_1\sigma}{\lambda\alpha_1\sigma+\mu}, \frac{\lambda(\sigma a_k+\alpha_1)-\mu}{(1+\sigma)\lambda a_k}\right\}\right] && \text{if } \lambda\alpha_1 < \mu; \\
\bar{M}_{I1} &= \frac{\sigma}{1+\sigma} && \text{if } \lambda\alpha_1 = \mu; \\
\bar{M}_{I1} &\in \left[\max\left\{\frac{\lambda\alpha_1\sigma}{\lambda\alpha_1\sigma+\mu}, \frac{\lambda(\sigma a_k+\alpha_1)-\mu}{(1+\sigma)\lambda a_k}\right\}, \frac{\lambda(\sigma a_1+\alpha_1)-\mu}{(1+\sigma)\lambda a_1}\right] && \text{if } \lambda\alpha_1 > \mu.
\end{aligned} \tag{4.4}$$

Figure 4 demonstrates the application of these bounds for the flu case study, showing the accuracy of the bounds obtained for the fraction of infected individuals in the endemic state. From the comparison of panels (a) and (b) in Fig. 4, the amplitude of the bounds seems to be influenced by the model parameters, yielding predictions of different accuracy for the fraction of infected individuals.

The epidemic curve, instead, could be estimated by using $k^* = 2$, which leads to a system of ODIs,



(a) $\sigma = 0.7$ (b) $\sigma = 0.3$

FIG. 5. Monte Carlo simulations (averaged over 200 trials) of the evolution of the fraction of infected individuals in the stochastic process (blue) for the flu case study [38] and bounds (red) obtained with $k^* = 2$, for different values of the behavioral parameter σ.

whose solutions can be bounded by coupling the first equation from (4.2) with the following two ancillary ODEs:

$$\dot{M}_{I2} = \lambda\big(\alpha_2 - \phi_{\varepsilon,M_{I2}}(M_{I1})\big)M_{I1} + (\lambda\sigma\alpha_1 - \mu)M_{I2} - \lambda\sigma M_{I2}^2, \tag{4.5}$$

$$\dot{M}_{I2} = \lambda\big(\alpha_2 - \phi_{\bar{\varepsilon},M_{I2}}(1 - M_{I1})\big)M_{I1} + (\lambda\sigma\alpha_1 - \mu)M_{I2} - \lambda\sigma M_{I2}^2, \tag{4.6}$$

in the limit $\varepsilon \to 0$. Here, $\phi_{\varepsilon,M_{I2}}(M_{I1})$ is a continuous monotone function such that

$$\phi_{\varepsilon,M_{I2}}(M_{I1}) = \begin{cases} a_1 M_{I2} & \text{if } M_{I1} < \sigma/(1+\sigma) - \varepsilon/2, \\ \min\{\alpha_2, a_k M_{I2}\} & \text{if } M_{I1} > \sigma/(1+\sigma) + \varepsilon/2. \end{cases} \tag{4.7}$$

Similar to Ex. 3.2, (4.5) and (4.6) are obtained by analyzing the dependence of M_{I1} on M_{I3}. Figure 5 illustrates the implementation of the bounds on the epidemic curve for the flu case study [38]. As already noticed in the case $k^* = 1$, the amplitude of the bounds is influenced by the model parameters. However, satisfactory predictions of the evolution of the fraction of infected individuals in the population are obtained for both the values of the behavioral parameter σ used in the simulations.

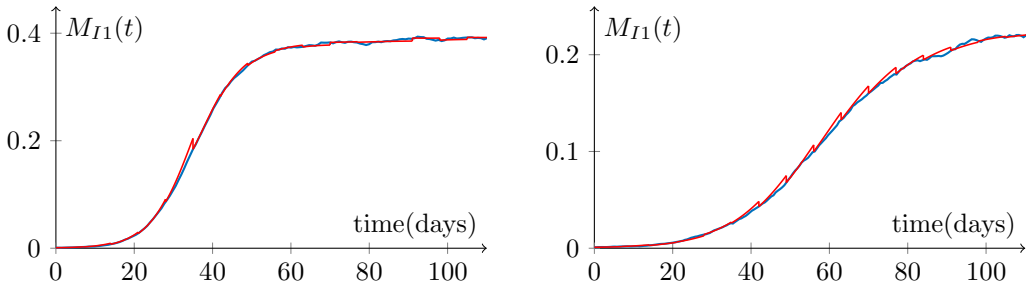
Finally, accurate finite-time-horizon predictions, based on the availability of few epidemic data at the population level at each fixed time (daily or weekly) can be gathered through Algorithm 3.3). In the case of an SIS model with behavioral changes due to infection, the equations for the first to macroscopic variables (3.19) reduces to

$$\begin{aligned} \dot{M}_{I1} &= \lambda\alpha_1 M_{I1} + \lambda\sigma M_{I2} - \mu M_{I1} - (1+\sigma)\lambda M_{I1} M_{I2}, \\ \dot{M}_{I2} &= \lambda\alpha_2 M_{I1} + (\lambda\sigma\alpha_1 - \mu)M_{I2} - \lambda K M_{I1}^2 - \lambda\sigma M_{I2}^2. \end{aligned} \tag{4.8}$$

This system presents a single proportionality constant $K = K^{(h)}$, which we initially set to $K^{(0)} = \alpha_2$, and dynamically update, according to our prediction algorithm, following:

$$K^{(h+1)} = K^{(h)}\left[1 + \beta\frac{M_{I1}((h+1)T) - M_{I1}^{(h)}((h+1)T)}{\rho - (1+\rho)M_{I1}((h+1)T)}\right]. \tag{4.9}$$

Similar to Ex. 3.4, this equation is obtained by considering the dependence of M_{I1} on M_{I2} and M_{I3} and introducing an update parameter $\beta > 0$. Figure 6 illustrates two examples of the application of



(a) $\sigma = 0.7$ (b) $\sigma = 0.3$

FIG. 6. Simulation of the evolution of the fraction of infected individuals in the stochastic process (blue) for the flu case study [38] and our predictions over a weekly time-horizon (red) for different values of σ.

this prediction technique, over a weekly time-horizon, for different values of the parameter σ. From this figure, we appreciate the accuracy of our predictions, which are only based on few epidemic data, sporadically sampled at the population level. From our findings, as σ decreases we may conjecture that the quality of the prediction slightly decreases, even if it is still accurate. This could be due to the fact that the lower σ is, the higher is the influence of the estimated term containing K in (4.8), which, differently from other terms, does not depend on σ.

4.2 *Presence of unconscious infected individuals*

As a second application, we study a more complex and realistic epidemic model, obtained by adding a new health state (compartment) for the individuals. Beyond the classical susceptible and infected compartments, we include the exposed one. The reason for this addition lies in the fact that, immediately after an individual has been exposed to the pathogen, his/her symptoms and signs may still not be evident. Therefore, an exposed individual could behave as a susceptible subject, generating contacts without reducing his/her activity. However, he/she has already contracted the infection and might be able to spread the epidemic through a contact with a susceptible individual. We consider the case in which an exposed individual is contagious within the SEIS model [58]. We note that the case where exposed individuals are latent can be easily obtained from the SEIS model, as we show further on.

In the SEIS model, we distinguish two possible types of contact: the contact between a susceptible individual and an exposed one, and the contact between a susceptible individual and an infected one. Therefore, we introduce two different infection probabilities, denoted by $\psi \in [0,1]$ and $\lambda \in [0,1]$, respectively. This distinction allows to model different scenarios. The case $\psi < \lambda$ models situations where the inception of symptoms increases the likelihood of contagion, such as in the case of influenza with coughing and sneezing. On the other hand, $\psi > \lambda$ models those cases where individual awareness about personal health state triggers precautions toward the others. When exposed individuals are not contagious, we set $\psi = 0$. The transition between exposed and infected compartments is spontaneous: after a time that is drawn from an exponentially distributed random variable with parameter ν, first symptoms appear and the exposed individual becomes conscious of his/her illness, turning into an infected individual. The parameter ν can be identified when specializing the model to a given pathogen, whereby ν^{-1} is the expected duration of the incubation period.

Hence, we set $\mathcal{A} = \{S,E,I\}$ and $\rho = [1,1,\sigma]$. The pairwise interaction kernel consists of the fol-

lowing matrices, where the first row/column refers to the compartment S, the second one to E, and the last one to I:

$$\Lambda^{(S)} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}, \quad \lambda^{(E)} = \begin{pmatrix} 1-\psi & \psi & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}, \quad \lambda^{(I)} = \begin{pmatrix} 1-\lambda & \lambda & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}, \tag{4.10}$$

while the spontaneous transition rates matrix is

$$\Theta = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & \nu \\ \mu & 0 & 0 \end{pmatrix}. \tag{4.11}$$

Next, in the thermodynamic limit (Lemma 2.2), the process can be approximated through the solution of the following system of ODEs:

$$\begin{aligned} \dot{\zeta}_{Ei} &= -\nu\zeta_{Ei} + (1-\zeta_{Ei}-\zeta_{Ii})\left[\psi(a_i\sum_{j\in\mathcal{K}}\eta_j\zeta_{Ej} + \sum_{j\in\mathcal{K}}\eta_j a_j\zeta_{Ej}) + \lambda(a_i\sum_{j\in\mathcal{K}}\eta_j\zeta_{Ij} + \sigma\sum_{j\in\mathcal{K}}\eta_j a_j\zeta_{Ij}\right], \\ \dot{\zeta}_{Ii} &= -\mu\zeta_{Ii} + \nu\zeta_{Ei}, \end{aligned} \tag{4.12}$$

where the variables referring to the compartment S, that is, ζ_{Si}, have been treated as linearly dependent on the others, based on the method for reducing the number of variables presented in Remark 2.1.

The system of ODEs in (4.12) can be rewritten in terms of the macroscopic variables M_{Ej} and M_{Ij}, according to the change of variables in Lemma 3.1, thereby leading to

$$\begin{aligned} \dot{M}_{E1} = \;& (\alpha_1\psi-\nu)M_{E1} + \alpha_1\lambda M_{I1} + \psi M_{E2} + \lambda\sigma M_{I2} - 2\psi M_{E1}M_{E2} \\ & -(1+\sigma)\lambda M_{I1}M_{I2} - (\lambda+\psi)M_{E2}M_{I1} - (\lambda\sigma+\psi)M_{E1}M_{I2}, \\ \dot{M}_{E2} = \;& \alpha_2\psi M_{E1} - \nu M_{E2} + \alpha_2\lambda M_{I1} + \psi\alpha_1 M_{E2} + \lambda\sigma\alpha_1 M_{I2} - \psi M_{E1}M_{E3} - \psi M_{E2}^2 \\ & -\lambda M_{I1}M_{I3} - \psi\sigma M_{I2}^2 - \psi M_{E1}M_{E3} - (\psi+\lambda\sigma)M_{E2}M_{I2} - \lambda M_{E3}M_{I1}, \\ \dot{M}_{Ej} = \;& \alpha_j\psi M_{E1} - \nu M_{Ej} + \alpha_j\lambda M_{I1} + \psi\alpha_{j-1}M_{E2} + \lambda\sigma\alpha_{j-1}M_{I2} \\ & -\psi M_{E1}M_{E(j+1)} - \psi M_{E2}M_{Ej} - \lambda M_{I1}M_{I(j+1)} - \psi\sigma M_{I2}M_{Ij} - \psi M_{E1}M_{I(j+1)} \\ & -\psi M_{E2}M_{Ij} - \lambda M_{E(j+1)}M_{I1} + \lambda\sigma M_{Ej}M_{I2}, \\ \dot{M}_{Il} = \;& -\mu M_{Il} + \nu M_{El}, \end{aligned} \tag{4.13}$$

with $j = 3,\dots,k$ and $l \in \mathcal{K}$.

The technique presented in Section 3.2 can be employed to estimate the fraction of infected (and exposed) individuals in the endemic state. Comparing with a binary dynamics, such as the SIS model or the one described in Section 4.1, a few more technical expedients should be put forward to tackle the increased dimensionality of this system. In particular, from (4.13), we notice that the ODEs for the jth order macroscopic variables referring to compartment I depend only on the two macroscopic variables M_{Ij} and M_{Ej} and not on higher-order variables.

Thus, to estimate the fraction of individuals per health state in the endemic state, we can consider a three-dimensional system of ODIs composed of the three equations describing the dynamics of M_{E1}, M_{I1}, and M_{I2} from (4.13), in which only the term M_{E2} is bounded using (3.12), that is, $a_1 M_{E1} \leqslant M_{E2} \leqslant \min\{a_k M_{E1}, \alpha_1\}$. The resulting system comprises two ODIs for $\dot{M}_{E1}$ and $\dot{M}_{I2}$ and an ODE for $\dot{M}_{I1}$.

The fraction of exposed and infected individuals at the endemic equilibrium ($\bar{M}_{E1}$ and $\bar{M}_{I1}$, respectively) can be estimated by considering that, at the equilibrium point, the two equations for the macroscopic variables M_{I1} and M_{I2} yield

$$M_{I1} = \frac{\nu}{\mu} M_{E1} \quad \text{and} \quad M_{I2} = \frac{\nu}{\mu} M_{E2}. \tag{4.14}$$

Hence, equilibrium conditions for the equation $\dot{M}_{E1}$ from (4.13) reduce to the computation of the roots of the following expression:

$$(\alpha_1 \psi + \alpha_1 \lambda \frac{\nu}{\mu} - \nu) M_{E1} + \left[\left(\psi + \lambda \sigma \frac{\nu}{\mu} \right) - \left(1 + \frac{\nu}{\mu} \right) \left(2\psi + (1+\sigma)\lambda \frac{\nu}{\mu} \right) M_{E1} \right] M_{E2}. \tag{4.15}$$

These roots can be conveniently estimated by recalling that $M_{E2} \in [a_1 M_{E1}, \min\{a_k M_{E1}, \alpha_1\}]$ and computing the zeros of (4.15) at each of the extremes of such bounding interval, thereby obtaining the following three solutions:

$$\begin{aligned}
M_{E1}^A &= \frac{\psi(\alpha_1 + a_1) + \lambda \frac{\nu}{\mu}(\alpha_1 + \sigma a_1) - \nu}{a_1 (1 + \frac{\nu}{\mu}) \left(2\psi + (1+\sigma)\lambda \frac{\nu}{\mu} \right)}, \\
M_{E1}^B &= \frac{\psi(\alpha_1 + a_k) + \lambda \frac{\nu}{\mu}(\alpha_1 + \sigma a_k) - \nu}{a_k (1 + \frac{\nu}{\mu}) \left(2\psi + (1+\sigma)\lambda \frac{\nu}{\mu} \right)}, \\
M_{E1}^C &= \frac{\left(\psi + \lambda \sigma \frac{\nu}{\mu} \right) \alpha_1}{\frac{\nu}{\mu} \alpha_1 \left(2\psi + (1+\sigma)\lambda \frac{\nu}{\mu} \right) + \alpha_1 \left(\psi + \lambda \sigma \frac{\nu}{\mu} \right) + \nu}.
\end{aligned} \tag{4.16}$$

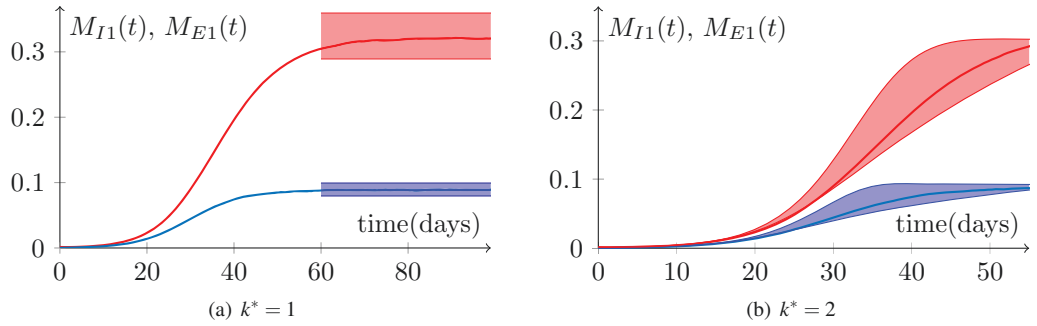
Using (4.14), these three solutions yield

$$M_{I1}^A = \frac{\nu}{\mu} M_{E1}^A, \quad M_{I1}^B = \frac{\nu}{\mu} M_{E1}^B, \quad \text{and} \quad M_{I1}^C = \frac{\nu}{\mu} M_{E1}^C. \tag{4.17}$$

In order to establish consistent bounds for the endemic equilibrium, roots (4.16) should be sorted. This can be done by analyzing the dependence of M_{E1} on M_{E2}. Specifically, we find from the first equation of (4.13) that the upper-bound for M_{E2} yields an upper-bound also for M_{E1} when $2\psi M_{E1} + (\lambda + \psi) M_{I1} < \psi$, otherwise it yields a lower-bound for M_{E1}. As a consequence, the following intervals for the endemic state can be derived:

- if $\alpha_1 \psi + \alpha_1 \lambda \nu / \mu > \nu$, then $\bar{M}_{E1} \in [\max\{M_{E1}^B, M_{E1}^C\}, M_{E1}^A]$ and $\bar{M}_{I1} \in [\max\{M_{I1}^B, M_{I1}^C\}, M_{I1}^A]$;
- if $\alpha_1 \psi + \alpha_1 \lambda \nu / \mu < \nu$, then $\bar{M}_{E1} \in [M_{E1}^A, \min\{M_{E1}^B, M_{E1}^C\}]$ and $\bar{M}_{I1} \in [M_{I1}^A, \min\{M_{I1}^B, M_{I1}^C\}]$; and
- if $\alpha_1 \psi + \alpha_1 \lambda \nu / \mu = \nu$, then we can explicitly compute

$$\bar{M}_{E1} = \frac{\psi + \lambda \sigma \frac{\nu}{\mu}}{\left(1 + \frac{\nu}{\mu} \right) \left(2\psi + (1+\sigma)\lambda \frac{\nu}{\mu} \right)} \quad \text{and} \quad \bar{M}_{I1} = \frac{\frac{\nu}{\mu}\psi + \lambda \sigma \frac{\nu^2}{\mu^2}}{\left(1 + \frac{\nu}{\mu} \right) \left(2\psi + (1+\sigma)\lambda \frac{\nu}{\mu} \right)}. \tag{4.18}$$



(a) $k^* = 1$ (b) $k^* = 2$

FIG. 7. Monte Carlo simulations (averaged over 200 trials) of the evolution of the fraction of infected individuals (solid red) and exposed individuals (solid blue) in the stochastic process for the flu case study [38] with unconscious infected individuals. Filled areas illustrate the bounds obtained with $k^* = 1$ (a) and $k^* = 2$ (b). The behavioral parameter σ has been set to 0.5, while parameters related to the exposed compartment, $\psi = 0.25$, and $\nu = 0.5$, are derived from [59]).

Figure 7(a) demonstrates the application of these bounds for the estimation of the endemic equilibrium in the flu case study with the addition of unconscious infected individuals. We observe that the addition of more compartments and transitions with respect to the SIS model in Fig. 4 does not challenge the feasibility of our approach. We are able to provide consistent bounds, whose accuracy depends on the specific model parameters. Also in this case, simulations with different model parameters, not reported here for brevity, yield bounds of different amplitude.

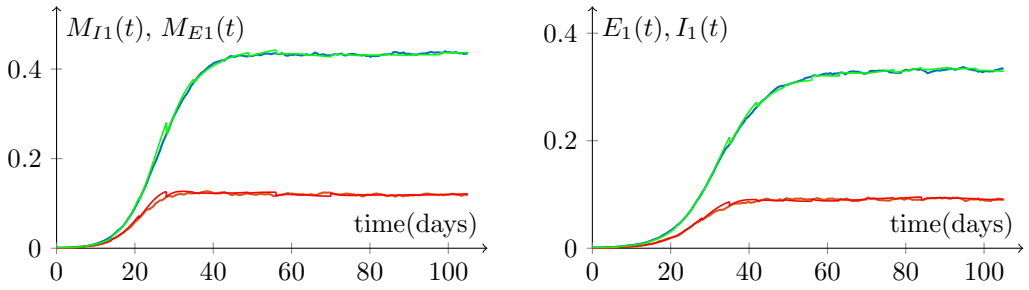
The epidemic curve, instead, can be estimated by using $k^* = 2$. Taking advantage of the simple structure of equation $\dot{M}_{I3}$ in (4.13), we use Remark 3.4 to construct the corresponding ODI, whereas the bounds on the macroscopic variables in (3.12) are only used to bound $a_1 M_{E2} \leqslant M_{E3} \leqslant \min\{\alpha_2, a_k M_{E2}\}$. Thus, the reduced ODI system derived from (4.13) comprises three ODEs, namely those governing the evolution of M_{E1}, M_{I1} and M_{I2}, and two ODIs, governing the evolution of M_{E2} and M_{I3}.

The upper- and lower-bounds on the first-order macroscopic variables M_{E1} and M_{I1} are derived by studying their dependence on M_{E3}. Considering the equation governing M_{E1} from (4.13), we note that if $2\psi M_{E1} + (\lambda + \psi) M_{I1} > \psi$, then M_{E1} has more marked increase when M_{E2} is smaller, and vice versa; the opposite behavior is exhibited if $2\psi M_{E1} + (\lambda + \psi) M_{I1} < \psi$. From the equation governing M_{E2}, we observe that M_{E2} has more marked increase when M_{E3} or M_{I3} are smaller. Finally, from the equation for M_{I3}, it is straightforward to note that M_{I3} has more marked increase when M_{E3} is larger. Hence, we conclude that when $2\psi M_{E1} + (\lambda + \psi) M_{I1} > \psi$, M_{E1} has more marked increase when M_{E3} is smaller, and vice versa.

In accordance with these observations, we define a continuous monotone function $\phi_{\varepsilon, M_{E2}}$ such that

$$\phi_{\varepsilon, M_{E2}}(M_{E1}, M_{I1}) = \begin{cases} a_1 M_{E2} & \text{if } 2\psi M_{E1} + (\lambda + \psi) M_{I1} < \psi - \varepsilon/2, \\ \min\{\alpha_2, a_k M_{E2}\} & \text{if } 2\psi M_{E1} + (\lambda + \psi) M_{I1} > \psi + \varepsilon/2, \end{cases} \tag{4.19}$$

and $\bar{\phi}_{\varepsilon, M_{E2}}(M_{E1}, M_{I1}) = a_1 M_{E2} + \min\{\alpha_2, a_k M_{E2}\} - \phi_{\varepsilon, M_{E2}}(M_{E1}, M_{I1})$. We introduce two pairs of ancillary ODEs and we couple each of them with the three equations for the evolution of M_{E1}, M_{I1}, and M_{I2} from (4.13), producing two five-dimensional systems of ODEs to bound the evolution of the fraction of



(a) $\sigma = 1$ (b) $\sigma = 0.5$

FIG. 8. Simulation of the evolution of the fraction of infected individuals (blue) and exposed ones (orange) in the stochastic process for the flu case study [38] with $\psi = 0.25$ and $\nu = 0.5$ [59], modeled as an SEIS, along with data-driven predictions over a weekly time-horizon (red and green, for exposed and infected, respectively) for different values of σ.

infected individuals (as $\varepsilon \to 0$). These two pairs of ancillary equations are, respectively,

$$\begin{aligned} \dot{M}_{E2} = \;& \alpha_2\psi M_{E1} - \nu M_{E2} + \alpha_2\lambda M_{I1} + \psi\alpha_1 M_{E2} + \lambda\sigma\alpha_1 M_{I2} - \psi M_{E2}^2 - \lambda M_{I1}M_{I3} \\ & -\psi\sigma M_{I2}^2 - \psi M_{E1}M_{I3} - (\psi+\lambda\sigma)M_{E2}M_{I2} - (\psi M_{E1} + \lambda M_{I1})\phi_{\varepsilon,M_{E2}}(M_{E1},M_{I1}), \\ \dot{M}_{I3} = \;& -\mu M_{I3} + \nu\phi_{\varepsilon,M_{I2}}(M_{E1},M_{I1}); \end{aligned} \tag{4.20}$$

and

$$\begin{aligned} \dot{M}_{E2} = \;& \alpha_2\psi M_{E1} - \nu M_{E2} + \alpha_2\lambda M_{I1} + \psi\alpha_1 M_{E2} + \lambda\sigma\alpha_1 M_{I2} - \psi M_{E2}^2 - \lambda M_{I1}M_{I3} \\ & -\psi\sigma M_{I2}^2 - \psi M_{E1}M_{I3} - (\psi+\lambda\sigma)M_{E2}M_{I2} - (\psi M_{E1} + \lambda M_{I1})\bar{\phi}_{\varepsilon,M_{E2}}(M_{E1},M_{I1}), \\ \dot{M}_{I3} = \;& -\mu M_{I3} + \nu\bar{\phi}_{\varepsilon,M_{I2}}(M_{E1},M_{I1}). \end{aligned} \tag{4.21}$$

The application of this technique for the flu case study is illustrated in Fig. 7(b), from which we obtain results comparable with those obtained for the simpler SIS model with behavioral changes due to infection, in Fig. 5.

In real-world applications, exposed individuals may not be conscious of their status, thereby hindering the possibility to measure the initial conditions for the macroscopic variables M_{Ej}. This problem can be tackled by estimating the fraction of exposed individuals (and higher-order macroscopic variables referring to the exposed compartment) through the combination of available historical data on breakouts of similar epidemics (for example, using the epidemic curve of previous seasons' breakouts of influenza).

Thus, short- and medium-time horizon predictions could be produced by using Algorithm 3.3, by considering a system of ODEs composed of $\dot{M}_{E1}$, $\dot{M}_{E2}$, $\dot{M}_{I1}$, $\dot{M}_{I2}$, and $\dot{M}_{I3}$, from (4.13), where each occurrence of the term M_{E3} is replaced with KM_{E1}, where $K = K^{(h)}$ is dynamically updated from an initial condition $K^{(0)} = \alpha_2$. At the beginning of each time period, epidemic data sampled at the population level are used to update the initial conditions and the proportionality constant K. Figure 8 demonstrates the accuracy of this technique for the flu case study, which does not seem to be affected by the increased number of compartments.

Clearly, the same arguments can be used to analyze more complex epidemic models, in which the infected compartment might be split into several sub-compartments, modeling progressive phases of the illness [51]. Moreover, the exposed compartment could be split into two distinct compartments, namely

latent and exposed. Susceptible individuals would enter the latent compartment immediately after the infection, while latent individuals are already infected but still not infectious, causing a natural delay in the dynamics. Latent individuals will spontaneously become exposed, being unconsciously infectious. For example, in the case of influenza, the incubation period lasts $1-3$ days, but only in the last day the infected individual becomes infectious [59].

4.3 *Presence of immunization*

Finally, we consider the case in which, after recovering, individuals become immune to the disease for a period of time, which can be finite (such as in influenza), or theoretically infinite (such as for smallpox and measles [60]).

At first, we consider the case in which immunization is temporary. The simplest model that includes this feature is the SIRS model [22]. Infected individuals, after recovering, enter the recovered compartment. Then, after a time that is the realization of a random variable exponentially distributed with parameter ν, they become susceptible again. Hence, ν^{-1} is the expected duration of the immunity. The limit case, $\nu \to 0$, models the case in which immunity is never lost (or the individual dies) and is treated later in this Section.

To cast this model in our framework, we set $\mathcal{A} = \{S, I, R\}$ and the activity reduction vector $\rho = [1, \sigma, 1]$. In addition, the pairwise interaction kernel is composed of the following matrices, where the first row/column refers to the compartment S, the second to I, and the third one to R:

$$\Lambda^{(S)} = \Lambda^{(R)} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}, \quad \Lambda^{(E)} = \begin{pmatrix} 1-\lambda & \lambda & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}. \tag{4.22}$$

The spontaneous transition rates matrix is

$$\Theta = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & \mu \\ \nu & 0 & 0 \end{pmatrix}. \tag{4.23}$$

The thermodynamic limit (Lemma 2.2) yields the following system of ODEs:

$$\begin{aligned} \dot{\zeta}_{Ii} &= -\mu\zeta_{Ii} + \lambda(1-\zeta_{Ii}-\zeta_{Ri})(a_i \sum_{j\in\mathcal{K}} \eta_j \zeta_{Ij} + \sigma \sum_{j\in\mathcal{K}} \eta_j a_j \zeta_{Ij}), \\ \dot{\zeta}_{Ri} &= \mu\zeta_{Ii} - \nu\zeta_{Ri}, \end{aligned} \tag{4.24}$$

where the variables referring to the compartment S have been written as linear combinations of the other variables, according to Remark 2.1.

Next, the system in (4.24) can be rewritten in terms of the macroscopic variables M_{Ij} and M_{Rl}, using the change of variables in Lemma 3.1, as follows:

$$\begin{aligned} \dot{M}_{I1} &= (\lambda\alpha_1 - \mu)M_{I1} + \lambda\sigma M_{I2} - (1+\sigma)\lambda M_{I1}M_{I2} - \lambda M_{R2}M_{I1} - \lambda\sigma M_{R1}M_{I2}, \\ \dot{M}_{I2} &= \lambda\alpha_2 M_{I1} + (\lambda\alpha_1\sigma - \mu)M_{I2} - \lambda M_{I1}M_{I3} - \lambda\sigma M_{I2}^2 - \lambda M_{R3}M_{I1} - \lambda\sigma M_{R2}M_{I2}, \\ \dot{M}_{Ij} &= \lambda\alpha_j M_{I1} + \lambda\alpha_{j-1}\sigma M_{I2} - \mu M_{Ij} - \lambda M_{I1}M_{I(j+1)} - \lambda\sigma M_{I2}M_{Ij} - \lambda M_{R(j+1)}M_{I1} - \lambda\sigma M_{Rj}M_{I2}, \\ \dot{M}_{Rl} &= \mu M_{Il} - \nu M_{Rl}, \end{aligned} \tag{4.25}$$

with $j = 3, \dots k$ and $l \in \mathcal{K}$.

The epidemic threshold of the SIRS model is equal to the one computed for the SIS model with behavioral changes due to infection, reported in (4.3). In fact, the linearization process of (4.25) close to the disease-free equilibrium cancels the terms related to the macroscopic variables of the recovered compartment.

The estimation of the fraction of infected and recovered individuals at the endemic equilibrium can be performed similar to the SEIS model. First, the simplicity of the equations M_{Rl} suggests to consider an additional ODI for the evolution of $M_{R(k^*+1)}$ instead of the bounds for that variable. Then, we consider the reduced system of ODIs with $k^* = 1$ and we prove that, at the endemic equilibrium,

$$M_{R1} = \frac{\mu}{\nu} M_{I1} \quad \text{and} \quad M_{R2} = \frac{\mu}{\nu} M_{I2}. \tag{4.26}$$

Hence, the equilibrium conditions reduce to bound the root of the following expression:

$$(\lambda\alpha_1 - \mu)M_{I1} + \lambda\left[\sigma - (1+\sigma)\left(1+\frac{\mu}{\nu}\right)M_{I1}\right]M_{I2}, \tag{4.27}$$

with $M_{I2} \in [a_1 M_{I1}, \min\{a_k M_{I1}, \alpha_1\}]$. The analysis of the dependence of M_{I1} on M_{I2} is carried out similar to the case of the SIS model with behavioral changes due to infection. Thus, we obtain the following intervals:

$$\begin{aligned}
\bar{M}_{I1} &\in \left[\frac{\lambda(a_1+\alpha_1\sigma)-\mu}{(1+\sigma)(1+\frac{\mu}{\nu})\lambda a_1}, \min\left\{\frac{\lambda\alpha_1\sigma}{\lambda\alpha_1\sigma(1+\frac{\mu}{\nu})+\mu+\lambda\alpha_1\frac{\mu}{\nu}}, \frac{\lambda(a_k+\alpha_1\sigma)-\mu}{(1+\sigma)(1+\frac{\mu}{\nu})\lambda a_1}\right\}\right] && \text{if } \lambda\alpha_1 < \mu;\\
\bar{M}_{I1} &= \frac{\sigma}{(1+\sigma)\left(1+\frac{\mu}{\nu}\right)} && \text{if } \lambda\alpha_1 = \mu;\\
\bar{M}_{I1} &\in \left[\max\left\{\frac{\lambda\alpha_1\sigma}{\lambda\alpha_1\sigma(1+\frac{\mu}{\nu})+\mu+\lambda\alpha_1\frac{\mu}{\nu}}, \frac{\lambda(a_k+\alpha_1\sigma)-\mu}{(1+\sigma)(1+\frac{\mu}{\nu})\lambda a_k}\right\}, \frac{\lambda(a_1+\alpha_1\sigma)-\mu}{(1+\sigma)(1+\frac{\mu}{\nu})\lambda a_1}\right] && \text{if } \lambda\alpha_1 > \mu.
\end{aligned} \tag{4.28}$$

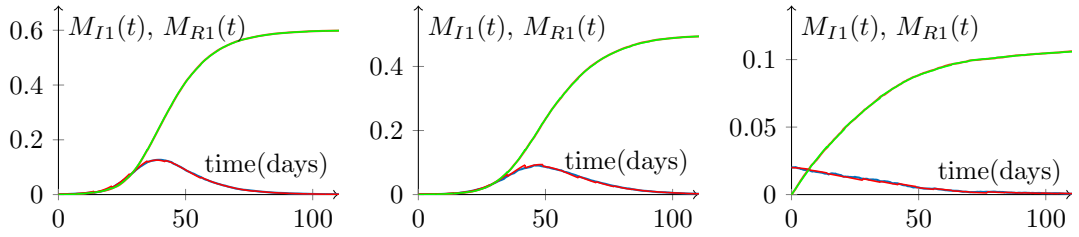
The corresponding intervals for $\bar{M}_{R1}$ can be written using (4.26).

The techniques from Section 3, both for the estimation of the epidemic curve and the derivation of accurate finite-time-horizon predictions, can be applied to this model, similar to the SEIS model. The epidemic curve can be estimated by first reducing (4.25) to a system of five ODIs. These five ODIs govern the evolution of M_{I1}, M_{I2}, M_{R1}, M_{R2}, and M_{R3}, using the bounds $a_1 M_{I2} \leqslant M_{I3} \leqslant \min\{\alpha_2, a_k M_{I2}\}$ from (3.12) and studying the dependence between the growth of M_{I1} and the magnitude of M_{I3}. Due to the absence of unconscious infected individuals, issues in the estimation of the initial conditions do not arise. In a similar vein, in our prediction algorithm we consider the five-dimensional system of ODEs obtained by substituting each occurrence of M_{I3} with KM_{I1}, where K is dynamically updated, along with the initial conditions of the macroscopic variables, at each iteration of the estimation process, from the available epidemic data.

The limit case in which immunization is permanent (for example, in the case of Ebola Virus Disease, measles, and smallpox [24, 60]) is obtained by setting $\nu = 0$. This model goes under the name of the susceptible-infected-recovered (SIR) model [22]. In the case of the SIR model, the system always converges to a disease-free state, in which some of the individuals are susceptible and some others, after having contracted the disease, recover (or are removed from the system, such as in the case of deaths).

To analyze the SIR model, we let $\nu = 0$ in (4.24), obtaining

$$\begin{aligned}
\dot{\zeta}_{Ii} &= -\mu\zeta_{Ii} + \lambda(1-\zeta_{Ii}-\zeta_{Ri})(a_i \sum_{j\in\mathcal{K}} \eta_j \zeta_{Ij} + \sigma \sum_{j\in\mathcal{K}} \eta_j a_j \zeta_{Ij}),\\
\dot{\zeta}_{Ri} &= \mu\zeta_{Ii};
\end{aligned} \tag{4.29}$$



(a) $\sigma = 1$ (b) $\sigma = 0.5$ (c) $\sigma = 0$

FIG. 9. Simulation of the evolution of the fraction of infected individuals (blue) and recovered ones (orange) in the stochastic process for the flu case study [38] modeled as an SIR for different values of σ, along with data-driven predictions over a weekly time-horizon (red and green, for infected and recovered, respectively).

and the system of ODEs in (4.25) reads

$$\begin{aligned}
\dot{M}_{I1} &= (\lambda\alpha_1 - \mu)M_{I1} + \lambda\sigma M_{I2} - (1+\sigma)\lambda M_{I1}M_{I2} - \lambda M_{R2}M_{I1} - \lambda\sigma M_{R1}M_{I2},\\
\dot{M}_{I2} &= \lambda\alpha_2 M_{I1} + (\lambda\alpha_1\sigma - \mu)M_{I2} - \lambda M_{I1}M_{I3} - \lambda\sigma M_{I2}^2 - \lambda M_{R3}Ix_1 - \lambda\sigma M_{R2}M_{I2},\\
\dot{I}_j &= \lambda\alpha_j M_{I1} + \lambda\alpha_{j-1}\sigma M_{I2} - \mu M_{Ij} - \lambda M_{I1}M_{I(j+1)} - \lambda\sigma M_{I2}M_{Ij} - \lambda M_{R(j+1)}M_{I1} - \lambda\sigma M_{Rj}M_{I2},\\
\dot{M}_{Rl} &= \mu M_{Il},
\end{aligned} \tag{4.30}$$

with $j = 3,\dots,k$ and $l \in \mathcal{K}$.

The technique used in the analysis of the SIRS model for the estimation of the epidemic curve can be adapted to this limit case, considering $k^* = 2$ with an additional ODI derived from the equation $\dot{M}_{R3} = \mu M_{I3}$, instead of the bounds on M_{R3}, according to Remark 3.4. Moreover, very accurate data-driven finite-time-horizon predictions can be obtained using Algorithm 3.3, as explained above for the SIRS model. As shown in Fig. 9, therein we apply this technique to the flu case study for different values of the behavioral parameter σ. Our data-driven finite-horizon prediction technique seems to offer better performance in the presence of immunization dynamics.

5. Conclusion

In this paper, we have developed a new continuous-time theory for stochastic epidemic models on ADNs with a discrete activity distribution, which allows for an analytical treatment of the spreading dynamics. One of the main contributions of this work is the derivation of a deterministic approximation valid in the thermodynamic limit of large population. We have developed several empowering mathematical techniques for the analysis of such models, which can be used to generate accurate predictions of the evolution of the epidemic process. We have demonstrated the possibility of studying a number of prominent modeling variants (such as a model with behavioral changes due to infection, SEIS, SIRS, and SIR) within our new framework. Numerical results confirm the effectiveness of our approach, which i) reproduces the dynamics of discrete-time Monte Carlo simulations avoiding the confounds associated with the selection of the discrete sampling time; and ii) affords the implementation of mathematical techniques to analyze the endemic equilibrium and of the evolution the spreading dynamics.

Several improvements to our model are envisaged and will be subject of further research. First, future work should seek to clarify the relationship between our predictions based on analytical bounds and model parameters, which is presently opaque. Second, future work should elucidate the truncating

procedure toward enhancing accuracy and minimize computational costs in treating several macroscopic variables. Third, a further effort should be devoted to cast the finite-time-horizon prediction techniques in an optimization context, to ease the selection of sampling periods and assess accuracy a priori. Finally, we aim at further empowering our framework with realistic phenomena, relevant from an epidemic modeling point of view, such as memory processes in link formation, spatial locality, formation of communities, and non-Markovianity.

The achievement of the envisaged research efforts would lead to the development of a general and powerful paradigm for the analytical treatment of realistic epidemic processes on time-varying networks, otherwise tractable only through extensive numerical simulations.

Funding

This work was supported by National Science Foundation under grant No. CMMI-1561134, the Army Research Office under grant No. W911NF-15-1-0267, with Drs. A. Garcia and S.C. Stanton as program managers, and Compagnia di San Paolo.

References

[1] HOLME, P. & SARAMÄKI, J. (2012) Temporal networks. *Phys. Rep.*, **519**, 97–125.

[2] PERRA, N., GONÇALVES, B., PASTOR-SATORRAS, R. & VESPIGNANI, A. (2012) Activity driven modeling of time varying networks.. *Sci. Rep.*, **2**, 469.

[3] RIZZO, A. & PORFIRI, M. (2016) Innovation diffusion on time-varying activity driven networks. *EPJ B*, **89**, 20.

[4] TIZZONI, M., SUN, K., BENUSIGLIO, D., KARSAI, M. & PERRA, N. (2015) The scaling of human contacts and epidemic processes in metapopulation networks. *Sci. Rep.*, **5**, 15111.

[5] SUN, K., BARONCHELLI, A. & PERRA, N. (2015) Contrasting effects of strong ties on SIR and SIS processes in temporal networks. *EPJ B*, **88**, 326.

[6] BARRAT, A., CATTUTO, C., COLIZZA, V., GESUALDO, F., ISELLA, L., PANDOLFI, E., PINTON, J. F., RAVA, L., RIZZO, C., ROMANO, M., STEH, J., TOZZI, A. E. & VAN DEN BROECK, W. (2013) Empirical temporal networks of face-to-face human interactions. *EPJ ST*, **222**, 1295–1309.

[7] IACOBELLI, G. & FIGUEIREDO, D. R. (2017) Edge-attractor random walks on dynamic networks. *J. Complex Netw.*, **5**, 84–110.

[8] RIZZO, A., FRASCA, M. & PORFIRI, M. (2014) Effect of individual behavior on epidemic spreading in activity driven networks. *Phys. Rev. E*, **90**, 042801.

[9] REN, G. & WANG, X. (2014) Epidemic spreading in time-varying community networks. *Chaos*, **24**, 023116.

[10] FRASCA, M., BUSCARINO, A., RIZZO, A., FORTUNA, L. & BOCCALETTI, S. (2006) Dynamical network model of infective mobile agents. *Phys. Rev. E*, **74**, 036110.

[11] VOLZ, E. & MEYERS, L. A. (2008) Epidemic thresholds in dynamic contact networks. *J. R. Soc. Int.*, **6**, 233–241.

[12] PASTOR-SATORRAS, R., CASTELLANO, C., VAN MIEGHEM, P. & VESPIGNANI, A. (2015) Epidemic processes in complex networks. *Rev. Mod. Phys.*, **87**, 925–979.

[13] VALDANO, E., FERRERI, L., POLETTO, C. & COLIZZA, V. (2015) Analytical computation of the epidemic threshold on temporal networks. *Phys. Rev. X*, **5**, 021005.

[14] OGURA, M. & PRECIADO, V. M. (2016) Stability of spreading processes over time-varying large-scale networks. *IEEE Trans. Netw. Sci. Eng.*, **3**, 44–57.

[15] VESTERGAARD, C., VALDANO, E., GÉNOIS, M., POLETTO, C., COLIZZA, V. & BARRAT, A. (2016) Impact of spatially constrained sampling of temporal contact networks on the evaluation of the epidemic risk. *Eur. J. of Appl. Math.*, **27**, 941–957.

[16] KOHER, A., LENTZ, H., HÖVEL, P. & SOKOLOV, I. (2016) Infections on temporal networks: a matrix-based approach. *PLOS One*, **11**, e0151209.

[17] BRAUNSTEIN, A. & INGROSSO, A. (2016) Inference of causality in epidemics on temporal contact networks.. *Sci. Rep.*, **6**, 27538.

[18] ANDERSSON, H. & BRITTON, T. (1998) Heterogeneity in epidemic models and its effect on the spread of infection. *J. Appl. Probab.*, **35**, 651–661.

[19] PASTOR-SATORRAS, R. & VESPIGNANI, A. (2001) Epidemic spreading in scale-free networks. *Phys. Rev. Lett.*, **86**, 3200.

[20] PARSHANI, R., CARMI, S. & HAVLIN, S. (2010) Epidemic threshold for the susceptible-infectious-susceptible model on random networks. *Phys. Rev. Lett.*, **104**, 258701.

[21] LIU, S., PERRA, N., KARSAI, M. & VESPIGNANI, A. (2014) Controlling contagion processes in activity driven networks. *Phys. Rev. Lett.*, **112**, 118702.

[22] BRAUER, F. & CASTILLO-CHAVEZ, C. (2011) *Mathematical models in population biology and epidemiology*. New York: Springer.

[23] STARNINI, M. & PASTOR-SATORRAS, R. (2014) Temporal percolation in activity-driven networks. *Phys. Rev. E*, **89**, 032807.

[24] RIZZO, A., PEDALINO, B. & PORFIRI, M. (2016) A network model for Ebola spreading. *J. Theor. Biol.*, **394**, 212 – 222.

[25] LEI, Y., JIANG, X., GUO, Q., MA, Y., LI, M. & ZHENG, Z. (2016) Contagion processes on the static and activity-driven coupling networks. *Phys. Rev. E*, **93**, 032308.

[26] LIU, C., ZHOU, L.-X., FAN, C.-J., HUO, L.-A. & TIAN, Z.-W. (2015) Activity of nodes reshapes the critical threshold of spreading dynamics in complex networks. *Phys. A*, **432**, 269–278.

[27] ZOU, Y., DENG, W., LI, W. & CAI, X. (2016) A study of epidemic spreading on activity-driven networks. *Int. J. Mod. Phys. C*, **27**, 1650090.

[28] AOKI, T., ROCHA, L. E. C. & GROSS, T. (2016) Temporal and structural heterogeneities emerging in adaptive temporal networks. *Phys. Rev. E*, **93**, 040301.

[29] GONZÁLEZ, M. C., HIDALGO, C. A. & BARABÁSI, A. (2008) Understanding individual human mobility patterns. *Nature*, **453**, 779–782.

[30] KARSAI, M., PERRA, N. & VESPIGNANI, A. (2014) Time varying networks and the weakness of strong ties. *Sci. Rep.*, **4**, 4001.

[31] STARNINI, M., BARONCHELLI, A. & PASTOR-SATORRAS, R. (2016) Model reproduces individual, group and collective dynamics of human contact networks. *Soc. Networks*, **47**, 130–137.

[32] ONAGA, T., GLEESON, J. P. & MASUDA, N. (2017) Concurrency-Induced Transitions in Epidemic Dynamics on Temporal Networks. *Phys. Rev. Lett.*, **119**, 108301.

[33] HAN, D., SUN, M. & LI, D. (2015) Epidemic process on activity-driven modular networks. *Phys. A*, **432**, 354–362.

[34] LIU, M., WANG, W., LIU, Y., TANG, M., CAI, S. & ZHANG, H. (2017) Social contagions on time-varying community networks. *Phys. Rev. E*, **95**, 052306.

[35] POZZANA, I., SUN, K. & PERRA, N. (2017) Epidemic Spreading on Activity-Driven Networks with Attractiveness. *preprint, arXiv:1703.02482*.

[36] ALESSANDRETTI, L., SUN, K., BARONCHELLI, A. & PERRA, N. (2017) Random walks on activity-driven networks with attractiveness. *Phys. Rev. E*, **95**, 052318.

[37] UBALDI, E., PERRA, N., KARSAI, M., VEZZANI, A., BURIONI, R. & VESPIGNANI, A. (2016) Asymptotic theory of time-varying social networks with heterogeneous activity and tie allocation. *Sci. Rep.*, **6**, 35724.

[38] ZINO, L., RIZZO, A. & PORFIRI, M. (2016) Continuous-time discrete-distribution theory for activity-driven networks. *Phys. Rev. Lett.*, **117**, 228302.

[39] GILLESPIE, D. T. (1976) A general method for numerically simulating the stochastic time evolution of coupled chemical reactions. *J. Comput. Phys.*, **22**, 403–434.

[40] VESTERGAARD, C. L. & GÉNOIS, M. (2015) Temporal Gillespie algorithm: fast simulation of contagion processes on time-varying networks. *PLoS Comput. Biol.*, **11**, e1004579.

[41] RIBEIRO, B., PERRA, N. & BARONCHELLI, A. (2013) Quantifying the effect of temporal resolution on time-varying networks.. *Sci. Rep.*, **3**, 3006.

[42] FENNELL, P. G., MELNIK, S. & GLEESON, J. P. (2016) Limitations of discrete-time approaches to continuous-time contagion dynamics. *Phys. Rev. E*, **94**, 052125.

[43] AJELLI, M., GONÇALVES, B., BALCAN, D., COLIZZA, V., HU, H., RAMASCO, J., MERLER, S. & VESPIGNANI, A. (2010) Comparing large-scale computational approaches to epidemic modeling: agent-based versus structured metapopulation models.. *BMC Inf. Dis.*, **10**, 190.

[44] BAJARDI, P., POLETTO, C., RAMASCO, J., TIZZONI, M., COLIZZA, V. & VESPIGNANI, A. (2011) Human mobility networks, travel restrictions, and the global spread of 2009 H1N1 pandemic. *PLOS One*, **6**, e16591.

[45] GOMES, M., PASTORE, A., ROSSI, L., CHAO, D., LONGINI, I., HALLORAN, M. & VESPIGNANI, A. (2014) Assessing the International Spreading Risk Associated with the 2014 West African Ebola Outbreak. *PLOS Cur. Out.*, (September 2).

[46] MERLER, S., AJELLI, M., FUMANELLI, L., GOMES, M., PIONTTI, A., ROSSI, L., CHAO, D., LONGINI, I., HALLORAN, M. & VESPIGNANI, A. (2015) Spatiotemporal spread of the 2014 outbreak of Ebola virus disease in Liberia and the effectiveness of non-pharmaceutical interventions: a computational modelling analysis. *The Lancet Inf. Dis.*, **3099**, 204–211.

[47] BORGS, C., CHAYES, J., GANESH, A. & SABERI, A. (2010) How to distribute antidote to control epidemics. *Random Struct. Algor.*, **37**, 204–222.

[48] LEYFFER, S. & SAFRO, I. (2013) Fast response to infection spread and cyber attacks on large-scale networks. *J. Complex Netw.*, **1**, 183–199.

[49] DRAKOPOULOS, K., OZDAGLAR, A. & TSITSIKLIS, J. N. (2014) An efficient curing policy for epidemics on graphs. *IEEE Trans. Netw. Sci. Eng.*, **1**, 67–75.

[50] NOWZARI, C., PRECIADO, V. & PAPPAS, G. (2017) Optimal resource allocation for control of networked epidemic models. *IEEE Trans. Control Netw. Syst.*, **4**, 159–169.

[51] LIN, Y., LUI, J. C. S., JUNG, K. & LIM, S. (2014) Modelling multi-state diffusion process in complex networks: theory and applications. *J. Complex Netw.*, **2**, 431–459.

[52] KURTZ, T. G. (1981) *Approximation of population processes*, vol. 36. Philadelphia: SIAM.

[53] KEELING, M. J. & EAMES, K. T. D. (2005) Networks and epidemic models. *J. R. Soc. Interface*, **2**, 295–307.

[54] VERELST, F., WILLEM, L. & BEUTELS, P. (2016) Behavioural change models for infectious disease transmission: a systematic review (2010-2015). *J. R. Soc. Interface*, **13**.

[55] AIELLO, W., CHUNG, F. & LU, L. (2000) A random graph model for massive graphs. in *Proceedings of the 32nd Annual ACM Symposium on Theory of Computing*, pp. 171–180. New York: Association for Computing Machinery.

[56] BAILEY, N. (1975) *The mathematical theory of infectious diseases and its applications. 2nd edn*. London: Griffin.

[57] BERETTA, E. & CAPASSO, V. (1986) On the general structure of epidemic systems. Global asymptotic stability. *Comput. Math. Appl.*, **12**, 677–694.

[58] FOXALL, E. (2016) The SEIS model, or, the contact process with a latent stage. *J. Appl. Probab.*, **53**, 783–801.

[59] COX, N. J. & SUBBARAO, K. (1999) Influenza. *Lancet*, **354**, 1277 – 1282.

[60] FERGUSON, N. M., KEELING, M. J., JOHN EDMUNDS, W., GANI, R., GRENFELL, B. T., ANDERSON, R. M. & LEACH, S. (2003) Planning for smallpox outbreaks. *Nature*, **425**, 681–685.

[61] HETHCOTE, H., ZHIEN, M. & SHENGBING, L. (2002) Effects of quarantine in six endemic models for infectious diseases. *Math. Biosci.*, **180**, 141–160.

[62] KATO, F., TAINAKA, K.-I., SONE, S., MORITA, S., IIDA, H. & YOSHIMURA, J. (2011) Combined effects of prevention and quarantine on a breakout in SIR model. *Sci. Rep.*, **1**, 10.