Controlling Epidemics Through Optimal Allocation of Test Kits and Vaccine Doses Across Networks

Mingtao Xia, Lucas Böttcher, and Tom Chou

***Abstract*—Efficient testing and vaccination protocols are critical aspects of epidemic management. To study the optimal allocation of limited testing and vaccination resources in a heterogeneous contact network of interacting susceptible, infected, and recovered individuals, we present a degree-based testing and vaccination model for which we derive optimal policies using control-theoretic methods. Within our framework, we find that optimal intervention policies first target high-degree nodes before shifting to lower-degree nodes in a time-dependent manner. Using such optimal policies, it is possible to delay outbreaks and reduce incidence rates to a greater extent than uniform and reinforcement-learning-based interventions, particularly on certain scale-free networks.**

***Index Terms*—Epidemics, infection networks, optimal control, reinforcement learning, testing, vaccination.**

I. Introduction

LIMITING the spread of novel pathogens such as SARS-CoV-2 requires efficient testing [1], [2] and quarantine strategies [3], especially when vaccines are not available or effective [4]. Even if effective vaccines are available at scale, their population-wide distribution is a complex and time-consuming endeavor, influenced by, for example, age-structure [5]–[7], vaccine hesitancy [8], and different objectives [9].

Until a sufficient level of immunity within a population is reached, distancing and quarantine policies can also be used to help slow the spread and evolutionary dynamics [10] of infectious diseases. Epidemic modeling and control-theoretic approaches are useful for identifying both efficient testing and vaccination policies. For an epidemic model of SARS-CoV-2 transmission, Pontryagin's maximum principle (PMP) has been used to derive optimal distancing and testing strategies that minimize the number of COVID-19 cases and intervention costs [11], [12]. Optimal control theory has also been applied to a multi-objective control problem that uses isolation and vaccination to limit epidemic size and duration [13]. These recent investigations describe the underlying infectious disease dynamics through compartmental models without underlying network structure, meaning that all interactions among different individuals are assumed to be homogeneous.

Multicompartment models that may be associated with contact networks have been investigated. For example, optimal vaccination strategies have been derived for a rapidly spreading disease in a highly mobile multicompartment susceptible-infected-recovered (SIR) model using PMP [14]. The application of optimal control methods and PMP to heterogeneous node-based susceptible-infected-recovered-susceptible (SIRS) models were also studied in the context of multiplex networks [15] and rumor spreading [16].

Complementing these control-theory-based investigations, reinforcement learning (RL) has been recently used to identify infectious high-degree nodes ("superspreaders") in temporal networks [17]. It has been found that RL was able to outperform intervention policies derived from purely structural node characterizations that are, for instance, based on centrality measures [17]. However, these RL methods could only be applied to rather small networks with about 400 nodes. For social networks describing much larger populations, early work by May and Anderson employed effective degree models to study the population-level dynamics of human immunodeficiency virus (HIV) infections [18]. These degree-based models and later variants [19]–[21] did not account for degree correlations. Effective degree models for susceptible-infected-susceptible (SIS) dynamics with degree correlations were derived in [22] and applied to SIR dynamics in [23]. A further generalization of these methods to model SIR dynamics with networked and well-mixed transmission pathways was presented in [24]. For a detailed summary of degree-based epidemic models, see [25].

In this work, we focus on formulating both optimal control and RL-based target policies on a degree-based epidemic model [26] that is constrained only by the maximum degree and not by the system size (*i.e.*, number of nodes). We construct effective control strategies to slow down disease spread across heterogeneous network models which include both degree distributions and higher-order correlations of the degree distribution. Our approach is not limited by size as agent-based models are [17], is simpler because we do not

Manuscript received July 19, 2021; revised December 6, 2021; accepted January 14, 2022. Date of publication February 18, 2022; date of current version May 23, 2022. This work was supported in part by NSF under Grant DMS-1814364 and in part by Army Research Office under Grant W911NF-18-1-0345. Recommended for acceptance by Prof. Xianbin Cao. *(Corresponding author: Tom Chou.)*

Mingtao Xia is with the Department of Mathematics, University of California, Los Angeles, Los Angeles, CA 90095 USA (e-mail: xiamingtao97@g.ucla.edu).

Lucas Böttcher is with the Department of Computational Medicine, University of California, Los Angeles, Los Angeles, CA 90095 USA, and also with the Frankfurt School of Finance and Management, 60322 Frankfurt, Germany (e-mail: L.Boettcher@fs.de).

Tom Chou is with the Department of Computational Medicine and Mathematics, University of California, Los Angeles, Los Angeles, CA 90095 USA (e-mail: tomchou@ucla.edu).

Digital Object Identifier 10.1109/TNSE.2022.3144624

resolve interpersonal contact times or other individual details, and is thus easier to solve. On the other hand, unlike simple multicompartment epidemic control models [11], [13], we take into account a heterogeneous contact network and control measures that depend on both time and node degree.

In the next section, we propose and justify a degree-based epidemic testing and quarantining model. An optimal control framework for this model is presented in Section III and, given limited testing resources, an optimal testing strategy is computed. We extend the same underlying disease model to include vaccination in Section IV and derive optimal vaccination strategies that minimize infection given a limited vaccination rate. We summarize our results and discuss how they depend on network and dynamical features of the model in Section V. For comparison, we also present in Appendix C a reinforcement-learning-based algorithm that is able to approximate optimal testing strategies for the model introduced in Section II.

Finally, we implemented a stochastic Monte-Carlo simulation of disease transmission, testing, and vaccination on networks. By using the optimal strategies computed using the PMP on ODE-based deterministic models, we find significant differences in the stochastic model. In Appendix D, we show that these differences arise from higher correlations in network connectivity that arise in the discrete stochastic model used.

II. Degree-Based Epidemic and Testing Model

For the formulation of optimal testing policies that allocate testing resources to different individuals in a contact network, we adopt an effective degree model of SIR dynamics with testing in a static network of N nodes. Nodes represent individuals, and edges between nodes represent corresponding contacts. Therefore, the degree of a node represents the number of its contacts. If K is the maximum degree across all nodes, we can divide the population into K distinct subpopulations, each of size N_k $(k = 1, 2, \ldots, K)$ such that all nodes in the k^{th} group have degree k. Therefore, $N = \sum_{k=1}^{K} N_k$.

In our epidemic model, we distinguish between untested and tested infected individuals. Let $S_k(t)$, $I_k^{\mathrm{u}}(t)$, $I_k^*(t)$, and $R_k(t)$ denote the numbers of susceptible, untested infected, tested infected, and recovered nodes with degree k at time t, respectively. Since these subpopulations together represent the entire population (the total number of nodes N), both N and N_k are constants in our model. Their values satisfy the normalization condition $S_k + I_k^{\mathrm{u}} + I_k^* + R_k = N_k$. The corresponding fractions are

$$\begin{aligned} s_k(t) &= S_k(t)/N, \quad i_k^{\mathrm{u}}(t) = I_k^{\mathrm{u}}(t)/N, \\ i_k^*(t) &= I_k^*(t)/N, \quad r_k(t) = R_k(t)/N, \end{aligned} \tag{1}$$

such that $\sum_k (s_k + i_k^{\mathrm{u}} + i_k^* + r_k) = 1$. Using an effective-degree approach [18], [24], we describe the evolution of the above subpopulations by

$$\frac{\mathrm{d}s_k(t)}{\mathrm{d}t} = -k s_k(t) \sum_{\ell=1}^{K} \frac{P(\ell|k)}{P(\ell)} \left(\beta^{\mathrm{u}} i_\ell^{\mathrm{u}}(t) + \beta^* i_\ell^*(t)\right), \tag{2}$$

$$\begin{aligned} \frac{\mathrm{d}i_k^{\mathrm{u}}(t)}{\mathrm{d}t} &= k s_k(t) \sum_{\ell=1}^{K} \frac{P(\ell|k)}{P(\ell)} \left(\beta^{\mathrm{u}} i_\ell^{\mathrm{u}}(t) + \beta^* i_\ell^*(t)\right) \\ &\quad - \gamma^u i_k^{\mathrm{u}}(t) - \frac{f_k(t)}{N_k} i_k^{\mathrm{u}}(t), \end{aligned} \tag{3}$$

$$\frac{\mathrm{d}i_k^*(t)}{\mathrm{d}t} = -\gamma^* i_k^*(t) + \frac{f_k(t)}{N_k} i_k^{\mathrm{u}}(t), \tag{4}$$

$$\frac{\mathrm{d}r_k(t)}{\mathrm{d}t} = \gamma^{\mathrm{u}} i_k^{\mathrm{u}}(t) + \gamma^* i_k^*(t), \tag{5}$$

where $P(\ell) = N_\ell/N$ is the degree distribution. $P(\ell|k)$ is the conditional probability that a chosen node with degree k is connected to a node with degree ℓ. By defining $E_{\ell,k}$ as the number of edges connecting a node with degree k with another node with degree ℓ in a given network, the conditional probability can be directly evaluated as $P(\ell|k) = E_{\ell,k}/(kN_k)$. Our degree-based formulation of SIR dynamics with testing, (2)–(5), is an approximation of the full node-based dynamics assuming that nodes of the same degree are equally likely to be infected at any given time [26].

Susceptible individuals become infected through contact with untested and tested infected individuals at rates β^{u} and β^*, respectively. Untested and tested infected individuals recover at rates γ^{u} and γ^*, respectively. Differences in the recovery rates γ^{u} and γ^* reflect differences in disease severity of and treatment options for untested and tested infected individuals. Once recovered, individuals develop long-lasting immunity that protects them from reinfection. Temporary immunity can be easily modeled by using an SIS type model with or without delays. Reduced transmissibility of tested infected (and potentially quarantined) individuals corresponds to setting $\beta^* \ll \beta^{\mathrm{u}}$.

The testing rate of nodes with degree k is defined as $f_k(t)$, such that $f_k(t)\Delta t$ is the total number of tests given to nodes with degree k in time window Δt. Tests given to recovereds, susceptibles, and already-tested infecteds do not lead to quarantining and will not affect the disease dynamics. However, a fraction $I_k^{\mathrm{u}}/(S_k + I_k^{\mathrm{u}} + I_k^* + R_k) \equiv I_k^{\mathrm{u}}/N_k$ of these $f_k(t)\Delta t$ tests will be administered to untested infecteds. Once infected nodes have been identified by testing, they can be quarantined and removed from the disease transmission dynamics. If infected individuals who already have been tested strictly avoid future testing, more tests will be available for the other subpopulations, increasing the rate at which the remaining untested infecteds will be tested. In this case, the fraction of tests administered to untested infecteds is modified: $I_k^{\mathrm{u}}/(S_k + I_k^{\mathrm{u}} + R_k) \equiv I_k^{\mathrm{u}}/(N_k - I_k^*)$. After normalizing by the total population N to write tested fractions in terms of (1), the testing term becomes $-f_k(t) i_k^{\mathrm{u}}/N_k$ (see (3)) or $-f_k(t) i_k^{\mathrm{u}}/[N_k(1 - I_k^*/N_k)]$, respectively.

Biased testing can also be represented by using a testing fraction of the form $I_k^{\mathrm{u}} e^b/(I_k^{\mathrm{u}} e^b + S_k + I_k^* + R_k)$, where $b > 0$ increases the fraction of tests given to infecteds. To correct for false-positive tests, (2)–(5) can be modified by including

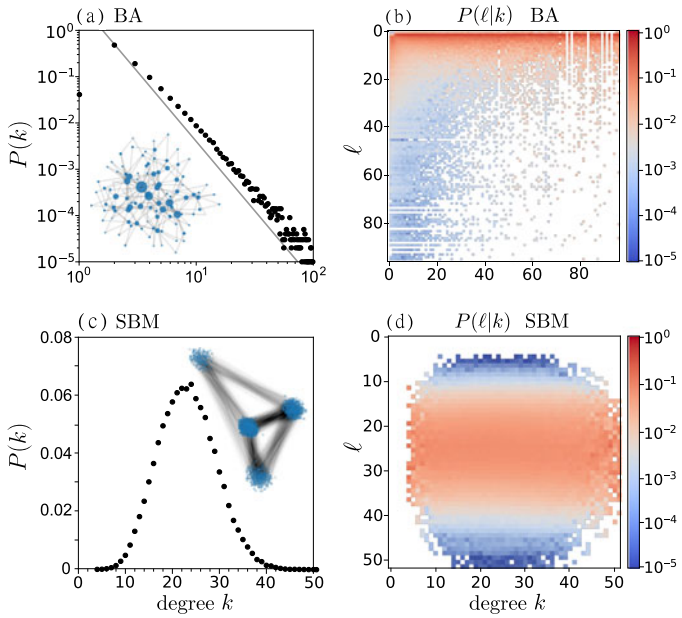

Fig. 1. Degree distribution of a Barabási–Albert network and a stochastic block model. (a) The degree distribution of a Barabási–Albert network with 99,817 nodes. To generate the network, we start with a dyad and iteratively add new nodes until we reach 100,000 nodes. Each new node has 2 edges that connect it to existing nodes using linear preferential attachment. Isolated nodes or nodes with degrees larger than 100 [32] are then removed from the network. The grey solid line is a guide-to-the-eye with slope -3 [30]. For illustration, the inset shows a realization of a Barabási–Albert network with 100 nodes. Node size scales with their betweenness centrality. (b) The conditional probability $P(\ell|k)$ associated with the Barabási–Albert network generated in (a). (c) The degree distribution of a stochastic block model with four blocks and 100,000 nodes. The inset shows a realization of a stochastic block model with 800 nodes, but using the same block probability matrix. (d) The conditional probability $P(\ell|k)$ associated with the SBM. In both (b) and (d), all elements that are strictly zero are uncolored.

an additional term that transfers the I_k^* population back to S_k. False negatives can be accounted for by a reduction in $f_k(t)/N_k$. For a detailed overview of statistical models that account for testing errors and bias, see [27], [28].

What remains is to assign network structures, extract $P(\ell|k)$ from them, and determine reasonable parameter values before calculating the optimal testing protocol $f_k(t)$. We apply our disease-control framework to (i) a Barabási-Albert (BA) network [29], [30] and (ii) a stochastic block model (SBM) [31] with four communities and a probability matrix

$$P = 10^{-4}\begin{pmatrix} 1 & 2 & 2 & 2 \\ 2 & 4 & 2 & 2 \\ 2 & 2 & 5 & 2 \\ 2 & 2 & 2 & 3 \end{pmatrix}. \tag{6}$$

These two network types exhibit properties, such as hub nodes with high degrees and community structure, that are observable in real-world contact networks [32], [33]. In the construction of the BA network, each new node is connected to 2 existing nodes using linear preferential attachment. Fig. 1(a) shows the degree distribution of a 99,817-node BA network that we use in this study. A heatmap of the conditional degree distribution matrix of the BA network with the degree distribution $P(k)$ shown in Fig. 1(a) is given in Fig. 1(b). The degree distribution and the conditional degree distribution matrix of the 100,000-node SBM network are shown in Fig. 1 (c) and (d), respectively. Taking into account empirical findings on the degree distributions in real-world contact networks [32], we use a degree cutoff of $k \leq K = 100$. We will use the specific configurations of the BA and SBM networks shown in Fig. 1 for our subsequent analysis of (2)-(5).

Next, to constrain the parameter values, we first invoke estimates of the basic reproduction number (*i.e.*, the average number of secondary cases that results from one case in a completely susceptible population), which for a network model is defined as [34], [35]

$$\mathcal{R}_0 = \rho(JV^{-1}) \tag{7}$$

in which $\rho(\cdot)$ is the largest eigenvalue (spectral radius), $V \equiv \mathrm{diag}(\gamma^{\mathrm{u}}) \in \mathbb{R}^{K\times K}$, and $J \in \mathbb{R}^{K\times K}$ is the Jacobian of the linearized dynamical system (see (2) and (3)) about the disease-free state with $s_k(t=0) = N_k/N$ and $f_k = 0$ corresponding to the initial, untested, and uncontrolled spread of the infection:

$$J_{ij} = iP(j|i)\frac{N_i}{N_j}\beta^{\mathrm{u}}, \quad i,j \leq K. \tag{8}$$

This "next-generation" method associates $\mathcal{R}_0$ with the largest eigenvalue inherent to the dynamical system. Additional expressions for $\mathcal{R}_0$ for an uncorrelated degree network are given in Appendix A.

Empirically, the basic reproduction number for COVID-19 varies across different regions. For the early outbreak in Wuhan [36], $\mathcal{R}_0$ was estimated to be 3.49, while for the early outbreak in Italy $\mathcal{R}_0 \sim 2.43 - 3.10$ [37]. Here we set $\mathcal{R}_0 = 4.5$ which was suggested in [38] as the basic reproduction number of early COVID-19 spread in the absence of any intervention. For a given value of the recovery rate γ^{u} of untested individuals, which can be inferred from empirical data [38], [39], we determine the transmissibility β^{u} by numerically solving $\mathcal{R}_0(\beta^{\mathrm{u}}) = 4.5$ for β^{u}. Our source codes are publicly available at https://gitlab.com/ComputationalScience/epidemic-control.

III. Allocating Limited Testing Resources

Without any testing constraints, it would be most effective for disease control to use a testing rate $f_k(t)$ sufficiently large to keep the fraction of untested individuals, $i_k^{\mathrm{u}}(t)$, close to zero. In general, the testing rates are constrained by

$$f_k^{\min} \leq \frac{f_k(t)}{N_k} \leq f_k^{\max}, \tag{9}$$

and the total testing rate is also bounded by availability and logistics of testing

$$\sum_{k=1}^{K} f_k(t) = F(t). \tag{10}$$

The goal is to determine, under these constraints, the function $f_k(t)$ or $f_k(t)/N_k$ that most effectively reduces the total

number of infections. In practice, high-degree nodes (*e.g.*, highly social individuals) might be subject to more testing (and quarantining if positive) than low-degree nodes because of their higher expected rate of infecting others. This rationale translates to $f_k(t)/N_k > f_{k'}(t)/N_{k'}$ if $k > k'$. In our numerical experiments, we use sufficiently broad bounds of $f_k(t)$ and set $f_k^{\min} = f_{\min}$ and $f_k^{\max} = f_{\max}$.

To minimize the number of total infections over time, while simultaneously stressing the importance of reducing early infections, we define a loss function as

$$L(T) = \int_0^T \mathrm{d}t\, \delta^t \sum_{k=1}^{K} k s_k(t) \sum_{\ell=1}^{K} \frac{P(\ell|k)}{P(\ell)} \left(\beta^{\mathrm{u}} i_\ell^{\mathrm{u}}(t) + \beta^* i_\ell^*(t)\right), \quad (11)$$

where $\delta \in (0, 1]$ denotes a discount factor, which describes how we balance between minimizing current infections and future infections. The smaller the parameter δ, the less attention we pay to future infections, and the more we focus on reducing early infections. For example, medical resources can better handle confirmed patients and new treatments can be given time to develop if the number of infections are spread over longer time periods. These effects can be effectively incorporated in the loss function by using $\delta < 1$. Minimizing the loss (11) is equivalent to minimizing the number of infections, weighted by the discount factor δ^t, in the time horizon $[0, T]$. To search for the optimal testing function $f_k(t)$ that minimizes (11), we invoke Pontryagin's maximum principle (PMP) and construct the associated Hamiltonian

$$\begin{aligned} H &= \delta^t \sum_{k=1}^{K} k s_k(t) \sum_{\ell=1}^{K} \frac{P(\ell|k)}{P(\ell)} \left(\beta^{\mathrm{u}} i_\ell^{\mathrm{u}}(t) + \beta^* i_\ell^*(t)\right) \\ &\quad + \sum_{k=1}^{K} \left(\lambda_k^s \frac{\mathrm{d}s_k(t)}{\mathrm{d}t} + \lambda_k^{\mathrm{u}} \frac{\mathrm{d}i_k^{\mathrm{u}}(t)}{\mathrm{d}t} + \lambda_k^* \frac{\mathrm{d}i_k^*(t)}{\mathrm{d}t} \right) \\ &= \sum_{k=1}^{K} (\delta^t - \lambda_k^s + \lambda_k^{\mathrm{u}}) k s_k(t) \sum_{\ell=1}^{K} \frac{P(\ell|k)}{P(\ell)} \left(\beta^{\mathrm{u}} i_\ell^{\mathrm{u}}(t) + \beta^* i_\ell^*(t)\right) \\ &\quad + \sum_{k=1}^{K} \left[\frac{f_k(t)}{N_k} (\lambda_k^* - \lambda_k^{\mathrm{u}}) i_k^{\mathrm{u}}(t) - \gamma^{\mathrm{u}} i_k^{\mathrm{u}}(t) \lambda_k^{\mathrm{u}} - \gamma^* i_k^*(t) \lambda_k^* \right], \end{aligned} \quad (12)$$

where λ_k^s, λ_k^{u}, and λ_k^* are adjoint variables associated with s_k, i_k^{u}, and i_k^*, respectively. PMP states that a necessary condition for the loss-minimizing control $f_k(t)$ is that it minimizes H (or maximizes $-H$) at every time point t. This method of optimal control has been applied to many other contexts, including control of economic growth [40]. In our problem, applying PMP under the total budget constraint $\sum_{k=1}^{K} f_k(t) = F(t)$, we explicitly find the minimizing testing function $(f_k^*) = \operatorname{argmin}_f H$, which we will assume to be optimal control that minimizes $L(T)$. The dynamics for $(\lambda_k^s, \lambda_k^{\mathrm{u}}, \lambda_k^*)$ obey

$$\begin{aligned} \frac{\mathrm{d}\lambda_k^s}{\mathrm{d}t} &= -\frac{\partial H}{\partial s_k} = -\delta^t k \sum_{\ell=1}^{K} \frac{P(\ell|k)}{P(\ell)} \left(\beta^{\mathrm{u}} i_\ell^{\mathrm{u}}(t) + \beta^* i_\ell^*(t)\right) \\ &\quad + \lambda_k^s k \sum_{\ell=1}^{K} \frac{P(\ell|k)}{P(\ell)} \left(\beta^{\mathrm{u}} i_\ell^{\mathrm{u}}(t) + \beta^* i_\ell^*(t)\right) \\ &\quad - \lambda_k^{\mathrm{u}} k \sum_{\ell=1}^{K} \frac{P(\ell|k)}{P(\ell)} \left(\beta^{\mathrm{u}} i_\ell^{\mathrm{u}}(t) + \beta^* i_\ell^*(t)\right), \end{aligned} \quad (13)$$

$$\begin{aligned} \frac{\mathrm{d}\lambda_k^{\mathrm{u}}}{\mathrm{d}t} &= -\frac{\partial H}{\partial i_k^{\mathrm{u}}} = -\frac{\beta^{\mathrm{u}}}{P(k)} \sum_{j=1}^{K} P(k|j) s_j(t) (\delta^t - \lambda_j^s + \lambda_j^{\mathrm{u}}) \\ &\quad + \gamma^{\mathrm{u}} \lambda_k^{\mathrm{u}} + \frac{f_k(t)}{N_k} (\lambda_k^{\mathrm{u}} - \lambda_k^*), \end{aligned} \quad (14)$$

$$\begin{aligned} \frac{\mathrm{d}\lambda_k^*}{\mathrm{d}t} &= -\frac{\partial H}{\partial i_k^*} \\ &= -\frac{\beta^*}{P(k)} \sum_{j=1}^{K} P(k|j) s_j(t) (\delta^t - \lambda_j^s + \lambda_j^{\mathrm{u}}) + \gamma^* \lambda_k^*, \end{aligned} \quad (15)$$

with end conditions $\lambda_k^s(T) = \lambda_k^{\mathrm{u}}(T) = \lambda_k^*(T) = 0$. To minimize H with respect to the testing rates $f_k(t)$, we have to minimize the term

$$\sum_{k=1}^{K} \frac{f_k(t)}{N_k} (\lambda_k^* - \lambda_k^{\mathrm{u}}) i_k^{\mathrm{u}}(t) \quad (16)$$

given the budget constraints (9) and (10). Hence, after giving each subpopulation the minimal testing resources $f_{\min} N_k$, we maximize the testing rates $f_k(t)$ with the smallest coefficients $(\lambda_k^* - \lambda_k^{\mathrm{u}}) i_k^{\mathrm{u}}(t)/N_k$ of $f_{\max} N_k$ as long as sufficient testing budget is available. In other words, we should give testing resources to those groups presumed to be at the highest risk, as quantified by the quantity $(\lambda_k^* - \lambda_k^{\mathrm{u}}) i_k^{\mathrm{u}}(t)/N_k$. We use the PMP-based algorithm outlined in Appendix B to iteratively calculate the loss function (11) and optimal testing strategy.

In accordance with empirical data on COVID-19 patients [39], [41], [42], we set $\gamma = \gamma^{\mathrm{u}} = \gamma^* = (1/14)/\mathrm{day}$ and $\beta^* = \beta^{\mathrm{u}}/10$. The transmissibility of untested individuals, β^{u}, is calculated according to (7) as $\beta^{\mathrm{u}} = 0.0411/\mathrm{day}$ for the BA network and $\beta^{\mathrm{u}} = 0.0130/\mathrm{day}$ for the SBM network. We set the discount factor $\delta = 0.95$ so that initial infections contribute more to the loss function (11). The total daily number of SARS-CoV-2 tests in the US after an initial ramping-up phase in 2020 is about 0.6%/day [28]. Hence, we set

$$\sum_k f_k(t) = 0.006N, \quad (17)$$

and $f_{\min} = 0$, $f_{\max} = 0.4N_k$. As initial condition, we use

$$\begin{aligned} s_k(0) &= P(k) - i_k^{\mathrm{u}}(0), \quad i_k^*(0) = 0, \\ i_k^{\mathrm{u}}(0) &= 10^{-6} P(k), \qquad r_k(0) = 0, \end{aligned} \quad (18)$$

corresponding to about 0.1 of an infected individual uniformly distributed on $N \approx 10^5$ susceptible nodes. The optimal testing strategy is supposed to identify those nodes that are most likely to be infected and transmit the disease to others. Upon

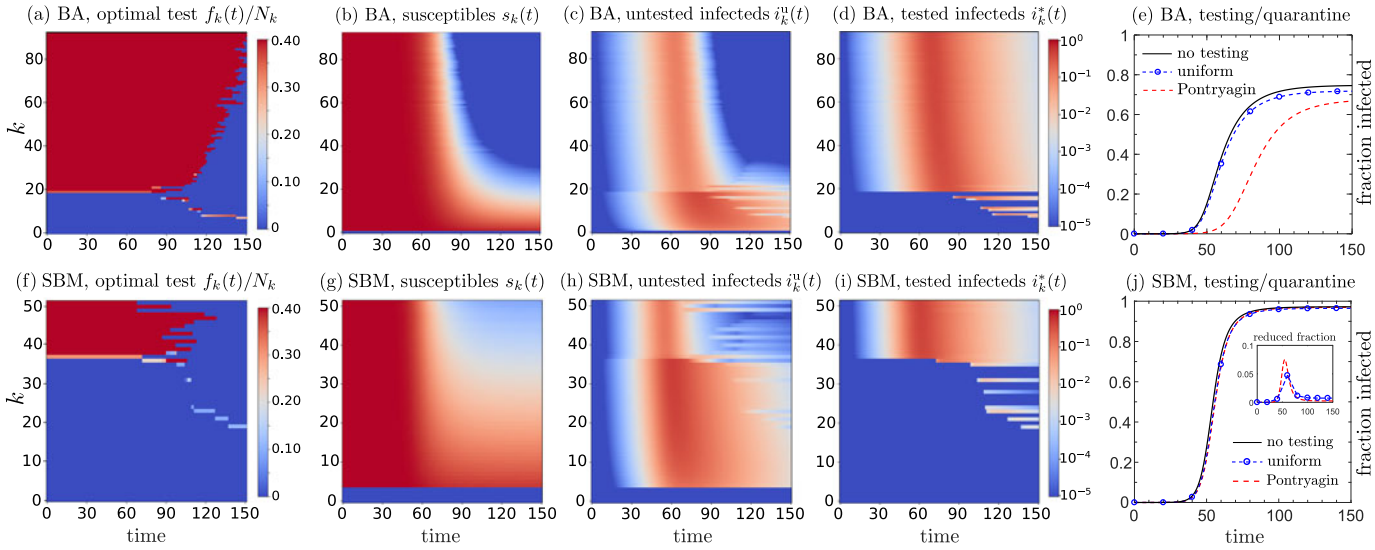

Fig. 2. Optimal testing and quarantining strategy for $T = 200$ and discount factor $\delta = 0.95$. We plot the optimal strategies and the corresponding susceptible, untested infected, and tested infected fractions at each degree k across time $t = n\Delta t$. (a) A heatmap of the PMP-optimal testing strategy (see Algorithm 1) for the BA network. The corresponding populations of degree-k susceptibles, untested infecteds, and tested infecteds are plotted in (b-d), respectively. (e) Time-evolution of the total fraction infected $1 - \sum_{k=1}^{K} s_k(t)$ under the PMP-optimal testing strategy (dashed red). The fractions infected under hypothetical uniform testing (dashed blue/circle) and no testing (black) scenarios are shown for comparison. For the BA network, optimal testing both delays and suppresses epidemic spreading more effectively than uniform testing. The bottom row (f-j) shows analogous results for the SBM network. Panels (f-i) show the corresponding optimal testing rates, susceptible, untested infected, and untested infected populations with degree k as a function of time. Panel (j) shows the fraction infected as a function of time. Although optimal testing and quarantining reduce the fraction infected relative to uniform or no testing, its effects are only modestly better. Given the same testing budget constraint, the effects of optimal testing strategies are greater in the BA network because its distribution of node degrees is more heterogeneous and testing and quarantining high-degree nodes can more effectively control disease spread. However, since the node degree distribution in the SBM network is sharply peaked, an optimal testing strategy is less effective overall.

using $T = 200, \Delta t = 0.1$ and $\delta = 0.95$, we find the optimal testing strategy $f_k(t)/N_k$ for our BA network and plot it in Fig. 2(a). Here (2)–(5) and (13)–(15) are solved using an improved Euler method. For the BA network, the value of the loss function defined in (11) is $L(T = 200) = 0.0109$ under the optimal testing strategy, while it is $L(T = 200) = 0.0325$ under uniform testing

$$f_k = F_0 \frac{N_k}{N}. \tag{19}$$

Fig. 2(b), (c), and (d) show the associated populations under optimal testing, while (e) shows the dynamics of the fraction of nodes infected, $1 - \sum_{k=1}^{K} s_k(t)$. The disease spread under optimal testing is significantly slowed relative to the no testing (black) and uniform testing (dashed blue/circle) cases. Fig. 2 (f) plots the optimal testing rate for the SBM network. Panels (g-i) show the corresponding subpopulations, and panel (j) plots the fraction of nodes infected under PMP-optimal, uniform, and no-testing conditions. For the SBM network, $L(T = 200) = 0.0564$ under the optimal strategy and $L(T = 200) = 0.0571$ under the uniform testing strategy, suggesting that the PMP approach yields better solutions than uniform testing. However, the improvement is modest and the SBM network is rather insensitive to testing and quarantining. The slight improvement from testing is shown by the *reduction* in the fraction infected relative to the no testing case (inset).

In both networks, nodes with larger degrees are more likely to be tested at the beginning of the outbreak [Fig. 2(a) and (f)], indicating that people with more contacts are more likely to infect others or get infected, and should be given priority to get tested. Yet, in both networks, as time evolves, the optimal testing strategy tends to shift focus from higher degree nodes to nodes with smaller degrees because testing those nodes that were infected and have already recovered is not meaningful in terms of disease control.

Comparing Fig. 2(e) and (j), we see that the differences between optimal and uniform testing are larger for the BA network compared to the SBM. A possible explanation for this behavior is that in the BA network, the degree distribution $P(k)$ decays algebraically. Therefore, as long as testing focuses primarily on high-degree nodes, the spreading of the disease can be controlled very effectively since the majority of nodes have small degrees and are more unlikely to be infected. On the other hand, for our SBM network, the degrees of most nodes are close to each other and larger than 10, indicating that nodes with a small degree are more likely to be infected compared to the BA network. Even if we use the same uniform testing rates (see (19)) in both networks, the proportion of infections in the BA network is less than that in the SBM network.

IV. Optimal Vaccination Policy

Optimal vaccination has also been studied within the classic SIR model [43]. However, devising vaccination strategies based on social network structure may provide a more refined and efficient way of administering vaccines and extinguishing an epidemic. Our simple testing model presented in the previous section can be straightforwardly

adapted to describe vaccination on a network. The goal is to determine the optimal allocation of vaccine doses to a population with heterogeneous contacts to minimize the impact of the infection across the entire population.

For COVID-19, there are a variety of vaccines that require one or two shots [44]. In our simulations, we assume that the administered vaccine provides full protection after one shot and that a vaccinated individual will instantly leave the susceptible group and enter the recovered group. This means that vaccinated individuals will no longer be infectious and can be treated as "recovered" after receiving one vaccination dose. Furthermore, we assume that only susceptible persons will be vaccinated. Other mechanisms such as prime-boost protocols and time delays between vaccination and onset of immune response can also be accounted for in similar models as detailed in [45].

We reformulate (2)-(5) to study optimal vaccination protocols that are constrained by vaccine supplies in a heterogeneous population. For simplicity, we do not take into account the effect of testing and quarantining when devising optimal vaccinating strategies, although testing and vaccination can be performed concurrently. The resulting rate equations are

$$\frac{\mathrm{d}s_k(t)}{\mathrm{d}t} = -\beta k s_k(t) \sum_{\ell=1}^{K} \frac{P(\ell|k)}{P(\ell)} i_\ell(t) - \frac{v_k(t)}{N}, \tag{20}$$

$$\frac{\mathrm{d}i_k(t)}{\mathrm{d}t} = \beta k s_k(t) \sum_{\ell=1}^{K} \frac{P(\ell|k)}{P(\ell)} i_\ell(t) - \gamma i_k(t), \tag{21}$$

$$\frac{\mathrm{d}r_k(t)}{\mathrm{d}t} = \gamma i_k(t) + \frac{v_k(t)}{N}, \tag{22}$$

where $v_k(t)$ is the rate of vaccination of susceptibles with degree k at time t. Once vaccinated, susceptibles become "recovered" because they are immunized and no longer susceptible to the infection. The total rate of administering vaccines at time t is defined as

$$\sum_{k=1}^{K} v_k(t) = V(t). \tag{23}$$

In other words, in time increment Δt at time t, we can administer only $V(t)\Delta t$ doses. Equation (20) assumes that vaccination is resource-limited and that the rate of protecting susceptibles is proportional only to the rate $v_k(t)$ of administering vaccines. In addition, we assume that the vaccination rates for different subpopulations are confined to the interval

$$v_{\min} \le \frac{v_k(t)}{N s_k(t)} \le v_{\max}, \tag{24}$$

where $v_{\min}, v_{\max} \in [0, 0.4]/\text{day}$ are minimum and maximum vaccination rates. Note that vaccines are allocated only to susceptibles, while tests are typically given to individuals of all categories: susceptible, infected, and recovered, according to their relative proportions. To formulate the vaccine distribution problem in a heterogeneous contact network, we use the following loss function

$$L(T) = \int_0^T \mathrm{d}t\, \delta^t \sum_{k=1}^{K} k s_k(t) \sum_{\ell=1}^{K} \frac{P(\ell|k)}{P(\ell)} \beta(t) i_\ell(t), \tag{25}$$

with the aim of minimizing the total number of infections over time (with a constant discount factor $\delta \in (0, 1]$) by appropriately distributing vaccines to groups with different degree k at different rates.

To minimize the loss function (25), we construct the Hamiltonian

$$\begin{aligned} H &= \beta \delta^t \sum_{k=1}^{K} k s_k(t) \sum_{\ell=1}^{K} \frac{P(\ell|k)}{P(\ell)} i_\ell(t) \\ &\quad + \sum_{k=1}^{K} \left(\lambda_k^s \frac{\mathrm{d}s_k(t)}{\mathrm{d}t} + \lambda_k^i \frac{\mathrm{d}i_k(t)}{\mathrm{d}t} \right) \\ &= \beta \sum_{k=1}^{K} (\delta^t - \lambda_k^s + \lambda_k^i) k s_k(t) \sum_{\ell=1}^{K} \frac{P(\ell|k)}{P(\ell)} i_\ell(t) \\ &\quad + \sum_{k=1}^{K} \left(\frac{v_k(t)}{N} \lambda_k^s(t) - \gamma i_k(t) \lambda_k^i(t) \right), \end{aligned} \tag{26}$$

Algorithm 1: Pseudo-Code for Determining Optimal Testing Strategies Based on Pontryagin's Maximum Principle.

Initialize $t = 0$, $s_k(0), i_k^{\mathrm{u}}(0), i_k^*(0)$, Δt, $T = n\Delta t$, $\beta^{\mathrm{u}}, \beta^*, \gamma^{\mathrm{u}}, \gamma^*, \delta$, initial strategy $F(k\Delta t)$, k, $f_{\max}$, $f_{\min}$, ϵ, $iter_{\max}$
for $k = 0 : n-1$ **do**
  Calculate $s_k(t), i_k^*(t), i_k^{\mathrm{u}}(t)$ under the strategy $F(k\Delta t)$ from (2)–(4)
end for
Set $\lambda_k^s, \lambda_k^{\mathrm{u}}, \lambda_k^* = 0, k = n$
Calculate the loss function L_1 in (11)
for $k = n-1 : 0$ **do**
  Calculate $\lambda_k^s, \lambda_k^{\mathrm{u}}, \lambda_k^*$ under the strategy $F(k\Delta t)$ from (13)–(15)
end for
for $k = 0 : n-1$ **do**
  First renew the strategy $F(k\Delta t)$, then calculate $s_k, i_k^{\mathrm{u}}, i_k^*$ under the strategy $F(k\Delta t)$ from (2)–(4)
end for
Calculate the loss function L_2 in (11)
$i \leftarrow 1$
while $|L_1 - L_2| > \epsilon$ && $i < iter_{\max}$ **do**
  $i \leftarrow i + 1$
  $L_1 \leftarrow L_2$
  Set $k = n, \lambda_k^s, \lambda_k^{\mathrm{u}}, \lambda_k^* = 0$
  for $k = n-1 : 0$ **do**
    Calculate $\lambda_k^s, \lambda_k^{\mathrm{u}}, \lambda_k^*$ under the strategy $F(k\Delta t)$ from (13)–(15)
  end for
  for $k = 0 : n-1$ **do**
    First renew the strategy $F(k\Delta t)$, then calculate $s_k, i_k^{\mathrm{u}}, i_k^*$ under the strategy $F(k\Delta t)$ from (2)–(4)
  end for
  Calculate the Loss function L_2 in (11)
end while

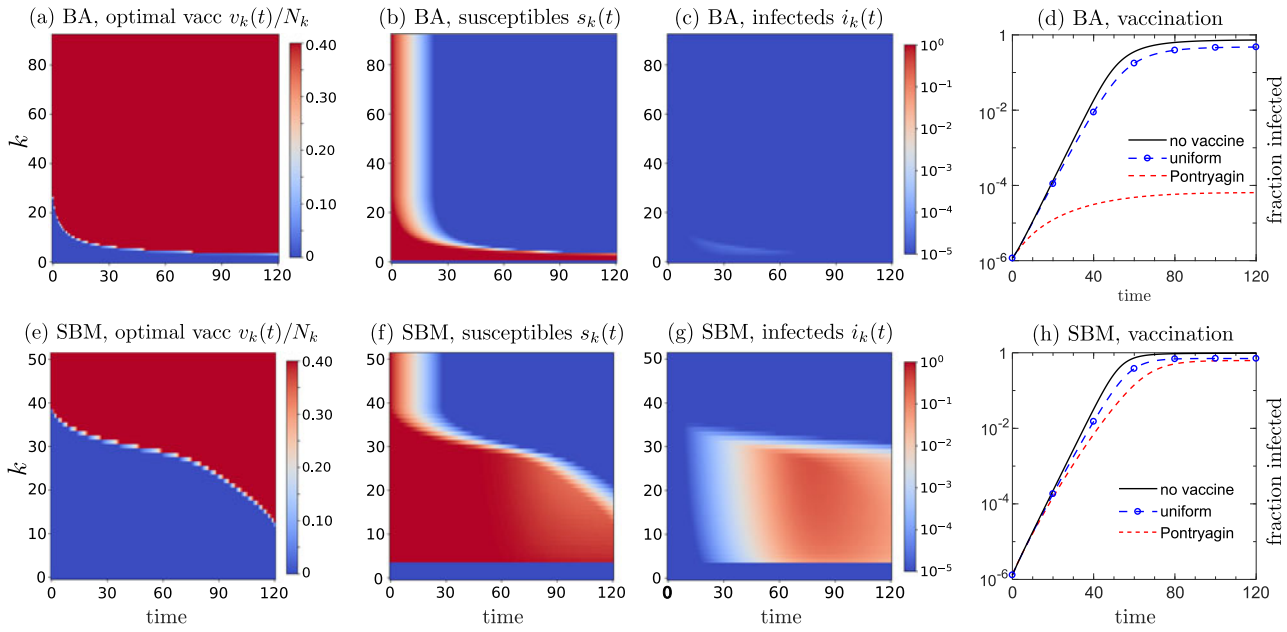

Fig. 3. Vaccination model optimized for $T = 150$ under different constraints. We plot the optimal strategies and the corresponding susceptible, untested infected, and tested infected fractions at each degree k across time $t = n\Delta t$. (a) Heatmap of the optimal vaccination strategy $v_k(t)/(s_k(t)N_k)$ for the BA network given by Algorithm 1. Panels (b,c) show the corresponding susceptible and infected subpopulations $s_k(t)$ and $i_k(t)$, while (d) plots the fraction infected as a function of time, derived from solving (20)–(22) under optimal vaccination using a discount factor $\delta = 0.95$. The dashed red curve indicates the fraction infected under optimal vaccination. For comparison, the infected population under no vaccination (solid black) and constant, uniform (dashed blue/circles) vaccination are also plotted and show how optimizing vaccination significantly suppresses infectivity. Panels (e-h) show the corresponding quantities for the SBM network. Optimal vaccination is less effective at decreasing infection in the SBM network than in the BA network, again because of the SBM's peaked (more homogeneous) node degree distribution. Note from the logarithmic scale that vaccination is qualitatively more effective in reducing infections than testing and quarantining.

where λ_k^s and λ_k^i are the adjoint variables satisfying the differential equations

$$\frac{\mathrm{d}\lambda_k^s}{\mathrm{d}t} = -\frac{\partial H}{\partial s_k} = -\beta k \sum_{\ell=1}^{K} \frac{P(\ell|k)}{P(\ell)} i_\ell(t)(\delta^t - \lambda_k^s + \lambda_k^i) \tag{27}$$

$$\frac{\mathrm{d}\lambda_k^i}{\mathrm{d}t} = -\frac{\partial H}{\partial i_k} = -\frac{\beta}{P(k)} \sum_{\ell=1}^{K} P(k|\ell)\ell s_\ell(t)(\delta^t - \lambda_\ell^s + \lambda_\ell^i) + \gamma\lambda_k^i. \tag{28}$$

Again, PMP explicitly generates the control that minimizes (26) for every t, which we assume also minimizes the target loss function (25). From the constraints (23) and (24), minimizing the Hamiltonian is achieved by giving vaccination doses to those subpopulations with the smallest λ_k^s. We use the same Algorithm 1 to solve the minimization problem (25) numerically and obtain the optimal strategy.

In the US, about two million doses of SARS-CoV-2 vaccines were delivered in May 2021 [46], most of which were two-dose vaccines. Since approximately 0.3% of the entire US population is fully vaccinated daily, we set $V(t) = 0.003N/\mathrm{day}$, $v_{\min} = 0/\mathrm{day}$, and $v_{\max} = 0.4/\mathrm{day}$ in the constraint (24). The infection rates β are set to be $0.0411/\mathrm{day}$ for the BA network and $0.0130/\mathrm{day}$ for the SBM network, and the recovery rate $\gamma = (1/14)/\mathrm{day}$. For comparison, we also simulate a vaccination strategy with a uniform vaccination rate

$$v_k(t) = \frac{s_k(t)V(t)}{\sum_{k=1}^{K} s_k(t)}. \tag{29}$$

In all simulations, we use the following initial condition:

$$i_k(0) = 10^{-6}P(k),\ r_k(0) = 0, s_k(0) = P(k) - i_k(0). \tag{30}$$

We plot the PMP-optimal vaccination strategy v_k/N_k in Fig. 3 (a) and the corresponding susceptible and infected k-degree subpopulations $s_k(t)$ and $i_k(t)$ in (b) and (c). We set $T = 150, \Delta t = 0.1$ and we use an improved Euler method to numerically solve (20)–(22), (27)–(28). Algorithm 1 is applied (without the infected and tested compartment) to determine the optimal vaccination strategy by the PMP approach. For the BA network, $L(T = 150) = 1.165 \times 10^{-5}$ under the PMP-optimal strategy and $L(T = 150) = 0.01953$ under a uniform vaccination rate. Fig. 3(d) shows that the optimal vaccination strategy on a BA network significantly reduces the fraction infected compared to the uniform vaccination strategy. Panels (e-h) show the corresponding quantities for the SBM network for which $L(T = 150) = 0.0210$ under the optimal vaccination strategy and $L(T = 150) = 0.0360$ under a constant, uniform vaccination strategy. In both networks, the optimal vaccination strategies obtained via Algorithm 1 tend to prioritize those nodes with higher degrees first and eventually expand to those nodes with smaller degrees [see Fig. 3(a) and (e)]. As with testing and quarantining, the reduction in the fraction infected by vaccination is greater in the BA network. Since the BA network has a degree distribution with algebraic decay, the effect of the

optimal vaccination strategy will be more pronounced than for the SBM, whose nodes have similar degrees.

V. Discussion

Effective testing and vaccination strategies are an essential part of epidemic management. In this paper, we derived optimal testing and vaccination policies by applying Pontryagin's maximum principle to a degree-based epidemic model in a heterogeneous contact network. We complemented our analytical results with reinforcement learning (RL) approaches that identify effective policies (see Appendix C). On occasions when the best optimal strategy can be analytically solved, the controls derived from Pontryagin's maximum principle outperform RL-based interventions. However, reinforcement learning is useful for epidemic management problems when an efficient procedure for computing optimal solutions solutions is not available. Our results show that the two approaches can complement each other; preliminary findings from optimal-control analysis can be used to pre-train and restrict the space of possible actions, which may lead to more efficient RL algorithms. In addition to RL-based control strategies, it may also be worthwhile to apply neural ODE control frameworks [47], [48] to resource allocation problems since they have exhibited better performance than RL and numerical adjoint system solvers.

Our analytical results show that optimal testing and vaccination policies under resource constraints initially tend to prioritize nodes with higher degrees to control the spread of the disease. In situations where the number of contacts of individuals is known or can be estimated with reasonable precision, Algs. 1 and 2 may be useful for identifying effective epidemic management strategies. Using our control-theoretic approach, we also explored the relative effectiveness of testing and vaccination under different conditions. If more information on contact patterns of individual nodes is available, it is possible to further refine the proposed policies using interventions that rely not only on node degrees but also on other structural features such as percolation and betweenness centrality [26], [49].

A. Effects of Delayed Intervention

First, we consider the effectiveness of interventions as a function of the time between the first infection and the implementation of testing or vaccination. The initial conditions are set to be the same as (18) and (30). Fig. 4 shows the total fraction infected and the loss functions at $T = 150$, for both the BA and SBM networks, as a function of intervention starting time t_0. We set $F = F(t)\mathbb{1}_{t>t_0}$ or $V = V(t)\mathbb{1}_{t>t_0}$ and explore the effects of different constant levels of test kits or vaccine availability, $F(t) = 0.002N, 0.004N, 0.006N, 0.008N/\mathrm{day}$ and $V(t) = 0.001N, 0.002N, 0.003N, 0.004N/\mathrm{day}$, respectively. The transmissibility rates $\beta^{\mathrm{u}}, \beta^*, \beta$ and the recovery rates $\gamma^{\mathrm{u}}, \gamma^*, \gamma$ are set to the same values as those used in Section III for the testing model and those used in Section IV for the vaccination model.

In the BA network, the total infected fraction shown in Fig. 4 (a) is fairly insensitive to starting times for all testing rates F, especially at small starting times $t_0 \lesssim 50$. However, the loss functions corresponding to all testing rates increase monotonically with the testing starting time t_0, as shown in Fig. 4(b). On the other hand, vaccination of a BA network leads to infected fractions that change significantly with delay time, but with an overall vaccination-rate-dependent starting time before which disease spread can be nearly completely suppressed, as shown in (c). For the vaccination model applied to both networks, an earlier intervention time will always lead to fewer infected nodes. Overall, we found that earlier and stronger intervention measures lead to more effective control of disease spread and a smaller loss function defined by (11), (25). For all cases, the testing loss functions monotonically increase with t_0. Similarly, for the SBM model, the final infections shown in Fig. 4(d) are insensitive to starting times $t_0 \lesssim 50$ for each of the four choices of total testing budgets F. Earlier intervention times t_0 lead to smaller testing loss functions which indicate more effective early-time disease control and fewer early infections (which are followed by larger later infections) than those associated with later start times t_0. Vaccination of the SBM network reveals more smoothly monotonically increasing infected fraction and loss functions and does not display the sigmoidal dependence on intervention time t_0 as exhibited by the infected fraction and loss function for the BA network. For the vaccination model applied to both networks, an earlier intervention time will always lead to fewer infected nodes.

Overall, higher levels of F and V lead to high-k nodes being addressed sooner and total infections can be reduced. In summary, for both networks, when the discount factor $\delta < 1$, earlier intervention starting times t_0 more effectively reduce early infections although it might be at the cost of larger later infections.

B. Dependence on Initial Conditions

Besides the start time of testing or vaccination, initial conditions may also affect the optimal strategy. For example, the initial propagation of the disease may depend on the degree k_{i} of the first infected individual [50]. Instead of an initial infectious source that is uniformly distributed across all nodes, as described in (18) and (30), we vary the degree of the first infected node and explore how the strategies change as a function of concentrated initial condition $i_k(0) = N_0\mathbb{1}_{k,k_{\mathrm{i}}}/N$. We take $N_0 = 10^{-6}N$ for both networks, $k_{\mathrm{i}} = 3, 20, 90$ for the BA network, and $k_{\mathrm{i}} = 5, 20, 30$ for the SBM network. These different initial conditions are denoted IC1, IC2, and IC3 for each network, respectively.

Optimal testing strategies are found to be subtly dependent on the initial conditions, *i.e.*, the degree of the initial infected patient. In Fig. 5 we show only the optimal strategy associated with IC2, but plot the time-dependent fraction infected under optimal strategies for all ICs. Under both optimal testing and vaccination, a smaller degree of the first infected source typically leads to a smaller subsequent infected population. Specifically for testing, this decrease is greatest at intermediate times because at early stages there are fewer infecteds. At later times, the testing strategy becomes insensitive to the initial

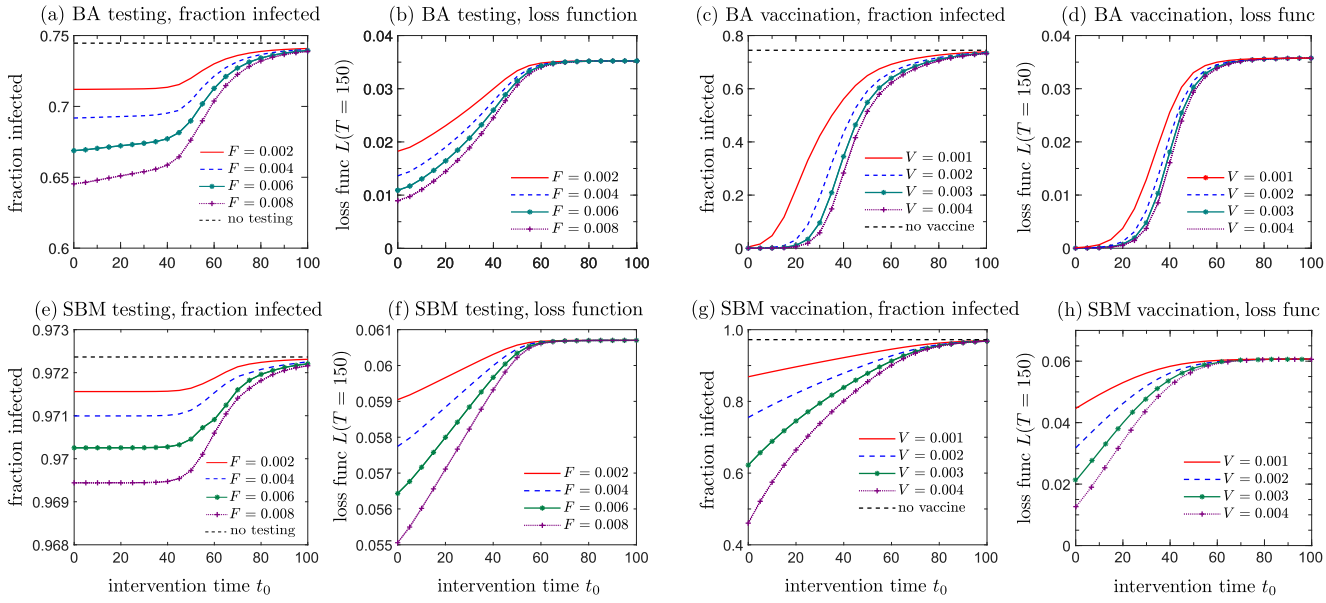

Fig. 4. Total fraction infected under testing or vaccination model as a function of different intervention starting times t_0. We minimize the corresponding loss function at $T = 150$ and use $\delta = 0.95$. (a) The fraction infected in the BA network as a function of start times for different testing amplitudes F. The total infected fraction is fairly insensitive to intervention starting times, especially for small intervention delays. The effect of delayed vaccination on the fraction infected is shown in (c), with the corresponding loss function shown in (d). For the SBM network, the fraction infected as a function of testing start time shown in (e) reflects the small effect of testing on the infected population. However, the loss functions shown in (f) are monotonic in the starting time. This implies that an early intervention time on the SBM network is able to "flatten" the curve by postponing infection so even if total infections stay roughly the same when t_0 varies in $\sim [0, 50]$, the earlier the intervention time, the fewer the earlier infections, with little change in the final total infected fraction. The starting time dependence of the fraction infected on an optimally vaccinated SBM network in (g) shows a monotonic and smooth decrease in effectiveness as vaccination is delayed. In (h), the loss function for vaccination on the SBM network also monotonically increases with the start time.

condition because persons with all degrees are infected and those with a higher degree tend to be infected sooner.

The optimal vaccination strategies obtained through Algorithm 1 are also relatively insensitive to initial conditions in both networks, particularly at longer times. Although not shown, the optimal strategies associated with different ICs are mostly the same because nodes with larger degrees tend to always be vaccinated first to minimize the loss function (25). Susceptibles with higher degrees are more vulnerable and should be vaccinated first to mitigate subsequent infection events. For vaccination, the different ICs lead to long-term differences in infected fractions because low-degree ICs allows more time for vaccination to more effectively remove susceptibles.

C. Monte-Carlo Simulation of Stochastic Network Model

Our results are derived from a mass-action ODE model and it is unclear how they apply to stochastic network dynamics. To compare these different representations of the disease, we define the discrete stochastic versions of our network models and impose a stochastic version of the optimal strategies found using the PMP on our ODEs. In Appendix D we implemented the optimal testing and vaccination strategies on the BA and SBM network realizations used in the PMP study. Infection, recovery, testing, and vaccination processes are described as Markov events in continuous time. Results from rejection-free, event-based Monte-Carlo simulation indicate that the degree-based mean-field ODE model tends to overestimate new infections because it assumes that all subpopulations interact in a well-mixed manner, thus neglecting certain structural features of the considered networks. The loss function derived from the stochastic model and using the PMP-derived optimal strategy is shown to be lower than that of the ODE model, except for vaccination on the SBM network for which they are comparable. Nonetheless, optimal strategies derived from the ODE model still outperforms a uniform or unstructured testing or vaccination strategy applied to the discrete stochastic model.

VI. Summary and Conclusions

Our overall results indicate that different network structures (*e.g.*, BA *vs.* SBM) have different susceptibilities to optimal intervention strategies. Thus, policies such as selective social distancing can potentially be used to shift network structure towards one that is more sensitive to direct testing and vaccination strategies.

We have analyzed testing and vaccination separately, but in practice, both are simultaneously implemented. The relative efforts of these two interventions, as a function of time, will depend on their constraints and costs as well as the desired loss function time T. A further generalization of either our mass-action or stochastic versions of our network models may be to derive different loss functions other than (11) and (25) to take into account factors such as economic effects or prioritization of certain groups (*e.g.*, healthcare workers or individuals with comorbidities). Formulating more specific loss functions would allow one to balance mitigation and suppression strategies, as studied in a well-mixed SIR model [12]. Another important and straightforward extension of our model is to consider the effects of waning protection of vaccination, which has become a relevant feature of disease control in the context of booster shots. Recovered individuals that include

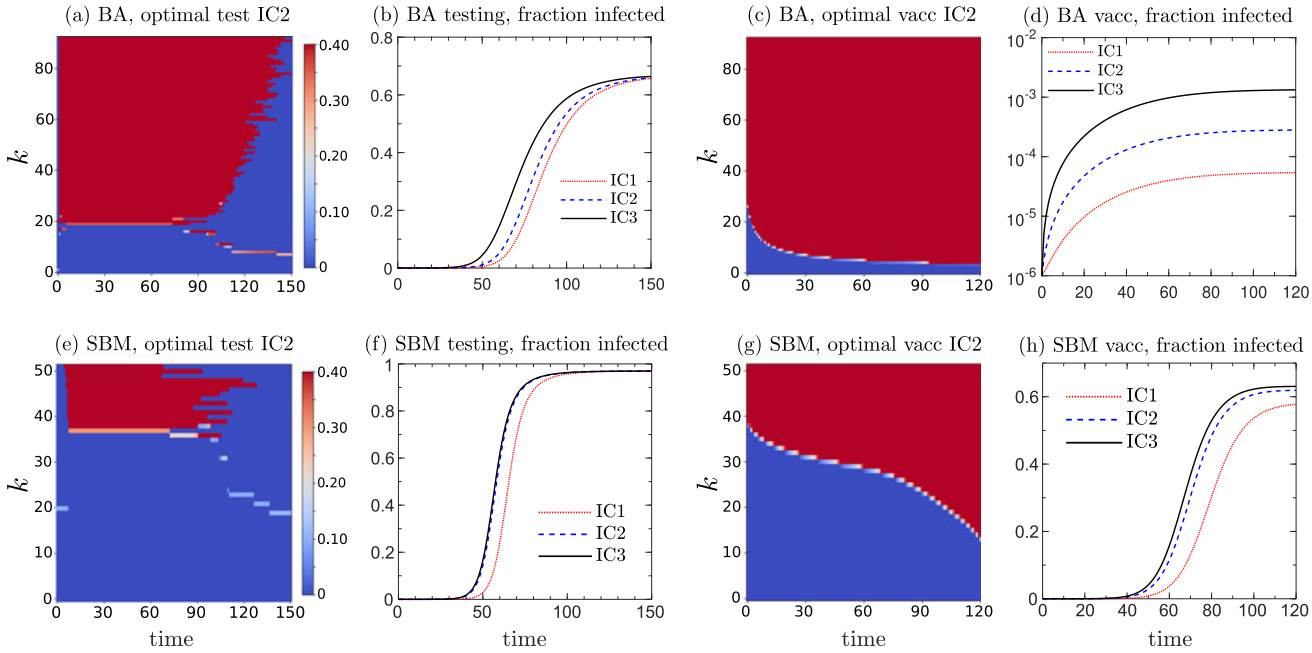

Fig. 5. Dependence of intervention effectiveness on the degree of the initial infected individual. (a) The PMP-optimal testing strategy computed using IC2 ($k_{\mathrm{i}} = 20$) on the BA network. Strategies for IC1 ($k_{\mathrm{i}} = 3$) and IC3 ($k_{\mathrm{i}} = 90$) are qualitatively similar (not shown) with small differences at the beginning leading to the different delays in the infection dynamics shown in (b). Specifically, for IC1 and IC3, the initial transient of the optimal testing strategy maximizes the testing rate for the subpopulation with the same degree as k_1 and k_3, respectively, indicating that the optimal testing strategy is sensitive to the degree properties of the initial seed infection. Once the disease spreads out, the testing strategies "forget" the initial condition and converge to each other. Despite optimal testing, initial infecteds with larger degrees, such as IC3, lead to the earlier spread of the epidemic. Results are found by using a discount factor $\delta = 0.95$, the optimal strategy given in Algorithm 1, and solving (2)–(5). (c-d) The optimal vaccination strategy for IC2 and the associated fraction infected for the BA network. As with testing, the vaccination strategies associated with IC1 ($k_{\mathrm{i}} = 5$) and IC3 ($k_{\mathrm{i}} = 30$) lead to differences in infection magnitudes. However, the optimal vaccination strategies are insensitive to different initial conditions, even at early times. Since the mechanism of vaccination is always to protect high-degree susceptibles, the vaccination strategies are not as dependent on the current infected population as the testing strategies are. Panel (e) shows the optimal testing strategy for the SBM network, assuming IC2 ($k_{\mathrm{i}} = 20$). (f) The fraction infected exhibits slower dynamics for smaller-degree initial conditions. (g) Optimal vaccination strategy for IC2 in the SBM network, and (h), the associated infected fraction showing both delay and amplitude changes with changes in the initial condition.

previously vaccinated or infected individuals can become susceptible again at a rate equal to the rate of loss of immunity. Thus, another timescale (months) is introduced which is comparable to timescales T that are used to define the loss function. We expect even wider variety and richness in the analysis of optimization problems under waning immunity.

Finally, the discrepancies between the effective degree ODE model and the Monte-Carlo simulations, under the same ODE-derived optimal strategies appear to arise from the differences in the underlying disease propagation. The discrete stochastic models tend to show lower infected fractions than the corresponding mass-action ODE models since its discreteness and finite infection lifetimes prevents high-degree nodes in some network regions to be infected while the mass-action model allows all nodes to be partially infected. Further analysis of fluctuations in real-world stochastic models could provide insight into a better estimation of optimal strategies without simulating the large space of intervention strategies. This and many other important extensions will be topics of future exploration.

Appendix A
Basic Reproduction Number

In this appendix, we analytically derive the basic reproduction number $\mathcal{R}_0$ for uncorrelated networks and compare the resulting values with those obtained using (7) and (8). As a starting point, we note that the conditional degree distribution $P(\ell|k)$ can be expressed in terms of a symmetric (for undirected networks) joint degree distribution $P(\ell, k)$, the probability that a randomly chosen edge connects two nodes with degrees ℓ and k. Marginalizing $P(\ell, k)$ over ℓ yields the distribution over edge ends [51] $P_{\mathrm{e}}(\ell) \equiv \sum_k P(\ell, k) = \ell P(\ell)/\langle k\rangle$, where $\langle k\rangle = \sum_k kP(k)$ is the mean degree. The conditional degree distribution is related to the joint distribution via

$$P(\ell|k) = \frac{P(\ell, k)}{P_{\mathrm{e}}(k)} = \frac{\langle k\rangle P(\ell, k)}{kP(k)} = \frac{E_{\ell,k}}{kP(k)N}, \tag{31}$$

which can be further simplified in the uncorrelated network limit where $P(\ell, k) \approx P_{\mathrm{e}}(k)P_{\mathrm{e}}(\ell)$:

$$P(\ell|k) \approx \frac{\ell P(\ell)}{\langle k\rangle}. \tag{32}$$

Equations (31) or (32) can be used as a simpler replacement for $P(\ell|k)$ in (2) and (3) if $E_{\ell,k}/(kN_k)$ is not directly accessible. For example, for an uncorrelated network (*i.e.*, for $P(\ell|k) = \ell P(\ell)/\langle k\rangle$), we find

$$\frac{\mathrm{d}i_k^{\mathrm{u}}(t)}{\mathrm{d}t} = \beta^{\mathrm{u}} \frac{ks_k(t)}{\langle k\rangle} \sum_\ell \ell i_\ell^{\mathrm{u}}(t) - \gamma^{\mathrm{u}} i_k^{\mathrm{u}}(t), \tag{33}$$

where we have set testing rates $f_k(0) = 0$ at the start of the infection. According to [24], we define

$$I^{\mathrm{u}}(t) := \sum_k i_k^{\mathrm{u}}(t), \quad J^{\mathrm{u}}(t) := \sum_k k i_k^{\mathrm{u}}(t) \tag{34}$$

and obtain

$$\begin{aligned} \frac{\mathrm{d}I^{\mathrm{u}}(t)}{\mathrm{d}t} &= \beta^{\mathrm{u}} J^{\mathrm{u}}(t) - \gamma^{\mathrm{u}} I^{\mathrm{u}}(t), \\ \frac{\mathrm{d}J^{\mathrm{u}}(t)}{\mathrm{d}t} &= \beta^{\mathrm{u}} \frac{\langle k^2 \rangle}{\langle k \rangle} J(t) - \gamma^{\mathrm{u}} J^{\mathrm{u}}(t). \end{aligned} \tag{35}$$

We perform a linear stability analysis about the disease-free state $(I^*, J^*) = (0, 0)$ and find the eigenvalues to (35):

$$\lambda_{\pm} = -\gamma^{\mathrm{u}} \pm \beta^{\mathrm{u}} \frac{\langle k^2 \rangle}{\langle k \rangle}. \tag{36}$$

The transition from negative to positive eigenvalues occurs for $-\gamma^{\mathrm{u}} + \beta^{\mathrm{u}} \langle k^2 \rangle / \langle k \rangle = 0$. Hence, the basic reproduction number is

$$\mathcal{R}_0 = \frac{\beta^{\mathrm{u}}}{\gamma^{\mathrm{u}}} \frac{\langle k^2 \rangle}{\langle k \rangle} = \frac{\beta^{\mathrm{u}}}{\gamma^{\mathrm{u}}} \left(\langle k \rangle + \frac{\mathrm{Var}[k^2]}{\langle k \rangle} \right). \tag{37}$$

If we use the conditional degree distribution $P(\ell|k) = (\ell - 1)P(\ell)/\langle k \rangle$ proposed by Kiss *et al.* [24] to account for a reduction in neighboring susceptible vertices, the corresponding basic reproduction number is modified to

$$\mathcal{R}_0^{\mathrm{Kiss}} = \frac{\beta^{\mathrm{u}}}{\gamma^{\mathrm{u}}} \left(\langle k \rangle - 1 + \frac{\mathrm{Var}[k^2]}{\langle k \rangle} \right). \tag{38}$$

The mean degrees of the BA and SBM networks are 3.77 and 23.14, and the variances for the BA and SBM networks are 20.40 and 36.62, respectively. Using the values $\gamma^{\mathrm{u}} = 14^{-1}/\mathrm{day}$ and $\beta^{\mathrm{u}} = 0.0411/\mathrm{day}$ for the BA network, we find that the basic reproduction numbers $\mathcal{R}_0 = 5.361$ and $\mathcal{R}_0^{\mathrm{Kiss}} = 4.777$ are larger than 4.5, the value we used to determine β^{u} according to the next-generation matrix method (see (7) and (8)). The observed approximation errors in (37) and (38) are a consequence of the assumption that the underlying network is uncorrelated. For the SBM network, we find $\mathcal{R}_0 = 4.499$ and $\mathcal{R}_0^{\mathrm{Kiss}} = 4.317$, close to the 4.5 value used to find $\beta^{\mathrm{u}} = 0.0130$ using (7) and (8).

To summarize, our comparison shows that in the SBM model where the degrees of neighbors are uncorrelated, (37) and (38) give close approximations of the actual reproduction number calculated from the next-generation matrix method (7). For the BA network, degree correlations make (37) and (38) overestimate the actual reproduction number. Therefore, we recommend using the next-generation matrix method to numerically determine the basic reproduction number unless degree correlations are weak and (37) and (38) can provide accurate estimates of $\mathcal{R}_0$.

Appendix B
Optimal Testing and Vaccination Algorithms

Below, we explicitly give the pseudo-code for the testing and quarantine model based on Pontryagin's maximum principle.

Algorithm 2: Pseudo-Code of Q-Learning in Testing Resource Allocation.

Initialize $F, \delta, C, s_k(0), i_k^u(0), i_k^*(0), \beta^u, \beta^*, \gamma^u, \gamma^*, M, \epsilon$
Initialize replay memory D
Randomly initialize the hyperparameter set $\Theta^- \leftarrow \Theta$ for evaluating the action value function $Q^*(\mathcal{S}, \mathcal{A}; \Theta)$
for episode $\ell = 1 : M$ **do**
  Initialize $\mathcal{S}_0$
  for $t = 0 : T_{\max} - 1$ **do**
    With probability ϵ, randomly select an action a_i
    otherwise select $\mathcal{A}_t = \mathrm{argmax}_{\mathcal{A}} Q(\mathcal{S}_t, \mathcal{A}; \Theta)$
    Execute action $\mathcal{A}_t$ and observe reward R_t and state $\mathcal{S}_{t+1}$
    Store transition $(\mathcal{S}_t, \mathcal{A}_t, R_t, \mathcal{S}_{t+1})$ in D
    Sample random minibatch of transitions $(\mathcal{S}_j, \mathcal{A}_j, R_j, \mathcal{S}_{j+1})$ from D
    if $j = T_{\max} - 1$ **then**
      Set $y_j = R_j$
    else
      Set $y_j = R_j + \delta \max_{\mathcal{A}'} \hat{Q}(\mathcal{S}_{j+1}, \mathcal{A}'; \Theta^-)$
      Perform a gradient descent step on the minibatch $\sum_j [y_j - Q(\mathcal{S}_j, \mathcal{A}_j; \Theta)]^2$ with respect to the network hyperparameter set Θ
    end if
  end for
  Every C steps reset $\Theta^- \leftarrow \Theta$
end for

Appendix C
Reinforcement-Learning Strategy

To identify effective testing and vaccination strategies, we also investigated reinforcement-learning (RL) approaches. RL explores the space of all possible actions and directly optimizes the loss functions for testing and vaccination defined in (11) and (25). Here, we use an RL approach with experience replay to learn both the optimal testing strategy in (2)–(5) and the optimal vaccination strategy in (20)–(22).

Typically, applying a policy-gradient method to a continuous action space will usually yield poor results due to the inability of such methods to explore the whole space. However, using our previous results based on PMP, we know that the optimal strategy is always obtained by maximizing the testing and vaccination rates for subpopulations presumed to be at a higher risk.

Therefore, we do not need to explore the whole space of all possible actions. Instead, from (16), (26), we can restrict our strategy space to the extreme points[1] of the set

$$\{(f_k)|_{k=1}^K | \sum_{k=1}^K f_k = F(t), f_{\min} \leq \frac{f_k}{N_k} \leq f_{\max}\} \tag{39}$$

for determining the testing-resource allocation and the extreme points of the set

[1] Extreme points are points in a set that cannot be written as a nontrivial convex linear combination of any other points in the same set.

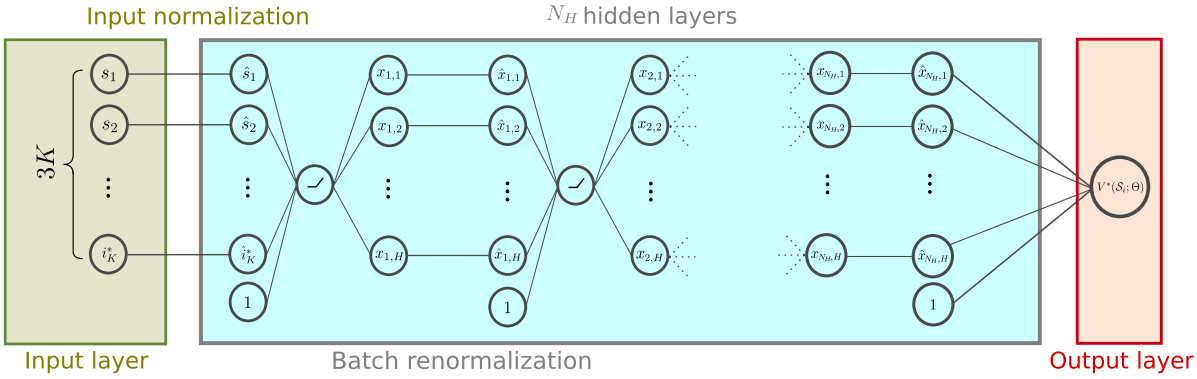

Fig. 6. Illustration of the neural network used to identify effective testing and vaccination strategies. The inputs of the input layer are $(s_1(t_i),\ldots,s_K(t_i),i_1^{\mathrm{u}}(t_i),\ldots,i_K^{\mathrm{u}}(t_i),i_1^*(t_i),\ldots,i_K^*(t_i))\in\mathbb{R}^{3K}$. For each hidden layer i ($1\le i\le N_H$), we normalize the corresponding outputs $x_{i,j}$ for all samples in a minibatch such that the resulting values $\hat{x}_{i,j}$ have zero mean and unit variance. These values are used as inputs to a rectified linear unit (ReLU) activation function in the next hidden layer. Neurons labeled 1 are bias terms. The output $V^*(\mathcal{S}_i;\Theta)$ is an estimate of the state-value function under the optimal policy (see (44)), where Θ denotes the set of hyperparameters.

$$\{(v_k)|_{k=1}^K|\sum_{k=1}^K v_k=V(t),v_{\min}\le\frac{v_k}{Ns_k(t)}\le v_{\max}\} \tag{40}$$

for determining vaccination-resource allocation at each step. The set of extreme points represents all strategies that maximize the testing/vaccination rates for some groups and minimize them for other groups. Such strategies also cannot be written as nontrivial convex combinations of other strategies. By confining ourselves to extreme points, the possible action space is reduced to a finite set on which we perform RL.

Since the curse of dimensionality increases the number of all possible strategies exponentially with K, we further restrict our RL approach to networks with degree cutoff $K=20$. This additional constraint allows us to perform RL with a computation time of about 30 days for the testing model on the BA network, 3 days for the testing model on the SBM network, 6 hours for the vaccination model on the BA network, and 2 hours for the vaccination model on an SBM network. All computations are performed using Python 3.8.10 on a laptop with a 4-core Intel(R) Core(TM) i7-8550 U CPU @ 1.80 GHz.

To identify effective testing and vaccination strategies, we use the reward functions (11) and (25). We define the reward at time $t_i=i\Delta t$ as

$$R(\mathcal{S}_i,\mathcal{A}_i)=\sum_{k=1}^K[s_k(t_{i+1})-s_k(t_i)], \tag{41}$$

the "negative" of the number of total infections during the time period $[t_i,t_{i+1})$. Here, the state $\mathcal{S}_i$ and action $\mathcal{A}_i$ are

$$\begin{aligned}\mathcal{S}_i&=(s_1(t_i),\ldots,s_K(t_i),i_1^{\mathrm{u}}(t_i),\ldots,i_K^{\mathrm{u}}(t_i),\\&\quad i_1^*(t_i),\ldots,i_K^*(t_i))\in\mathbb{R}^{3K},\\\mathcal{A}_i&=(f_1(t_i),\ldots,f_K(t_i))\in\mathbb{R}^K\end{aligned} \tag{42}$$

for the testing model (2)–(5) and

$$\begin{aligned}\mathcal{S}_i&=(s_1(t_i),\ldots,s_K(t_i),i_1(t_i),,i_K(t_i))\in\mathbb{R}^{2K},\\\mathcal{A}_i&=(v_1(t_i),\ldots,v_K(t_i))\in\mathbb{R}^K\end{aligned} \tag{43}$$

for the vaccination model (20)–(22). We recursively define the state-value function under a certain policy π to be

$$V^\pi(\mathcal{S}_i,i)=\begin{cases}V^\pi(\mathcal{S}_{i+1})\delta+R(\mathcal{S}_i,\pi(\mathcal{S}_i)), & t_i<T_{\max},\\0, & t_i=T_{\max},\end{cases} \tag{44}$$

where $\pi(\mathcal{S}_i)$ is the action determined under policy π given $\mathcal{S}_i$ and $\delta\in(0,1]$ is a discount factor. We also define the action-value function to be

$$Q^\pi(\mathcal{S}_i,\mathcal{A}_i,i)=\begin{cases}V^\pi(\mathcal{S}_{i+1})\delta+R(\mathcal{S}_i,\mathcal{A}_i), & t_i<T_{\max}-1,\\R(\mathcal{S}_i,\mathcal{A}_i), & t_i=T_{\max}-1.\end{cases} \tag{45}$$

We use Q^* and V^* to denote the action-value and state-value functions, respectively, under the best policy and apply the deep Q-learning algorithm, which has been used to find the RL strategies that can approximate optimal strategies of certain Atari 2600 games [52]. Here, we use a neural network with a hyperparameter set Θ, representing neural-network weights and biases to estimate the action-value function under the best policy $Q^*(\mathcal{S},\mathcal{A};\Theta)$, which is improved over epochs by Algorithm 2.

An illustration of the neural network, its layers, and activation functions, is shown in Fig. 6. We use another neural network with a hyperparameter set Θ^- updated every $C=4$ steps to match Θ. The neural network contains $N_H=4$ hidden layers with $H=30$ neurons in each layer. The input data is the state at the i^{th} step $\mathcal{S}_i$, and the output is $V^*(\mathcal{S}_i;\Theta)$, the prediction for the optimal state-value function generated by the neural network. In each layer, the batch normalization technique is used before a rectified linear unit (ReLU) function is applied as an activation function. We compare the optimal strategies based on the PMP approach from Algorithm 1 with the RL strategies that are based on Algorithm 2. We set $T=100$ and $\Delta t=1$ so that the strategy is updated every day. Here, we use $f_{\min}=0.002/\mathrm{day}$, $f_{\max}=0.4/\mathrm{day}$. We use (7) with $\gamma^{\mathrm{u}}=(1/14)/\mathrm{day}$ to calculate $\beta^{\mathrm{u}}=0.0703/\mathrm{day}$ for the

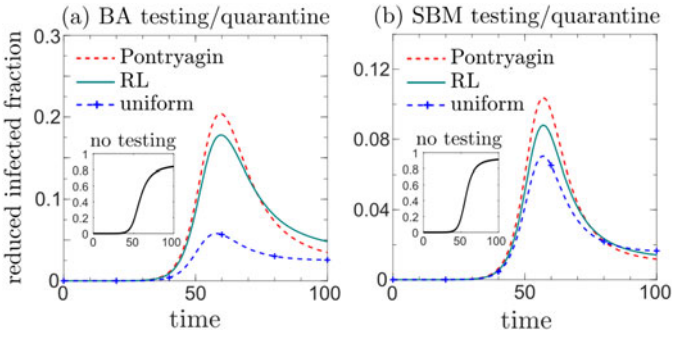

Fig. 7. Reduction in fractions of infected individuals calculated as the difference between the fractions infected obtained with testing and without testing for the BA network shown in (a) and the SBM network shown in (b). The optimal control approach based on PMP reduces early infections the most. RL outperforms uniform testing in reducing the number of early-stage infections. Additionally, the effect of the optimal strategy is more striking in the BA network because it has a more heterogeneous node degree distribution.

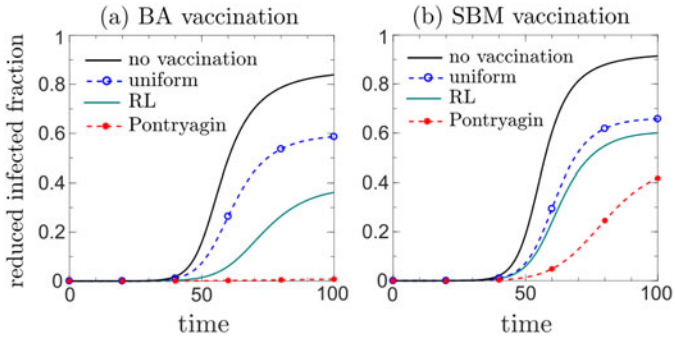

Fig. 8. Fractions of infected individuals with vaccination and without vaccination for the BA network is shown in (a) and the SBM network is shown in (b). The optimal control approach using PMP can most effectively reduce infections for both networks and successfully suppress the spreading of the disease in the BA network. On the other hand, although not as good as the PMP-optimal strategies, the strategies obtained by the RL algorithm Algorithm 2 can obviously reduce infections compared to the uniform vaccination rate strategy. As with testing, we observe that the effect of optimal vaccination is more pronounced in the BA network than in the SBM network.

$K = 20$ BA network and $\beta^{\mathrm{u}} = 0.0632/\mathrm{day}$ for the $K = 20$ SBM network. Both PMP and RL strategies are also compared to the uniform testing strategy (19). For RL, we train the underlying neural network for $M = 100$ epochs using Algorithm 2. Fig. 7 shows the differences between the infected fractions in simulations with and without testing. The PMP-based optimal control reduces early infections the most for both BA and SBM networks. Early infections contribute more to the loss function (11) since we set the discount factor to $\delta = 0.95$. We also observe that RL-based testing strategies outperform uniform testing in reducing early-stage infections. Comparing Fig. 7(a) and (b), the effect of the optimal vaccination strategy in the BA network is more pronounced than that in the SBM network. In the BA network, node degrees are more heterogeneous and most nodes have small degrees, indicating that epidemic spreading can be controlled effectively as long as the few high-degree nodes are monitored and tested. Finally, comparing the result of the optimal-control approach in Fig. 7 with Fig. 2, we observe that with a smaller K in the SBM network, the effect of the optimal vaccination strategy is less apparent because node degrees are more homogeneous.

Next, we compared the PMP approach with the RL approach for the optimal vaccination strategy model (20)–(22). Here, we set $v_{\min} = 0.0001, v_{\max} = 1$. For both networks, the optimal vaccination strategy obtained using PMP can most effectively reduce the initial infections because early infections have a higher weight in the loss function (25). Reinforcement-learning-based vaccination policies can also reduce initial infections, but the reduction is less than that of the PMP approach. Comparing Fig. 8(a) and (b), we again observe that the effect of the optimal vaccination strategy for the BA network is more pronounced than that of the SBM network because the BA network has a more heterogeneous degree and is dominated by small-degree nodes.

To summarize, the controls derived from PMP are more effective than those based on RL. One limitation of RL-based interventions is that the possible action space that needs to be explored is usually large. However, based on our PMP results, we can constrain the action space before the learning process. Such PMP-informed constraints allow us to explore just the extreme points of the whole action space and thus make the training more efficient. Yet, the total number of possible actions grows exponentially with the maximal degree K and the strategy obtained by the RL approach will probably be only locally optimal, violating the PMP condition and thus under performing PMP. Nonetheless, RL could be useful if a procedure for computing an explicit solution cannot be formulated.

Appendix D
Simulations of Corresponding Stochastic Models

We impose the optimal testing and vaccination strategies derived from applying PMP to the ODE system (2)–(5) and (20)–(22) on a simple discrete stochastic model and compare the resulting total infections. The corresponding optimal testing or vaccination is implemented by probabilistically testing or vaccinating each selected subpopulation. For example, in the testing model, we can employ a rejection-free event-based Monte-Carlo (MC) algorithm [53] that implements a testing strategy.

For initial conditions, we randomly choose two nodes with degree $k = 10$ to be infected. Correspondingly, for the deterministic ODE models, we set $s_k(0) = p(k) - \frac{2}{N}\mathbb{1}_{k,10}$, $i_k^{\mathrm{u}}(0) = \frac{2}{N}\mathbb{1}_{k,10}$, $i_k^*(0) = 0$ for the testing model and $s_k(0) = p(k) - \frac{2}{N}\mathbb{1}_{k,10}$, $i_k(0) = \frac{2}{N}\mathbb{1}_{k,10}$ for the vaccination model. We set the recovery rates $\gamma = \gamma^{\mathrm{u}} = \gamma^* = (14)^{-1}/\mathrm{day}$ and use the same reproduction number $\mathcal{R}_0(\beta^{\mathrm{u}}) = 4.5$ to calculate the unconstrained infection rates for the two networks from (7). The loss functions defined for the testing and vaccination models in (11) and (25) are plotted below.

From Fig. 9, the deterministic ODE models tend to overestimate the loss functions since all subpopulations are well mixed by the conditional degree distribution function $P(\ell|k)$ and therefore a single infected node could have an impact on the whole system. This difference arises because in a fully discrete realization of a BA or SBM network, each node can be in only one of three or four states and the disease may never arrive at certain

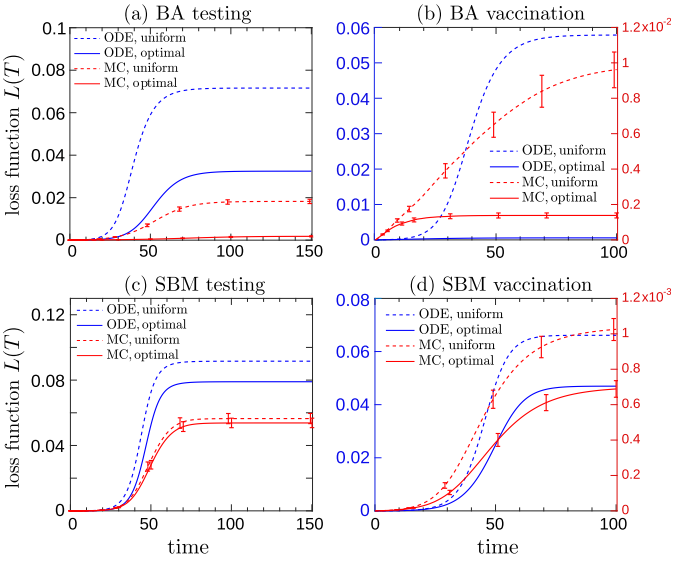

Fig. 9. Loss functions associated with the deterministic ODE models (2)–(5) and (20)–(22), and the corresponding stochastic models. We apply PMP-based (solid lines) and uniform (dashed lines) testing and vaccination protocols. Panels (a) and (b) show the loss functions (11) and (25) associated with testing and vaccination interventions in a BA network. Results from the ODE models are shown in blue while the loss functions derived from the simulated stochastic model are shown in red. Panels (c) and (d) show loss functions for the testing and vaccination models in the SBM network. Note the different scales for the ODE (blue, left) and the MC (red, right) results. The loss functions of the discrete stochastic models are obtained by averaging over 100 trajectories with the standard error of the mean (standard deviation of means divided by $\sqrt{N}$) indicated by the error bars. For both networks, the deterministic ODE models yield larger losses than those obtained from averaging MC trajectories. For both deterministic ODEs and stochastic systems, the loss functions during optimal testing and vaccination are much smaller than when testing and vaccination are uniformly applied.

critical nodes, significantly delaying its spread allowing the overall infection to dissipate before ever reaching portions of the network. In contrast, the mass-action ODE model allows all nodes to be partially infected, allowing continuous transmission of the disease. Therefore, more network measures may be needed to accurately quantify the dynamics of disease spread across discrete agent-based network models. Higher-order interactions beyond the pairwise conditional degree distribution [54]–[56] could be helpful in explaining the discrepancy between deterministic ODE and stochastic models and in estimating optimal policies in the fully stochastic context.

Nonetheless, Fig. 9 shows that the PMP-based interventions that we derived in the main text are also more effective than uniform testing and vaccination strategies in the stochastic agent-based model. This loss function reduction arises for both the BA and SBM networks. Thus, the optimal testing and vaccination strategies obtained from the deterministic model outperforms uniform testing and vaccination strategies even when applied on discrete stochastic network models, representing a reasonable starting point for approximating optimal strategies within agent-based discrete systems.

ACKNOWLEDGMENT

The authors thank S. K. Lyons for help in editing the manuscript and the reviewers for their valuable comments.

REFERENCES

[1] B. Abdalhamid, C. R. Bilder, E. L. McCutchen, S. H. Hinrichs, S. A. Koepsell, and P. C. Iwen, "Assessment of specimen pooling to conserve SARS-CoV-2 testing resources," *Amer. J. Clin. Pathol.*, vol. 153, no. 6, pp. 715–718, 2020.

[2] I. Yelin *et al.*, "Evaluation of COVID-19 RT-qPCR test in multi sample pools," *Clin. Infect. Dis.*, vol. 71, no. 16, pp. 2073–2078, 2020.

[3] B. J. Quilty *et al.*, "Quarantine and testing strategies in contact tracing for SARS-CoV-2: A modelling study," *Lancet Public Health*, vol. 6, pp. e175–e183, 2021.

[4] T. Schneider *et al.*, "Epidemic management and control through risk-dependent individual contact interventions," 2021, *arXiv:2109.10970*.

[5] S. Moore, E. M. Hill, L. Dyson, M. J. Tildesley, and M. J. Keeling, "Modelling optimal vaccination strategy for SARS-CoV-2 in the U.K.," *PLoS Comput. Biol.*, vol. 17, 2021, Art. no. e1008849.

[6] J. Müller, "Optimal vaccination patterns in age-structured populations," *SIAM J. Appl. Math.*, vol. 59, no. 1, pp. 222–241, 1998.

[7] Z. Zhao *et al.*, "The optimal vaccination strategy to control COVID-19: A modeling study based on the transmission scenario in Wuhan City, China," *Infect. Dis. Poverty*, vol. 10, 2020, Art. no. 140.

[8] R. P. Curiel and H. G. Ramírez, "Vaccination strategies against COVID-19 and the diffusion of anti-vaccination views," *Sci. Rep.*, vol. 11, 2021, Art. no. 6626.

[9] M. A. Acuña-Zegarra, S. Díaz-Infante, D. Baca-Carrasco, and D. Olmos-Liceaga, "COVID-19 optimal vaccination policies: A modeling study on efficacy, natural and vaccine-induced immunity responses," *Math. Biosci.*, vol. 337, 2021, Art. no. 108614.

[10] CDC, "New COVID-19 variants," 2021. Accessed: Jan. 15, 2021. [Online]. Available: https://www.cdc.gov/coronavirus/2019-ncov/transmission/variant.html

[11] W. Choi and E. Shim, "Optimal strategies for social distancing and testing to control COVID-19," *J. Theor. Biol.*, vol. 512, 2021, Art. no. 110568.

[12] S. Nowak, P. N. de Lima, and R. Vardavas, "Should we mitigate or suppress the next pandemic? Time-horizons and costs shape optimal social distancing strategies," *medRxiv*, 2021, doi: 10.1101/2021.11.14.21266322.

[13] L. Bolzoni, E. Bonacini, R. Della Marca, and M. Groppi, "Optimal control of epidemic size and duration with limited resources," *Math. Biosci.*, vol. 315, 2019, Art. no. 108232.

[14] P. Ogren and C. F. Martin, "Optimal vaccination strategies for the control of epidemics in highly mobile populations," in *Proc. 39th IEEE Conf. Decis. Control*, 2000, pp. 1782–1787.

[15] Z. Wang, C. Xia, Z. Chen, and G. Chen, "Epidemic propagation with positive and negative preventive information in multiplex networks," *IEEE Trans. Cybern.*, vol. 51, no. 3, pp. 1454–1462, Mar. 2021.

[16] F. Liu and M. Buss, "Optimal control for heterogeneous node-based information epidemics over social networks," *IEEE Trans. Control Netw. Syst.*, vol. 7, no. 3, pp. 1115–1126, Sep. 2020.

[17] B. Wang, Y. Sun, T. Q. Duong, L. D. Nguyen, and L. Hanzo, "Risk-aware identification of highly suspected COVID-19 cases in social IoT: A joint graph theory and reinforcement learning approach," *IEEE Access*, vol. 8, pp. 115655–115661, 2020.

[18] R. M. May and R. M. Anderson, "The transmission dynamics of human immunodeficiency virus (HIV)," *Philos. Trans. Roy. Soc. London, B. Biol. Sci.*, vol. 321, no. 1207, pp. 565–607, 1988.

[19] M. Barthélemy, A. Barrat, R. Pastor-Satorras, and A. Vespignani, "Velocity and hierarchical spread of epidemic outbreaks in scale-free networks," *Phys. Rev. Lett.*, vol. 92, no. 17, 2004, Art. no. 178701.

[20] R. Pastor-Satorras and A. Vespignani, "Epidemic dynamics and endemic states in complex networks," *Phys. Rev. E*, vol. 63, no. 6, 2001, Art. no. 066117.

[21] R. Pastor-Satorras and A. Vespignani, "Epidemic spreading in scale-free networks," *Phys. Rev. Lett.*, vol. 86, no. 14, 2001, Art. no. 3200.

[22] M. Boguná, R. Pastor-Satorras, and A. Vespignani, "Absence of epidemic threshold in scale-free networks with degree correlations," *Phys. Rev. Lett.*, vol. 90, no. 2, 2003, Art. no. 028701.

[23] M. Barthélemy, A. Barrat, R. Pastor-Satorras, and A. Vespignani, "Dynamical patterns of epidemic outbreaks in complex heterogeneous networks," *J. Theor. Biol.*, vol. 235, no. 2, pp. 275–288, 2005.

[24] I. Z. Kiss, D. M. Green, and R. R. Kao, "The effect of contact heterogeneity and multiple routes of transmission on final epidemic size," *Math. Biosci.*, vol. 203, no. 1, pp. 124–136, 2006.

[25] J. Lindquist, J. Ma, P. Van den Driessche, and F. H. Willeboordse, "Effective degree network disease models," *J. Math. Biol.*, vol. 62, no. 2, pp. 143–164, 2011.

[26] M. Newman, *Networks*. London, U.K.: Oxford Univ. Press, 2018.
[27] L. Böttcher, M. R. D'Orsogna, and T. Chou, "Using excess deaths and testing statistics to determine COVID-19 mortalities," *Eur. J. Epidemiol.*, vol. 36, pp. 545–558, 2021.
[28] L. Böttcher, M. R. D'Orsogna, and T. Chou, "A statistical model of COVID-19 testing in populations: Effects of sampling bias and testing errors," *Philos. Trans. Roy. Soc. A*, vol. 380, 2021, Art. no. 20210121.
[29] A.-L. Barabási and R. Albert, "Emergence of scaling in random networks," *Science*, vol. 286, no. 5439, pp. 509–512, 1999.
[30] R. Albert and A.-L. Barabási, "Statistical mechanics of complex networks," *Rev. Modern Phys.*, vol. 74, no. 1, pp. 47–97, 2002.
[31] P. W. Holland, K. B. Laskey, and S. Leinhardt, "Stochastic blockmodels: First steps," *Social Netw.*, vol. 5, no. 2, pp. 109–137, 1983.
[32] C. Brown, A. Noulas, C. Mascolo, and V. Blondel, "A place-focused model for social networks in cities," in *Proc. Int. Conf. Social Comput.*, 2013, pp. 75–80.
[33] Y. Zhao *et al.*, "Consistency of community detection in networks under degree-corrected stochastic block models," *Ann. Statist.*, vol. 40, no. 4, pp. 2266–2292, 2012.
[34] P. Van den Driessche and J. Watmough, "Reproduction numbers and sub-threshold endemic equilibria for compartmental models of disease transmission," *Math. Biosci.*, vol. 180, no. 1-2, pp. 29–48, 2002.
[35] O. Diekmann, J. A. P. Heesterbeek, and J. A. Metz, "On the definition and the computation of the basic reproduction ratio R_0 in models for infectious diseases in heterogeneous populations," *J. Math. Biol.*, vol. 28, no. 4, pp. 365–382, 1990.
[36] Y. Wang *et al.*, "Estimating the basic reproduction number of COVID-19 in Wuhan, China," *Chin. J. Epidemiol.*, vol. 41, no. 4, pp. 476–479, 2020.
[37] M. D'Arienzo and A. Coniglio, "Assessment of the SARS-CoV-2 basic reproduction number, R_0, based on the early phase of COVID-19 outbreak in Italy," *Biosafety Health*, vol. 2, no. 2, pp. 57–59, 2020.
[38] G. G. Katul, A. Mrad, S. Bonetti, G. Manoli, and A. J. Parolari, "Global convergence of COVID-19 basic reproduction number and estimation from early-time SIR dynamics," *PLoS One*, vol. 15, no. 9, 2020, Art. no. e0239800.
[39] M. P. Barman, T. Rahman, K. Bora, and C. Borgohain, "COVID-19 pandemic and its recovery time of patients in India: A pilot study," *Diabetes Metabolic Syndrome: Clin. Res. Rev.*, vol. 14, no. 5, pp. 1205–1211, 2020.
[40] S. M. Aseev and A. V. Kryazhimskii, "The Pontryagin maximum principle and optimal economic growth problems," *Proc. Steklov Inst. Math.*, vol. 257, no. 1, pp. 1–255, 2007.
[41] C. Wilasang *et al.*, "Reduction in effective reproduction number of COVID-19 is higher in countries employing active case detection with prompt isolation," *J. Travel Med.*, vol. 27, no. 5, pp. 1–3, 2020.
[42] A. Hyafil and D. Moriña, "Analysis of the impact of lockdown on the reproduction number of the SARS-CoV-2 in Spain," *Gaceta Sanitaria*, vol. 35, pp. 453–458, 2020.
[43] G. Zaman, Y. H. Kang, G. Cho, and I. H. Jung, "Optimal strategy of vaccination & treatment in an SIR epidemic model," *Math. Comput. Simul.*, vol. 136, pp. 63–77, 2017.
[44] N. Peiffer-Smadja, S. Rozencwajg, Y. Kherabi, Y. Yazdanpanah, and P. Montravers, "COVID-19 vaccines: A race against time," *Anaesthesia, Crit. Care Pain Med.*, vol. 40, no. 2, 2021, Art. no. 100848.
[45] L. Böttcher and J. Nagler, "Decisive conditions for strategic vaccination against SARS-CoV-2," *Chaos: An Interdiscipl. J. Nonlinear Sci.*, vol. 31, no. 10, 2021, Art. no. 101105.
[46] E. Mathieu *et al.*, "A global database of COVID-19 vaccinations," *Nature Hum. Behav.*, vol. 5, pp. 947–953, 2021.
[47] T. Asikis, L. Böttcher, and N. Antulov-Fantulin, "Neural ordinary differential equation control of dynamics on graphs," *Phys. Rev. Res.*, to be published.
[48] L. Böttcher, N. Antulov-Fantulin, and T. Asikis, "AI Pontryagin or how neural networks learn to control dynamical systems," *Nature Commun.* vol. 13, 2022, Art. no. 333.
[49] M. Piraveenan, M. Prokopenko, and L. Hossain, "Percolation centrality: Quantifying graph-theoretic impact of nodes during percolation in networks," *PLoS One*, vol. 8, no. 1, 2013, Art. no. e53095.
[50] G. Ódor, D. Czifra, J. Komjáthy, L. Lovász, and M. Karsai, "Switchover phenomenon induced by epidemic seeding on geometric networks," *Proc. Nat. Acad. Sci. USA*, vol. 118, no. 41, 2021, Art. no. e2112607118.
[51] S. Weber and M. Porto, "Generation of arbitrarily two-point-correlated random networks," *Phys. Rev. E*, vol. 76, no. 4, 2007, Art. no. 046111.
[52] V. Mnih *et al.*, "Human-level control through deep reinforcement learning," *Nature*, vol. 518, no. 7540, pp. 529–533, 2015.
[53] D. T. Gillespie, "Exact stochastic simulation of coupled chemical reactions," *J. Phys. Chem.*, vol. 81, no. 25, pp. 2340–2361, 1977.
[54] I. Iacopini, G. Petri, A. Barrat, and V. Latora, "Simplicial models of social contagion," *Nature Commun.*, vol. 10, no. 1, 2019, Art. no. 2485.
[55] F. Battiston *et al.*, "Networks beyond pairwise interactions: Structure and dynamics," *Phys. Rep.*, vol. 874, pp. 1–92, 2020.
[56] W. Li, X. Xue, L. Pan, T. Lin, and W. Wang, "Competing spreading dynamics in simplicial complex," *Appl. Math. Computation*, vol. 412, 2022, Art. no. 126595.

Mingtao Xia received his bachelor's degree in information and computing science from Peking University, Beijing, China, in 2019. He is currently working toward his Ph.D. degree in the Department of Mathematics, University of California, Los Angeles, CA, USA. His research interests include mathematical modeling and computational methods.

Lucas Böttcher received his Ph.D. degree in theoretical physics and applied mathematics from Eidgenössische Technische Hochschule (ETH) Zurich, Zurich, Switzerland, in 2018. He is an Assistant Professor of computational social science in the Frankfurt School of Finance & Management, Frankfurt, Germany. He joined the Department of Computational Medicine, University of California, Los Angeles, CA, USA, as a Swiss National Fund fellow in 2020. His research interests include applied mathematics, statistical mechanics, and machine learning.

Tom Chou received his Ph.D. degree in physics from Harvard University, Cambridge, MA, USA. He was a Postdoctoral Researcher at Cornell University, Ithaca, NY, USA, University of Cambridge, Cambridge, U.K., and Stanford University, Stanford, CA, USA. He is currently a Professor in the Departments of Computational Medicine and Mathematics, University of California, Los Angeles, CA, USA. His research interests include statistical physics, applied mathematics, and mathematical biology.