Spreading processes with layer-dependent population heterogeneity over multilayer networks

Yurun Tian and Osman Yağan
Department of Electrical and Computer Engineering
Carnegie Mellon University, Pittsburgh, PA, 15213 USA
{yurunt, oyagan}@andrew.cmu.edu

***Abstract*—The study of spreading processes on complex networks has gained significant attention recently. For example, bond percolation models considering population heterogeneity have been used to provide insights into disease spread and misinformation control. However, most of these studies focus on single-layer contact networks. In our work, we examine how the spreading process is impacted by multiple layers, considering layer-dependent population heterogeneity from a principled, mathematical perspective. Using SIR dynamics, we derive expressions for three key epidemiological measures: the probability of emergence, the epidemic threshold, and the expected epidemic size. Through extensive simulations, we demonstrate that our analytical results match the numerical results near-perfectly in the finite nodes regime. Moreover, we show that the reductions to single-layer network models that discount the multi-layer network structure will lead to incorrect predictions, which warrants a separate analysis for the multi-layer network model. Furthermore, we investigate the impact of layer-dependent population heterogeneity and identify important factors for developing effective and economical layer-oriented spreading control strategies. These findings reveal the interplay among the multi-layer network structures, transmission dynamics, and population heterogeneity in determining the final outcome of the spreading process. Overall, our work provides insights into developing and analyzing mitigation and control strategies for disease spread and information diffusion across multi-layer complex networks.**

***Index Terms*—Heterogeneous bond percolation, Branching Process, Population Heterogeneity, Multi-layer networks, Network Epidemics**

I. Introduction

The attention towards studies on spreading processes over complex networks has grown in recent years, driven by the impact of pandemics like COVID-19 and SARS, as well as concerns regarding misinformation diffusion [1]. Researchers have extensively examined mathematical models over complex networks to provide insights into the dynamics of spreading pathogens or information [2]–[5]. The susceptible-infectious-recovered (SIR) compartmental model, in particular, has received significant interest due to its ability to capture the propagation of both pathogens and information [6]–[8]. Additionally, its steady-state analysis is closely linked to *bond-percolation* over networks [8], [9].

This work was supported in part by the National Science Foundation through grants RAPID-2026985, RAPID-2026982, CCF-1813637, DMS-1811724, CNS-2027908, CCF-1917819, CNS-2041952, RAPID-2142997; the Army Research Office through grants # W911NF-20-1-0204, # W911NF-17-1-0587, and # W911NF-18-1-0325; the C3.ai Digital Transformation Institute; Google, LLC; James S. McDonnell Foundation 21st Century Science Initiative Collaborative Award in Understanding Dynamic and Multi-scale Systems.

More recently, there has been interest on studying the SIR spreading process with increasing complexity of the underlying contact network (e.g., clustered networks [10]–[12] and multi-layer networks [7], [13], [14]) and the *heterogeneity* of the population. For example, Tian et al. [15] investigated a SIR model with population heterogeneity that manifest from different types of masks that the individuals in the population might be wearing. More broadly, population heterogeneity can also arise from factors such as age, gender, socio-economic status, and access to healthcare and other resources [16]–[18] in the population. In the context of information diffusion, population heterogeneity becomes relevant as individuals may exhibit different tendencies in accepting and transmitting information based on their personalities and fact-checking behaviors [1], [19]. Allard et al. [6] also studied the SIR model with population heterogeneity and showed that their steady-state can be analyzed through a *semi-directed* bond percolation model.

This paper is motivated by the fact that most studies on spreading processes with population heterogeneity consider single-layer networks, including [15] - [16]. However, most real-world spreading processes take place over *multi-layer* networks. In viral spreading, different layers might represent viral spreading paths in different environments, e.g., community, school, workplace, etc, each with a different rate of viral transmissibility [20]. Similarly, (mis)information tends to spread over multiple social media platforms, each with different rates and dynamics of propagation. To our best knowledge, there have only been a few prior efforts [6], [21], [22] on studying the SIR model while incorporating both population heterogeneity and the multi-layer nature of the contact network. Bongiorno and Zino [21] proposed a model that incorporates both population heterogeneity and a multi-layer contact network, but they do not provide mathematical analysis for the three epidemic quantities and instead rely on simulation results. The work by Allard et al. [6] consider multi-type networks with arbitrary joint degree distribution. However, their work does not provide a detailed analysis on the impact of multi-layer network structures and the associated multi-layer transmission dynamics on the final spreading results. Tian et al. [22] studied the *Multi-layer Mask model* which takes into account both population heterogeneity and multi-layer contact network, but they assume population heterogeneity is layer-independent and thus lack of capability to investigate spreading mitigation strategies that target specific

layers, e.g., in some regions masks are mandated at school but not communities during pandemics [23].

Inspired by these, our main contribution is to provide a thorough analysis of the spreading process in a class of multi-layer networks considering layer-dependent population heterogeneity. Following Ref. [22], for illustrative purposes, we suppose population heterogeneity results from different types of masks (with different efficiencies) that individuals are wearing in the viral spreading context. We also reserve one *mask-type* to represent individuals who do not wear any mask. Specifically, we present the analytical solution of the multi-layer mask model for three key epidemiological quantities: probability of emergence (PE), epidemic threshold, and expected epidemic size (ES). The *emergence* of *epidemics* represents situations where the spreading process leads to a positive fraction of the population being infected in the limit of the number of nodes going to infinity. *Epidemics* refer to *large-scale* spreading events such as viral pandemics or information *memes*. Our analytical solutions disentangle the impact of multiple factors, including the multi-layer network structure, transmission dynamics, and layer-dependent population heterogeneity distribution, on these three quantities of interest. We also show that our results incorporate the existing layer-independent population heterogeneity models and single-layer contact network models. Extensive simulations validate our analytical results with a near-perfect match under different types of degree distributions.

Utilizing the analytical results, we first compare the dynamics of multi-layer networks and their monoplex (single-layer) projections. There has been recent interest [14] in understanding whether monoplex projection of a multi-layer network can still capture the essential properties of a spreading process. We find that projecting a multi-layer network into a single-layer network leads to significant differences in the dynamics warranting a separate multi-layer network structure analysis. Moreover, it is seen that by different ways of monoplex projection, the resulting single-layer networks might lead to bounds of the key epidemiological quantities such as probability and size of epidemics. Second, we explore layer-dependent population heterogeneity by investigating layer-oriented mitigation control policies. A comparison metric quantifying the expected cost of mask allocation is proposed and has been shown useful in characterizing different layer-oriented mitigation policies. We identify the *transmission power* of each layer in the multi-layer contact network, and the participation rate of nodes in the secondary layer as two crucial factors in developing *effective* and *economical* mitigation strategies. We believe these results provide fundamental insights into the spreading process over multi-layer complex networks when taking into account population heterogeneity. Thus, they might help develop mitigation and control strategies for disease spread and information diffusion.

II. Model

A. Contact network model

We create our two-layer contact network as follows. Given a population of size n, each node in the multi-layer network corresponds to an individual in $\mathcal{N}$. An edge exists between two nodes if there is a chance to transmit the spreading item (e.g., a piece of news, a virus) between them once in contact. The pattern of these potential transmission-causing contacts forms a network. Let $\mathbb{C}$ represent the first contact layer defined on the node set $\mathcal{N}$. Let $\mathbb{S}$ represent the second contact layer with the assumption that each node in $\mathcal{N}$ is a member of $\mathbb{S}$ with probability $\alpha \in (0, 1]$. Formally, we let

$$\mathbb{P}\left[i \in \mathcal{N}_S\right] = \alpha, \quad i = 1, \ldots, n \tag{1}$$

where $\mathcal{N}_S$ denotes the set of individuals who also participate the school layer. Edges belonging to network $\mathbb{C}$ (resp., $\mathbb{S}$) are noted as type-c (resp. type-s) edges.

We generate network $\mathbb{C}$ and $\mathbb{S}$ independently via the *configuration model* in line with prior work on stochastic epidemic models [24], [25]. The network topology is generated *randomly* from the given degree distribution. Besides their degree distributions, the graphs generated are assumed to be entirely random. The degree distributions for $\mathbb{C}$ and $\mathbb{S}$ are given as $\{p_k^c\}$ and $\{p_k^s\}$, where $k = 0, 1, ...$ p_k^c (resp. p_k^s) denotes the probability that an arbitrary node on network $\mathbb{C}$ (resp. $\mathbb{S}$) has degree k, i.e., it is connected to k other nodes via an *undirected* type-c (resp. type-s) edge. Our analytical solution is valid for any degree distributions with moments of arbitrary order being *finite*, e.g., Poisson degree distributions, power law degree distributions with exponential cut-off, etc [25]. Here we describe the procedure of generating layer $\mathbb{C}$ given its degree distribution $\{p_k^c\}$ using the configuration model. Generation of layer $\mathbb{S}$ follows the same. Given the degree distribution for layer $\mathbb{C}$, we draw a set of random numbers of size n from the degree distribution $\{p_k^c\}$, noted $\{k_c^i\}, i = 1, ..., n$. Each number can be viewed as *stubs* of edges, emerging from their perspective nodes. Then we randomly choose pairs of these stubs and place edges on the graph joining them up. This requires the sum $\sum_{i=1}^{n} k_c^i$ to be even because each edge added to the graph must have two ends. We draw a new set if the set $\{k_c^i\}$ sum to an odd number. This graph generation method guarantees that the graph generated is chosen uniformly at random from the set of all graphs with the selected degree sequence. Our results are averaged over the ensemble of possible graphs generated in this way. Note that the chances of a component containing a closed loop of edges goes as n^{-1}, which is negligible in the limit of large n [24]. In other words, the graph defined by the given degree distribution is locally tree-like when n goes to infinity. Throughout the paper, we are interested in the asymptotic case where n tends to infinity, and our claims are exact for sufficiently large n.

After generating layer $\mathbb{C}$ and layer $\mathbb{S}$ independently, the multi-layer network $\mathbb{H}$ is formed by taking the *disjoint* union of $\mathbb{C}$ and $\mathbb{S}$, i.e., $\mathbb{H} = \mathbb{C} \coprod \mathbb{S}$. In this setting, an arbitrary node i in $\mathcal{N}$ will have a *colored* degree represented by an integer vector $\boldsymbol{d^i} = \left[k_c^i, k_s^i\right]$, where k_c^i (resp., k_s^i) stands for its number of type-c edges (resp. type-s). The colored degree distribution for node i is thus given by:

$$p_{\boldsymbol{d^i}} = \left(\alpha p_{k_s^i}^c + (1-\alpha)\mathbf{1}\left[k_s^i = 0\right]\right) \cdot p_{k_c^i}^s, \quad \boldsymbol{d^i} = \left(k_c^i, k_s^i\right). \tag{2}$$

where the term $(1-\alpha)\mathbf{1}\left[k_s^i=0\right]$ accounts for the case where node i is not a member of layer $\mathbb{S}$, and its number of type-s edges is automatically zero.

B. Layer-dependent population heterogeneity

In their seminal work [8], Newman studied the SIR (susceptible-infectious-recovered) model over a contact network generated by the configuration model through bond percolation theory. Newman's model captures complex viral transmission mechanisms via the average transmissibility parameter *T*. Many works have incorporated various node-level heterogeneity based on Newman's model [15], [17], [22], among which Tian et al. [22] studied a SIR model with layer-independent population heterogeneity over multi-layer networks. Specifically, they study a two-layer contact network, where each layer is denoted as layer-$\mathbb{C}$ and layer-$\mathbb{S}$, respectively. They assume separate baseline transmissibilities T_c and T_s for layer-$\mathbb{C}$ and layer-$\mathbb{S}$. In other words, T_c and T_s are the probability of transmission (i.e., *transmissibility*) over type-c and type-s edges without any masks. Population heterogeneity is modeled via the inward and outward efficiency of different types of masks that individuals might wear. Assuming M types of masks, the mask distribution is given by $\boldsymbol{p}=\{p_1,...,p_M\}$ where p_i represents the fraction of individuals who wear masks of type-i. Let $\epsilon_{out,i}$ (resp. $\epsilon_{in,i}$) denote the outward (resp. inward) efficiency of mask type-i, where $0\leq\epsilon_{out,i}\leq 1$ and $0\leq\epsilon_{in,i}\leq 1$ for all $1\leq i\leq M$. Two transmissibility matrices $\mathbf{T_c}$ and $\mathbf{T_s}$, each of size $M\times M$, represent all types of transmission over layer $\mathbb{C}$ and $\mathbb{S}$, respectively. More specifically,

$$\mathbf{T_c}[i,j]=\left(1-\epsilon_{\text{out},i}\right)\left(1-\epsilon_{\text{in},j}\right)T_c,\ 1\leq i,j\leq M \quad (3)$$

$$\mathbf{T_s}[i,j]=\left(1-\epsilon_{\text{out},i}\right)\left(1-\epsilon_{\text{in},j}\right)T_s,\ 1\leq i,j\leq M \quad (4)$$

where $\mathbf{T_c}[i,j]$ (resp. $\mathbf{T_s}[i,j]$) gives the probability that, an infected node wearing a type-i mask transmits the virus/information to a susceptible node wearing a type-j mask given that they are connected by a type-c (resp. type-s) link.

However, assumptions are made regarding population heterogeneity and the corresponding transmission dynamics here. Mask distribution is assumed to be independent of the network structure and the spreading process. Each node keeps a consistent mask choice across layers. Nevertheless, in real life, population heterogeneity can be correlated with the network structure. For example, cases existed where surgical masks were mandated in schools but not in communities during the COVID-19 pandemic [23]; in information spreading, some people are more willing to speak out (or be silent) online than in real life, or vice versa [26]. Modeling layer-dependent population heterogeneity will provide a better understanding of the role of the multi-layer network structure in the spreading processes, and assist in developing layer-specific spread control strategies.

In this work, we incorporate layer-dependent population heterogeneity by associating node type with the node's layer-dependent mask-wearing behavior. In other words, nodes can wear different types of masks if the transmission occurs over different network layers. In particular, given $M\geq 1$ different types of masks, for an arbitrary node v in the population $\mathcal{N}$, we call it a type-ij node, if it wears a type-i mask when the transmission happens over layer-$\mathbb{C}$ ($1\leq i\leq M$); and it wears a type-j ($1\leq j\leq M$) mask when the transmission occurs over layer-$\mathbb{S}$. We then have M^2 node types in total. The node type distribution is given as $\boldsymbol{m}=\{m_{ij}\}, 1\leq i,j\leq M$. Individual node type is independently drawn from this distribution. We assume the node type is pre-assigned before the spreading process starts and does not depend on whether the node participates in layer-$\mathbb{S}$. We shall see that $p_i=\sum_{j=1}^{M}m_{ij}$. For integers $1\leq i,j,r,t\leq M$, the transmissibility from a type-ij infected node to a type-(r,t) susceptible node is thus given by:

$$(1-\epsilon_{\text{out},i})(1-\epsilon_{\text{in, r}})T_c$$

if the transmission occurs via a type-c edge, and

$$(1-\epsilon_{\text{out},j})(1-\epsilon_{\text{in, t}})T_s$$

If the transmission occurs via a type-s edge.

Even though there are M^2 node types, $M\times M$ (rather than $M^2\times M^2$) possible types of transmission exist on each layer. This is because the transmissibility for an infectious-susceptible node pair only depends on the mask types and the baseline transmissibility given the edge type that connects them, which are fully encoded in Equation (3) and (4). Therefore, we continue using the same $\mathbf{T_c}$ and $\mathbf{T_s}$ in Equation (3) and (4) to represent all possible types of transmission for each layer, but with extended semantics in usage. Specifically, for a type-ij infected node, it transmits the spreading item to a type-rt susceptible node with probability $\mathbf{T_c}[i,r]$ if the transmission occurs over a type-c edge; with probability $\mathbf{T_s}[j,t]$ if the transmission occurs over a type-c edge, where

$$\mathbf{T_c}[i,r]=\left(1-\epsilon_{\text{out},i}\right)\left(1-\epsilon_{\text{in},r}\right)T_c,\ 1\leq i,r\leq M \quad (5)$$

$$\mathbf{T_s}[j,t]=\left(1-\epsilon_{\text{out},j}\right)\left(1-\epsilon_{\text{in},t}\right)T_s,\ 1\leq j,t\leq M \quad (6)$$

Notice that layer-dependent population heterogeneity returns to layer-independent when $i=j$ and $r=t$ for all $1\leq i,j,r,t\leq M$.

III. Analytical Results

This section presents the derivation of the probability of emergence (PE), the epidemic threshold (R_0), and the expected epidemic size (ES). *Emergence* is defined as the event where the spreading process leads to a positive fraction of the infected population in the limit of the number of nodes n going to infinity. *Epidemics* refer to *large-scale* spreading events such as viral pandemics or information *memes*. Formally, with $S(n)$ denoting the *final* fraction of infected nodes in the population size of n, the probability of emergence with a random initiator is given by $\text{PE}=\lim_{n\to\infty}\mathbb{P}[S>0]$. Further, we are also interested in the epidemic threshold that separates the parameter space where $\lim_{n\to\infty}\mathbb{P}[S>0]=0$ from those that yield $\lim_{n\to\infty}\mathbb{P}[S>0]>0$. Finally, we compute the expected epidemic size when they take place, i.e., $\lim_{n\to\infty}\mathbb{E}[S|S>0]$.

A. Probability of Emergence and Epidemic Threshold

Consider random graphs $\mathbb{C}(n, \{p_k^c\})$ and $\mathbb{S}(n; \alpha, \{p_k^s\})$ as introduced in Section II-A. In order to study the viral transmission in the multi-layer network $\mathbb{H} = \mathbb{C} \coprod \mathbb{S}$, we consider a branching process that starts by giving the pathogen to an arbitrary node and the recursively expose the set of nodes that are reached and *infected* by exploring its neighbors. As mentioned in Section II-B, a type-ij infected node transmits the pathogen to a type-rt susceptible neighbor with probability $\mathbf{T_c}[i,r] = T_c(1-\epsilon_{i,out})(1-\epsilon_{r,in})$ if the link connecting them is type-c (or, with probability $\mathbf{T_s}[j,t] = T_s(1-\epsilon_{j,out})(1-\epsilon_{t,in}$ if the link between them is type-s), independently from all the other neighbors.

The *survival probability* of the aforementioned branching process is derived through a *mean-field* approach utilizing the method of generating functions [8], [25]. For integers $1 \leq i,j,r,t \leq M$, let $h_{c,ij}(x)$ (resp. $h_{s,ij}(x)$) denote the generating function for "the *finite* number of nodes reached and infected by following a *randomly* selected type-c (resp. type-s) edge coming from a type-ij infected node." Put differently, we have $h_{c,ij}(x) = \sum_{m=0}^{\infty} v_m x^m$ where v_m denotes the "probability that an arbitrary type-c edge coming from a type-ij infected node leads to a component of size m." Similarly, let $H_{ij}(x)$ denote the generating function for "the finite number of nodes reached and infected by following a randomly selected type-ij node."

Now we derive $h_{c,ij}(x)$ and $h_{s,ij}(x)$. For integers $1 \leq i,j,r,t \leq M$, we find that the following self-consistency equations hold:

$$h_{c,ij} = \sum_{r=1}^{M}\sum_{t=1}^{M} m_{rt}\Bigg(1 - \mathbf{T_c}[i,r] + \mathbf{T_c}[i,r]x\sum_{\boldsymbol{d}} \frac{p_{\boldsymbol{d}}k_c}{\langle k_c\rangle} h_{c,rt}(x)^{k_c-1}h_{s,rt}(x)^{k_s}\Bigg) \tag{7}$$

$$h_{s,ij} = \sum_{r=1}^{M}\sum_{t=1}^{M} m_{rt}\Bigg(1 - \mathbf{T_s}[j,t] + \mathbf{T_s}[j,t]x\sum_{\boldsymbol{d}} \frac{p_{\boldsymbol{d}}k_s}{\langle k_s\rangle} h_{c,rt}(x)^{k_c}h_{s,rt}(x)^{k_s-1}\Bigg) \tag{8}$$

We now explain each term in Eq. (7). Consider an infected type-ij node, say node v, and consider a type-c edge incident on it. Condition on the *type* of the node on the other end of this edge, say node u. Since the node type assignment is completed before the spread process and is drawn independently for all the nodes, the direct neighbor node u is of type-rt with probability m_{rt}. Conditioning on node u being type-rt, it will become infected through the type-c edge from v, with probability $\mathbf{T_c}[i,r]$. If the transmission fails with probability $1 - \mathbf{T_c}[i,r]$, then node v will have zero offspring through this edge to u, which explains the first part of Eq. (7). If the transmission is successful with probability $\mathbf{T_c}[i,r]$, the number of nodes reached and infected by node v increases by one (i.e., node u). This is captured by the multiplicative term x in the second half of Eq. (7). Additionally, the total size of this branch will also include all subsequent nodes that are reached and infected by u, which leads to the following term:

$$\sum_{\boldsymbol{d}} \frac{p_{\boldsymbol{d}}k_c}{\langle k_c\rangle} h_{c,rt}(x)^{k_c-1}h_{s,rt}(x)^{k_s}.$$

This term is explained as follows. Firstly, condition on the *colored degree* of node u, i.e., number of edges in both network layers. The term $p_{\boldsymbol{d}}k_c/\langle k_c\rangle$ gives the probability of colored *excess degree* of u [8]. It is the normalized probability that a type-c edge is attached to a node at the other end when the node has colored degree $\boldsymbol{d} = (k_c, k_s)$. Therefore, following the type-c edge from v that reaches u, u can infect other nodes with the remaining $k_c - 1$ type-c edges and k_s type-s edges. Recall that the number of nodes reached and infected by a type-rt node by following a type-c (resp. type-s) edge attached is generated by $h_{c,rt}$ (resp. $h_{s,rt}$). Collecting all the sub-branches, we obtain the term $h_{c,rt}(x)^{k_c-1}h_{s,rt}(x)^{k_s}$ utilizing the *powers property* of generating functions [25]. The validity of Equation (8) can be seen in a very similar way and is omitted here for brevity.

Utilizing equations (7) and (8), we now derive the generating function $H_{ij}(x)$ for the entire size of the branching process. For $1 \leq i,j \leq M$, we have

$$H_{ij}(x) = x\sum_{\boldsymbol{d}} p_{\boldsymbol{d}} h_{c,ij}(x)^{k_c}h_{s,ij}(x)^{k_s} \tag{9}$$

Here, the factor x corresponds to the initial node selected arbitrarily and infected. The selected node has colored degree $\boldsymbol{d} = (k_c, k_s)$ with probability $p_{\boldsymbol{d}}$. The number of nodes it reaches and infects by each of its k_c (resp. k_s) links of type-c (resp. type-s) is generated through $h_{c,ij}(x)$ (resp. $h_{s,ij}(x)$). Summing over all the possible colored degrees, we obtain Equation (9).

With equations (7)-(9) in hand, the generating function $H_{ij}(x)$ can be computed in the following manner. Given any x, we can solve for the recursive relations (7)-(8) to obtain $h_{c,ij}(x)$ and $h_{s,ij}(x)$ for integers $1 \leq i,j \leq M$, which in turn will yield H_{ij} for integers $1 \leq i,j \leq M$ in light of (9).

We are interested in cases where the number of nodes reached and infected by the initial node is *infinite*, representing cases where a randomly chosen infected node triggers an *epidemic*. The *conservation of probability* property of generating functions indicates that there exists a trivial fixed point $h_{c,ij}(1) = h_{s,ij}(1) = 1$ (yielding $H_{ij}(1) = 1$) when the number of nodes reached and infected is always *finite*. In other words, the underlying branching process is in the *subcritical* regime, and *all* infected components have finite size. However, the fixed point $h_{c,i}(1) = h_{s,i}(1) = 1$ may not be a stable solution to the recursion (7) to (9). We can check the stability of this fixed point by the linearization of recursion (7) to (9) around $h_{c,ij}(1) = h_{s,ij}(1) = 1$. This yields the Jacobian matrix $\boldsymbol{J}$ with the form $\boldsymbol{J} = \begin{bmatrix} \boldsymbol{J}_{cc} & \boldsymbol{J}_{cs} \\ \boldsymbol{J}_{sc} & \boldsymbol{J}_{ss} \end{bmatrix}_{2M^2\times 2M^2}$ in which

$$\begin{aligned} \boldsymbol{J}_{cc}(a,b) &= \frac{\partial h_{c,ij}(1)}{\partial h_{c,rt}(1)}; \quad & \boldsymbol{J}_{cs}(a,b) &= \frac{\partial h_{c,ij}(1)}{\partial h_{s,rt}(1)} \\ \boldsymbol{J}_{ss}(a,b) &= \frac{\partial h_{s,ij}(1)}{\partial h_{s,rt}(1)}; \quad & \boldsymbol{J}_{sc}(a,b) &= \frac{\partial h_{s,ij}(1)}{\partial h_{c,rt}(1)} \end{aligned} \tag{10}$$

where $a = \text{M2D}(i,j), b = \text{M2D}(r,t), 1 \leq i,j,r,t \leq M$. M2D is an M-base to 10-base converter, to map the tuple $ij, 1 \leq i,j \leq M$ to an integer ranging from 1 to M^2 for the ease of matrix indexing. Without loss of generality, we place the order of the four sub-matrices $\boldsymbol{J}_{cc}$ to $\boldsymbol{J}_{ss}$ as above. Swapping the locations of the sub-matrices will not change the spectral radius of $\boldsymbol{J}$ because two of the sub-matrices can always commute, and thus share the same polynomial characteristics based on Theorem 3 in Ref. [27]. We will use this property again to show PE and ES share the same transition points later in Sec. III-B.

If all eigenvalues of $\boldsymbol{J}$ are less than one in absolute value, i.e., if the spectral radius $\rho(\boldsymbol{J})$ of $\boldsymbol{J}$ satisfies $\rho(\boldsymbol{J}) \leq 1$, then the solution $h_{c,ij}(1) = h_{s,rt}(1) = 1$ is stable and $H_{ij}(1) = 1$ becomes the physical solution. In this case, the fraction of nodes that are infected will tend to zero as as the number of nodes n goes to infinity. In contrast, if $\rho(\boldsymbol{J}) > 1$, the trivial fixed point is not stable, which indicates that the branching process is in the *supercritical* regime; i.e., there is a positive probability that the branching process will lead to an *infinite* component. In this case, the fraction nodes that are infected will be strictly greater than zero as as the number of nodes n goes to infinity.

When $\rho(\boldsymbol{J}) > 1$, a nontrivial fixed point exists and becomes the attractor of the recursions (7) to (9), leading to a solution with $h_{c,ij}(1), h_{s,ij}(1) < 1$ which in turn yields $H_{ij}(1) < 1$. In that case, $1 - H_{ij}(1)$ gives the probability that the spreading process initiated by a seed node of type-ij yields an *epidemic*. Recall that S denotes the *final* fraction of infected nodes. The probability of epidemic emergence PE (with a random initiator) is thus given by $\text{PE} = \lim_{n\to\infty} \mathbb{P}[S > 0] = \sum_{i,j=1}^{M} m_{ij}(1 - H_{ij}(1))$. Finally, we conclude that the epidemic threshold, i.e., the boundary that separates the parameter regions where $\lim_{n\to\infty} \mathbb{P}[S > 0] = 0$ from those that yield $\lim_{n\to\infty} \mathbb{P}[S > 0] > 0$ is given by $\rho(\boldsymbol{J}) = 1$. $\rho(\boldsymbol{J})$ is also known as the *reproduction number*, R_0 in the disease spreading context. It is known that in the special case where $M = 1$, if R_0 is greater than one then the PE is positive, i.e., epidemics can take place. When $R_0 \leq 1$ on the other hand, the PE is zero [8] meaning that there is zero chance for a spreading process to reach a positive fraction of the population. Beyond marking a phase transition, the metric R_0 measures, in a sense, the speed at which the epidemic grows and is often used by policymakers when deciding on mitigation strategies. Thus, it is of significant importance to characterize R_0 for our model.

To further look into the implications of Eq. (10), utilizing Equation (7) to (8), we have

$$\begin{aligned} \boldsymbol{J}_{cc} &= \left[\lambda_{cc} \cdot \mathbf{T}_\mathbf{c}' \cdot \mathbf{m}\right]; \quad \boldsymbol{J}_{cs} = \left[\lambda_{cs} \cdot \mathbf{T}_\mathbf{c}' \cdot \mathbf{m}\right] \\ \boldsymbol{J}_{sc} &= \left[\lambda_{sc} \cdot \mathbf{T}_\mathbf{s}' \cdot \mathbf{m}\right]; \quad \boldsymbol{J}_{ss} = \left[\lambda_{ss} \cdot \mathbf{T}_\mathbf{s}' \cdot \mathbf{m}\right] \end{aligned}$$

$$\begin{aligned} \lambda_{cc} &= \frac{\langle k_c^2 \rangle - \langle k_c \rangle}{\langle k_c \rangle}, \quad \lambda_{cs} = \frac{\langle k_c k_s \rangle}{\langle k_c \rangle} \\ \lambda_{ss} &= \frac{\langle k_s^2 \rangle - \langle k_s \rangle}{\langle k_s \rangle}, \quad \lambda_{sc} = \frac{\langle k_c k_s \rangle}{\langle k_s \rangle} \end{aligned} \tag{11}$$

where $\mathbf{T_c}'[a,b] = \mathbf{T_c}[i,r]$, $\mathbf{T_s}'[a,b] = \mathbf{T_c}[j,t], \mathbf{m} = diag(\boldsymbol{m})$, where $\mathbf{m}[a,a] = \boldsymbol{m}[i,j], a = \text{M2D}(i,j), b = \text{M2D}(r,t), 1 \leq i,j,r,t \leq M$. $\mathbf{m}$ is a diagonal matrix of the node type distribution $\boldsymbol{m} = \{m_{ij}\}$. In matrix $\mathbf{T_c}'$ (resp. $\mathbf{T_s}'$), each element in $\mathbf{T_c}$ (resp. $\mathbf{T_s}$) will be at least duplicated for M^2 times. $\mathbf{T}_\mathbf{c}'$ and $\mathbf{T}_\mathbf{s}'$ are functions of the mask efficiencies and baseline transmissibilities, encoding all the possible transmission scenarios when taking into account layer-dependent population heterogeneity over on layer-$\mathbb{C}$ and $\mathbb{S}$, respectively. Note $\mathbf{T}_\mathbf{c}'$ and $\mathbf{T}_\mathbf{s}'$ are not symmetric unless $\mathbf{T_c}$ and $\mathbf{T_s}$ are symmetric, which only stands when $\mathbf{1} - \boldsymbol{\epsilon}_{\text{out}}$ and $\mathbf{1} - \boldsymbol{\epsilon}_{\text{in}}$ are colinear.

Observe that the four sub-matrices of $\boldsymbol{J}$ in Eq. (11) follow a same pattern, i.e., they are composed of three parts from left to right: i) a parameter that attributes the multi-layer network structure; ii) a heterogenous edge-type and node-type dependent transmissibility term; iii) a node type distribution term. This result disentangles the impact of multiple factors on the final state of the spreading process: the degree distribution of different layers and network structure of the multi-layer contact network as a whole, viral transmission dynamics, and layer-dependent population heterogeneity.

In a particular case where population heterogeneity is layer-independent, $\mathbf{T_c}$ and $\mathbf{T_s}$ are rank-1 and can be decomposed into $T_c \cdot \boldsymbol{\epsilon_{out}} \cdot \boldsymbol{\epsilon_{in}}^\top$ and $T_s \cdot \boldsymbol{\epsilon_{out}} \cdot \boldsymbol{\epsilon_{in}}^\top$, respectively. In other words, the node type here is not correlated with the network structure. $\boldsymbol{\epsilon_{out}} \cdot \boldsymbol{\epsilon_{in}}^\top$ is a property of each type of masks. The Jacobian matrix for the layer-independent population heterogeneity can be further simplified as follows:

$$\left[\begin{bmatrix} T_c & \\ & T_s \end{bmatrix} \cdot \begin{bmatrix} \lambda_{cc} & \lambda_{cs} \\ \lambda_{sc} & \lambda_{ss} \end{bmatrix}\right] \otimes \left((\mathbf{1} - \boldsymbol{\epsilon_{out}}) \cdot (\mathbf{1} - \boldsymbol{\epsilon_{in}}^\top) \cdot \mathbf{p}\right) \tag{12}$$

where $\mathbf{p} = diag(\boldsymbol{p})$ (recall $\boldsymbol{p}$ is mask type distribution), and $\otimes$ is the Kronecker product sign. The spectral radius $\rho(\boldsymbol{J})$ is thus

$$\rho\left(\begin{bmatrix} T_c & \\ & T_s \end{bmatrix} \cdot \begin{bmatrix} \lambda_{cc} & \lambda_{cs} \\ \lambda_{sc} & \lambda_{ss} \end{bmatrix}\right) \cdot \rho\left((\mathbf{1} - \boldsymbol{\epsilon_{out}}) \cdot (\mathbf{1} - \boldsymbol{\epsilon_{in}}^\top) \cdot \mathbf{p}\right) \tag{13}$$

Eq. (13) further disentangles between network-related and mask-related factors and shows the trade-off between them upon the critical behavior of the spreading process. From (13) it is straightforward to derive the spectrum radius for the *single* layer network that has degree distribution $\{p_k\}$ with mean degree $\langle k \rangle$ and baseline transmissibility T with population heterogeneity by the same set of masks, as shown in Eq. (14). The same expression is obtained in Ref. [15] for the single-layer contact network.

$$\rho(\boldsymbol{J}_{\text{single-layer}}) = \frac{\langle k^2 \rangle - \langle k \rangle}{\langle k \rangle} \cdot T \cdot \rho\left((\mathbf{1} - \boldsymbol{\epsilon_{out}}) \cdot (\mathbf{1} - \boldsymbol{\epsilon_{in}}^\top) \cdot \mathbf{p}\right) \tag{14}$$

B. Expected epidemic size

In this section, we compute the expected size of epidemics when they take place, i.e., $\lim_{n\to\infty} \mathbb{E}[S \mid S > 0]$. We will also compute the fraction of infected nodes in each type. Our approach is similar that used in [3], [4], [17], [22], [28]. Since the multi-layer network $\mathbb{H}$ is locally tree-like as the network size approaches infinity [29], we can consider it as a tree-structure, where there is a single node of type-ij at the top

level (referred to as the *root*). We label the levels of the tree from $\ell = 0$ at the bottom to the top $\ell = \infty$. Without loss of generality, we assume that the spreading event starts at the bottom of the tree and proceeds toward the top. In other words, we assume that a node at level ℓ can only be infected by one of its neighbors in level $\ell - 1$. Let $q_{c,\ell}^{ij}$ (respectively, $q_{s,\ell}^{ij}$) denote the probability of a type-ij node at level ℓ who is connected to its parent at level $\ell+1$ through a type-c (respectively, type-s) edge is *not* infected. Our essential goal is to compute q_{∞}^{ij} which represents the probability that the *root* node, which is of type-ij, is *not* infected. Given that the root node is arbitrary, q_{∞}^{ij} also gives the expected fraction of type-ij nodes that will be infected during the spreading process. Put differently, we have $\lim_{n\to\infty} \mathbb{E}[S_{ij} \mid S > 0] = 1 - q_{\infty}^{ij}$ with S_{ij} denoting the fraction of nodes of type-ij that are infected in the spreading process; we also have $S = \sum_{i,j=1}^{M} m_{ij} S_{ij}$.

Now we derive $q_{c,\ell}^{ij}$ and $q_{s,\ell}^{ij}$ in a recursive manner. For each $i, j = 1, \ldots, M$, we find that

$$q_{c,\ell+1}^{ij} = \sum_{\boldsymbol{d}=(k_c,k_s)} \frac{p_{\boldsymbol{d}} k_c}{\langle k_c \rangle} f_{ij}(\boldsymbol{q}_{c,\ell}, \boldsymbol{q}_{s,\ell}, k_c - 1, k_s) \tag{15}$$

$$q_{s,\ell+1}^{ij} = \sum_{\boldsymbol{d}=(k_c,k_s)} \frac{p_{\boldsymbol{d}} k_s}{\langle k_s \rangle} f_{ij}(\boldsymbol{q}_{c,\ell}, \boldsymbol{q}_{s,\ell}, k_c, k_s - 1) \tag{16}$$

where $\boldsymbol{q}_{c,\ell} = [q_{c,\ell}^{11}, q_{c,\ell}^{12}, ..., q_{c,\ell}^{MM}]$, $\boldsymbol{q}_{s,\ell} = [q_{s,\ell}^{11}, q_{s,\ell}^{12}, ..., q_{s,\ell}^{MM}]$, and $f_{ij}(\boldsymbol{q}_{c,\ell}, \boldsymbol{q}_{s,\ell}, k_c, k_s)$ is given by

$$\left(\sum_{r=1}^{M} \sum_{t=1}^{M} m_{rt} (1 - \mathbf{T_c}[r,i] + q_{c,\ell}^{rt} \mathbf{T_c}[r,i]) \right)^{k_c} \cdot \left(\sum_{r=1}^{M} \sum_{t=1}^{M} m_{rt} (1 - \mathbf{T_s}[t,j] + q_{s,\ell}^{rt} \mathbf{T_s}[t,j]) \right)^{k_s} \tag{17}$$

In order to see why (15) holds, let u be a type-ij node at level $\ell+1$ who is connected to its unique parent at level $\ell+2$ with an edge of type-c. As already mentioned, $q_{c,\ell+1}^{ij}$ gives the probability that u is *not* infected. As before, we first condition on the colored degree of u being $\boldsymbol{d} = (k_c, k_s)$ which has probability $\frac{p_{\boldsymbol{d}} k_c}{\langle k_c \rangle}$. Under the assumption that nodes can only be infected by neighbors in the layers below, node u can be infected through either one of $k_c - 1$ edges of type-c and k_s edges of type-s in layer ℓ (given that one of its type-c edges is used to connect it to the parent node in layer $\ell + 2$). We establish (15) by noting that $f_{ij}(\boldsymbol{q}_{c,\ell}, \boldsymbol{q}_{s,\ell}, k_c, k_s)$ represents the probability that a type-ij node with k_c edges of type-c and k_s edges of type-s with nodes in layer ℓ is *not* infected. Equation (16) can be seen to hold in a similar way.

We now explain why (17) holds. First, for a node with k_c edges of type-c and k_s edges of type-s in layer ℓ to be not infected, it should not receive the pathogen from any of these neighbors. Given the independence of infection events, we see that $f_{ij}(\boldsymbol{q}_{c,\ell}, \boldsymbol{q}_{s,\ell}, k_c, k_s)$ should be of the form $f_{ij}(\boldsymbol{q}_{c,\ell})^{k_c} f_{ij}(\boldsymbol{q}_{s,\ell})^{k_s}$ with $f_{ij}(\boldsymbol{q}_{c,\ell})$ (respectively, $f_{ij}(\boldsymbol{q}_{s,\ell})$) defined as the probability that a type-ij node with only one edge of type-c (respectively, type-s) with nodes in layer ℓ is *not* infected. In order to compute $f_{ij}(\boldsymbol{q}_{c,\ell})$, we condition on the type of the node that is connected in layer ℓ, which is type-rt with probability m_{rt}. Then, we note that for a type-ij node in layer $\ell + 1$ to be infected by a type-rt neighbor in layer ℓ that it is connected via a type-c link if both of the following events hold: the node in layer ℓ is infected, which happens with probability $(1 - q_{c,\ell}^{rt})$, *and* the pathogen is transmitted from the node in layer ℓ to its parent in layer $\ell + 1$, which happens with probability $\mathbf{T_c}[r,i]$. Collecting, we see that the probability of a type-ij node in layer $\ell+1$ to be *not* infected by a type-rt neighbor in layer ℓ that it is connected via a type-c is given by

$$1 - (1 - q_{c,\ell}^{rt}) \mathbf{T_c}[r,i] = 1 - \mathbf{T_c}[r,i] + q_{c,\ell}^{rt} \mathbf{T_c}[r,i]$$

Proceeding similarly for $f_{ij}(\boldsymbol{q}_{s,\ell})$, we establish (17).

We are now able to compute q_{∞}^{ij} for each $i, j = 1, \ldots, M$. First, solving (15)-(16) in the limit of $\ell \to \infty$ we compute $\boldsymbol{q}_{c,\infty}$ and $\boldsymbol{q}_{s,\infty}$. Using these, we then get

$$q_{\infty}^{ij} = \sum_{\boldsymbol{d}=(k_c,k_s)} p_{\boldsymbol{d}} f_{ij}(\boldsymbol{q}_{c,\infty}, \boldsymbol{q}_{s,\infty}, k_c, k_s) \tag{18}$$

by conditioning on the colored degree of the root node. Finally, we have $\lim_{n\to\infty} \mathbb{E}[S_{ij} \mid S > 0] = 1 - q_{\infty}^{ij}$ and the expected epidemic size is given by $\text{ES} = \lim_{n\to\infty} \mathbb{E}[S \mid S > 0] = \sum_{i,j=1}^{M} m_{ij}(1 - q_{\infty}^{ij})$.

Convergence and Phase transition of ES : Next, we show that PE and ES share the same transition point at the epidemic threshold $\rho(\boldsymbol{J}) = 1$. Note the convergence to the fixed point $\boldsymbol{q}_{c,\infty}$ and $\boldsymbol{q}_{s,\infty}$ is guaranteed. Namely, taking the limit $\ell \to \infty$ in Equation (17) is a well-defined operation since the multi-type branching process corresponding to Equation (15) and (16) is positive regular and non-singular under the assumption that $m_{ij} > 0$, $\mathbf{T_c}[i,j] > 0$, $\mathbf{T_s}[i,j] > 0$ for all integers $1 \leq i, j \leq M$. Then the convergence to the fixed point $\boldsymbol{q}_{c,\infty}$ and $\boldsymbol{q}_{s,\infty}$ is guaranteed as long as the initial condition $\boldsymbol{q}_{c,0}$ and $\boldsymbol{q}_{s,0}$ has positive entries [30, Theorem V.3.2]. Conditioned on the event where the epidemic emerges, this can be safely assumed.

Just as in the PE, there is a phase transition between an ES of zero ($\boldsymbol{q}_{c,\infty} = \mathbf{1}$, $\boldsymbol{q}_{s,\infty} = \mathbf{1}$) and a positive ES ($\boldsymbol{q}_{c,\infty} < \mathbf{1}$, $\boldsymbol{q}_{s,\infty} < \mathbf{1}$). Noting that the formulas for Equation (7) and (15) are identical, except that $\mathbf{T_c}[i,r]$ is replaced with $\mathbf{T_c}[r,i]$. Similar to Equation (8) and (16). Let $\boldsymbol{J'}$ denote the Jacobian matrix of relations (15) and (16) , we show the threshold for the ES is given by $\rho(\boldsymbol{J'}) = 1$ where $\boldsymbol{J'} = \begin{bmatrix} \boldsymbol{J'}_{cc} & \boldsymbol{J'}_{cs} \\ \boldsymbol{J'}_{sc} & \boldsymbol{J'}_{ss} \end{bmatrix}_{2M^2 \times 2M^2}$ with

$$\begin{aligned} \boldsymbol{J'}_{cc}(a,b) &= \frac{\partial q_{c,\infty}^{ij}(1)}{\partial q_{c,\infty}^{rt}(1)}; \quad \boldsymbol{J'}_{cs}(a,b) = \frac{\partial q_{c,\infty}^{ij}(1)}{\partial q_{s,\infty}^{rt}(1)} \\ \boldsymbol{J'}_{ss}(a,b) &= \frac{\partial q_{s,\infty}^{ij}(1)}{\partial q_{s,\infty}^{rt}(1)}; \quad \boldsymbol{J'}_{sc}(a,b) = \frac{\partial q_{s,\infty}^{ij}(1)}{\partial q_{c,\infty}^{rt}(1)} \end{aligned} \tag{19}$$

where $a = \text{M2D}(i,j), b = \text{M2D}(r,t); 1 \leq i, j, r, t \leq M$. More specifically,

$$\begin{aligned} \boldsymbol{J}'_{cc} &= \left[\lambda_{cc} \cdot \mathbf{T_c}'^{\top} \cdot \mathbf{m} \right]; \quad \boldsymbol{J}'_{cs} = \left[\lambda_{cs} \cdot \mathbf{T_c}'^{\top} \cdot \mathbf{m} \right] \\ \boldsymbol{J}'_{sc} &= \left[\lambda_{sc} \cdot \mathbf{T_s}'^{\top} \cdot \mathbf{m} \right]; \quad \boldsymbol{J}'_{ss} = \left[\lambda_{ss} \cdot \mathbf{T_s}'^{\top} \cdot \mathbf{m} \right] \end{aligned} \tag{20}$$

When $\rho(\boldsymbol{J'})$ is less than or equal to 1, $\boldsymbol{q}_{c,\infty} = \mathbf{1}$, $\boldsymbol{q}_{s,\infty} = \mathbf{1}$. When $\rho(\boldsymbol{J'})$ is greater than 1, $q_{c,\infty}^{ij}, q_{s,\infty}^{ij} < 1$ for all $1 \leq i, j \leq M$. It can be seen that $\boldsymbol{J'}$ shares the exactly same transition points for R_0 defined in (10) and (11).

We show this by proving $\boldsymbol{J}^\top$ and $\boldsymbol{J}'$ have the same spectrum, which in turn implies that $\rho(\boldsymbol{J}) = \rho(\boldsymbol{J'})$ because $\rho(\boldsymbol{J}) = \rho(\boldsymbol{J}^\top)$. Note that $\boldsymbol{J}^\top$ has the following form: $\boldsymbol{J}^\top = \begin{bmatrix} \boldsymbol{J}^\top{}_{cc} & \boldsymbol{J}^\top{}_{sc} \\ \boldsymbol{J}^\top{}_{cs} & \boldsymbol{J}^\top{}_{ss} \end{bmatrix}$ where

$$\begin{aligned} \boldsymbol{J}_{cc}^\top &= \left[\lambda_{cc} \cdot \mathbf{T_c}'^\top \cdot \mathbf{m}\right]; \quad \boldsymbol{J}_{cs}^\top = \left[\lambda_{cs} \cdot \mathbf{T_c}'^\top \cdot \mathbf{m}\right] \\ \boldsymbol{J}_{sc}^\top &= \left[\lambda_{sc} \cdot \mathbf{T_s}'^\top \cdot \mathbf{m}\right]; \quad \boldsymbol{J}_{ss}^\top = \left[\lambda_{ss} \cdot \mathbf{T_s}'^\top \cdot \mathbf{m}\right] \end{aligned} \tag{21}$$

The only difference between (21) and (20) is the anti-diagonal blocks (left-bottom to upper-right) swapping the positions. This does not change the spectrum radius because it is just a matter of how we place them, as it is in (10). To be more precise, this can also be seen through the stronger property that the characteristic polynomials of the two matrices are identical. Specifically, Based on Theorem 3 of the determinant of block matrix in Ref. [27], for block matrix (11), where $\boldsymbol{J}_{sc}\boldsymbol{J}_{ss} = \boldsymbol{J}_{ss}\boldsymbol{J}_{sc}$, we have $\det(\boldsymbol{J}) = \det(\boldsymbol{J}_{cc}\boldsymbol{J}_{ss} - \boldsymbol{J}_{cs}\boldsymbol{J}_{sc}) = \det(\boldsymbol{J}_{ss}^\top\boldsymbol{J}_{cc}^\top - \boldsymbol{J}_{sc}^\top\boldsymbol{J}_{cs}^\top)$. Similarly, due to $\boldsymbol{J}'_{cc}\boldsymbol{J}'_{cs} = \boldsymbol{J}'_{cs}\boldsymbol{J}'_{cc}$, $\det(\boldsymbol{J}') = \det(\boldsymbol{J}_{ss}^\top\boldsymbol{J}_{cc}^\top - \boldsymbol{J}_{sc}^\top\boldsymbol{J}_{cs}^\top)$. Putting together the results of this section and Section III-A, we have shown that when $R_0 \leq 1$, the epidemic dies out in finite time, whereas when $R_0 > 1$, the epidemic eventually infects a positive fraction of the population.

C. Opposite trends between PE and ES

Next, we discuss the behavior of PE and ES. We point out that PE and ES are not always showing the same trends, even though in most literature they do show the same monotonic trend [4], [6], [22].

As discussed, PE and ES share a common epidemic threshold $R_0 = 1$. Besides, Eq. (7) -(8) and (17) are identical, except that $\mathbf{T_c}[i, r]$ is replaced with $\mathbf{T_c}[r, i]$. However, this difference may lead to the opposite trends between PE and ES even though they have the same transition points. Fig. 1 provides evidence for it.

In this analysis, we use vectors $\boldsymbol{\epsilon}_{in}$ and $\boldsymbol{\epsilon}_{out}$ to represent the inward and outward efficiencies for all types of masks. When the inward efficiency of a mask is better than its outward efficiency, we call them inward-good masks. Similarly, we call masks with higher outward efficiency as outward-good masks. For simplicity, we only consider layer-independent population heterogeneity.

Assume that there are three types of masks among the population: inward-good mask wearers, outward-good mask wearers, and people who don't wear masks, represented as type-1, type-2, and type-3 nodes, respectively. We have the proportion vector of three types as $\boldsymbol{m} = [m_{\text{outward-good}}, m_{\text{inward-good}}, m_{\text{no-mask}}]$. We fix the proportion of no-mask-wearers at 0.1, and vary the proportion of outward-good-mask-wearers $m_{\text{outward-good}}$ from 0.1 to 0.9. Based on Eq. (13), to compare inward-good and outward-good masks fairly, the parameter choice for the inward-good mask and outward-good mask should follow that $(1-\epsilon_{\text{out},o})(1-\epsilon_{\text{in},o}) = (1-\epsilon_{\text{out},i})(1-\epsilon_{\text{in},i})$, where $\epsilon_{\text{out},o}$ and $\epsilon_{\text{in},o}$ ($\epsilon_{\text{out},i}$ and $\epsilon_{\text{in},i}$ resp.) represent the outward and inward efficiencies for the outward-good mask (inward-good mask resp.). The efficiency parameters of the masks are selected as $\boldsymbol{\epsilon}_{out} = [0.7, 0.4, 0]$, and $\boldsymbol{\epsilon}_{in} = [0.4, 0.7, 0]$.

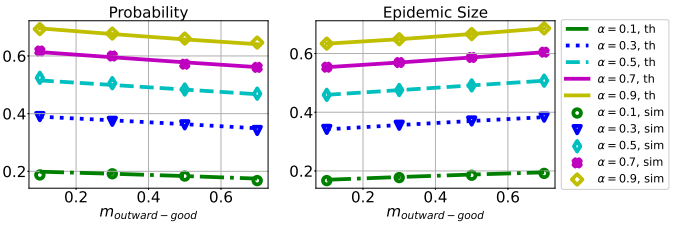


Fig. 1. *Probability of emergence (left) and expected epidemic size given emergence (right) show opposite trends as* $m_{\text{outward-good}}$ *increases.* $\boldsymbol{\epsilon}_{out} = [0.7, 0.4, 0]$, *and* $\boldsymbol{\epsilon}_{in} = [0.4, 0.7, 0]$. $m_{\text{no-mask}} = 0.1$. $T_c = 0.6, T_s = 0.5$. $\{p_k^c\} \sim$ *Poisson*(6), $\{p_k^c\} \sim$ *Poisson*(8).

Fig. 1 shows that as $m_{\text{outward-good}}$ increases, PE decreases, and ES increases, with different α values. This implies that a source-control-oriented strategy is crucial to prevent the epidemic from emerging over a multi-layer network at the early stages of the virus spreading, i.e., when the infection fraction has not reached a significant percentage. However, if an epidemic has already occurred and a significant fraction of the population has already been infected, it becomes most effective to implement a self-protection-oriented strategy to reduce the final fraction of the infected nodes. This trend is first reported in Ref. [15] over single-layer networks. Here we validated it over multi-layer networks under different second-layer participation rates.

IV. Model Validation

In this section, we present the simulation results in the finite nodes regime with an eye toward validating our analytical results. Note that the analytical results are exact in the limit of the number of nodes n goes to infinity. In doing so, we also aim to shed light on how the node type distribution affects the spreading process, e.g., in terms of probability and expected size of epidemics. Throughout, we fix the number of nodes as $n = 10^6$. We run 10,000 independent experiments for each parameter setting and report the average of these independent trials. The network $\mathbb{C}$ and $\mathbb{S}$ are generated based on configuration model using degree distributions with finite moments. We validate our results using two types of degree distributions: Poisson and power law with exponential cutoff. Poisson distribution is simple and it is one of the most widely-used and analyzed degree distributions to generate random networks. We choose power law with exponential cutoff distribution too because they are applied to a wide range of real-world networks [4], [31], and they are *well-behaved*, i.e., have finite moments. The validation is under the layer-dependent population heterogeneity setup. Ref. [22] provides more validations with layer-independent population heterogeneity.

The fraction of nodes that participate in network layer $\mathbb{S}$ is denoted by α. The baseline transmissibilities are denoted by T_c and T_s for networks $\mathbb{C}$ and $\mathbb{S}$, respectively. In this experiment, we assume there are two types of masks: surgical mask (type-1) and cloth-mask (type-2). The outward and inward efficiencies are denoted by $\boldsymbol{\epsilon}_{\text{out}} = [\epsilon_{\text{out},1}, \epsilon_{\text{out},2}]$, and

$\boldsymbol{\epsilon}_{\text{in}} = [\epsilon_{\text{in},1}, \epsilon_{\text{in},2}]$. The entire population is thus split into four non-overlapping categories: type-1 on $\mathbb{C}$ and type-1 on $\mathbb{S}$, type-2 on $\mathbb{C}$ and type-1 on on $\mathbb{S}$, type-1 on $\mathbb{C}$ and type-2 on $\mathbb{S}$, and type-2 on $\mathbb{C}$ and type-2 on $\mathbb{S}$. $\boldsymbol{m} = [m_{11}, m_{21}, m_{12}, m_{22}]$ denotes the proportion of the above 4 types of nodes. Here we fix $m_{11} = 0.1$, and vary m_{21} and m_{12}, and we have $m_{22} = 0.9 - m_{21} - m_{12}$.

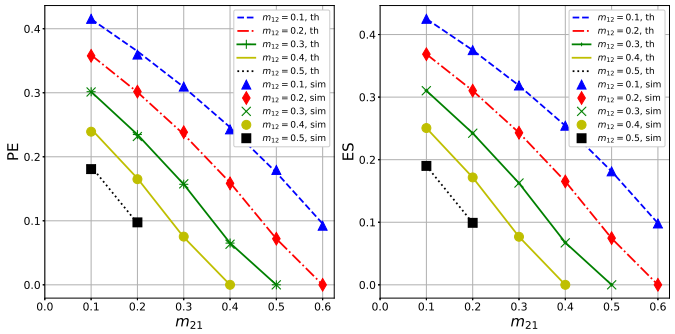


Fig. 2. *Probability of emergence (left) and epidemic size given emergence (right) varying m_{21} for different m_{12} values. We set $m_{11} = 0.1$ and $m_{22} = 0.9 - m_{21} - m_{12}$. $\{p_k^c\} \sim$ Poisson(6), $\{p_k^c\} \sim$ Poisson(8) and $\alpha = 0.6$. $T_c = 0.6$, $T_s = 0.5$, $\boldsymbol{\epsilon}_{out} = [0.8, 0.5]$, and $\boldsymbol{\epsilon}_{in} = [0.7, 0.5]$. Node set size $n = 1,000,000$, and each data point of the simulation result is averaged over 10,000 trials. Analytical results (marked th) show a near-perfect match with the simulation results (marked sim).*

In Fig. 2, we investigate how the probability of emergence (PE) and expected epidemic size given emergence (ES) change as we vary m_{21} and m_{12} when the degree distributions follow Poisson degree distributions. We see that the simulation results match the analytical solutions with near-perfect accuracy, confirming our analytical results' usefulness in the finite node regime. Moreover, the results are also helpful in understanding the impact of node-type distributions. Is is seen that with fixed m_{11} and m_{12}, as m_{21} increases (m_{22} decreases), PE and ES decrease. Namely, while the fraction of nodes wearing cloth masks on layer-$\mathbb{C}$ and surgical masks on layer-$\mathbb{S}$ increases, PE and ES decrease. This shows that masks with better efficiencies on layer-$\mathbb{S}$ help in reducing the risk and size (if it already exists) of an epidemic. On the other hand, with fixed m_{11} and m_{21}, increasing m_{12} (decreasing m_{22}) decreases PE and ES. Similarly, this indicates that better masks on layer-$\mathbb{C}$ are also helpful. In Sec. VI, we further investigate the interplay between the node-type distribution and multi-layer network structure.

Fig. 3 reports the results when generating the networks using power law with exponential cutoff. We see a near-perfect match with a different type of degree distribution. Similar trends in terms of varying m_{21} and m_{12} are observed in Fig. 3. These demonstrate the validation of our analytical results. Besides, we see that PE and ES in Fig. 3 are lower than those in Fig. 2. It is due to the fact that the mean degrees of the networks generated by the chosen parameter for power law with exponential cutoff being lower. This also helps understand the impact of strategies that reduce the mean degrees of contact networks in disease spread, such as social distancing and quarantines.

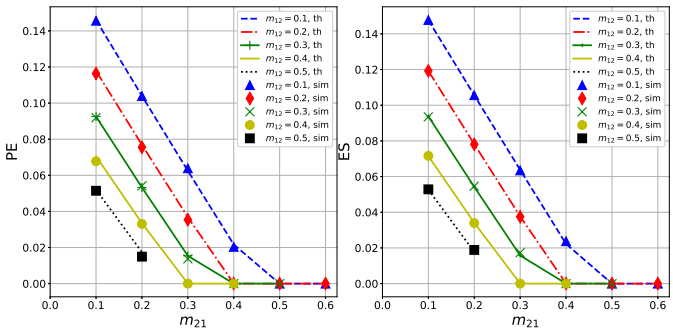


Fig. 3. *Probability of emergence (left) and epidemic size given emergence (right) varying m_{21} for different m_{12} values. We set $m_{11} = 0.1$ and $m_{22} = 0.9 - m_{21} - m_{12}$. $\{p_k^c\}$ and $\{p_k^s\}$ follow power law with exponential cutoff where the power exponent equals 2.5, and the cutoff equals 10. The mean degree is 1.028. $\alpha = 0.6$. $T_c = 0.4$, $T_s = 0.7$, $\boldsymbol{\epsilon}_{out} = [0.5, 0]$, and $\boldsymbol{\epsilon}_{in} = [0.4, 0]$. Node set size $n = 1,000,000$, and each data point of the simulation result is averaged over 10,000 trials. Analytical results (marked th) show a near-perfect match with the simulation results (marked sim).*

V. Comparison between multi-layer and single-layer networks

This section compares the dynamics of spreading processes over a single-layer network with that over a multi-layer network. In other words, we consider whether projecting a multi-layer network into a single-layer network leads to any significant differences in the dynamics that warrant a separate analysis of the multi-layer network structure. Poisson degree distribution is used for simplicity. We show that there exists such significant difference, and it is necessary to model a multi-layer network while having a single-layer network solution.

As mentioned in Sec. II-A, we generate layer-$\mathbb{S}$ and layer-$\mathbb{C}$ separately based on the configuration model, given their respective degree distributions $\{p_k^s\}$ and $\{p_k^c\}$. This results in the *colored* degree distribution as shown in Eq. (2). One approach of projecting the multi-layer network degree distribution to a single-layer network degree distribution is to *ignore* the color of the edges and match all stubs randomly with each other. Specifically, based on bond percolation theory [8], for the multi-layer network with mean degrees λ_c and λ_s and baseline transmissibilities T_c and T_s, it is equivalent to be considered as a multi-layer network with mean degrees $\lambda_c T_c$ and $\lambda_s T_s$ with transmissibilities 1. Given these, for a node in the multi-layer network, its degree in the projected single-layer network is the sum of the node's colored degrees in the multi-layer network. Formally, for an arbitrary node i in the multi-layer network $\mathbb{H}$ where the degree distribution follows (2), the projected single-layer network degree disitrution has the following form:

$$p_{d_{\text{SL}}^i} = \sum_{k_c^i + k_s^i = d_{\text{SL}}^i} p_{\boldsymbol{d}^i = (k_c^i, k_s^i)} \tag{22}$$

Assuming there are two types of masks in the population for convenience, we consider layer-independent population heterogeneity for the multi-layer network model for the single-layer projection. Given the same mask efficiencies and distribution, we compare PE and ES for multi-layer (abbrev. ML)

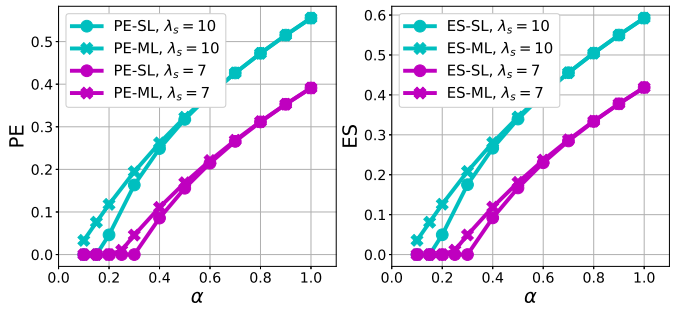


Fig. 4. *Comparison of PE and ES for the multi-layer network model (abbrev. ML) and the single-layer projection (abbrev. SL) by degree projection.* $\lambda_c = 7$, $T_c = 0.6$ *and* $T_s = 0.5$. *Suppose there are only two types of masks with* $\boldsymbol{\epsilon}_{out} = [0.8, 0.5]$, $\boldsymbol{\epsilon}_{in} = [0.7, 0.5]$. *For the purpose of projection, we consider layer-independent population heterogeneity with mask type distribution* $\boldsymbol{m}_{ML} = \boldsymbol{m}_{SL} = [0.4, 0.6]$.

and projected single-layer (abbrev. SL) models. Fig. 4 shows the results when $\lambda_s = 7$ and 10 with fixed $\lambda_c = 7$ while varying α. We see that PE-SL and ES-SL do not match the corresponding ML results. First, the SL results are not able to accurately predict the epidemic threshold where $\rho(\boldsymbol{J}) = 1$, which postpones the transition points. Second, there is a discrepancy between ML and SL results. It is observed that when both PE-ML and PE-SL (ES-ML and ES-SL follow the same trend) are non-zero, increasing α decreases this discrepancy. Meanwhile, the discrepancy increases under the same α value when λ_s increases from 7 to 10.

We now explain why there exists a discrepancy between the predictions and the trend when varying α. The key difference between the ML and SL network models is the resulting assortativity. When α is small, for example, $\alpha = 0.1$, only 10% of the nodes have type-c edges, and they can only be randomly matched with other type-c edges in the ML case. In this case, a small fraction of the population will have statistically *higher* degrees than the rest, and the additional links they have only connect nodes with high degrees together. This results in a positive correlation, i.e., assortativity, between the degrees of pairs of connected nodes. High-degree nodes provide a higher probability of transmitting the spreading item to the rest compared to the low-degree nodes, and higher assortativity will amplify such an effect. This is also the reason that transition points of SL come later than ML as α increases, and the ML predictions for PE and ES are above the SL ones. Besides, with the same α, when λ_s is larger, the assortativity creates a more significant difference.

We further consider the case where the single-layer and multi-layer models match the epidemic threshold by utilizing Eq. (13) and (14). Both Eq. (13) and (14) disentangle three impact factors on the critical transition behavior of a spreading process: a term that attributes the network structure, a term incorporating baseline transmissibilities, and a population heterogeneity term. This provides another method of projection that keeps the same R_0 for ML and SL: given the same the population heterogeneity term, and let

$$\frac{\langle k^2\rangle - \langle k\rangle}{\langle k\rangle} \cdot T = \rho\left(\begin{bmatrix} T_c & \\ & T_s \end{bmatrix} \cdot \begin{bmatrix} \lambda_{cc} & \lambda_{cs} \\ \lambda_{sc} & \lambda_{ss} \end{bmatrix}\right).$$

Fig. 5 shows the results of the second projection method by matching R_0 with varying α. We can see that, matching R_0, still, SL can not accurately predict PE and ES for ML. Similar to Fig. 4, when α increases, the difference between the predictions of SL and ML decreases in most cases. Moreover, we see that SL-PE and SL-ES provides an upper-bound for the corresponding ML-PE and ML-ES. This is due to the fact that we match R_0 for SL and ML, and thus the SL model can at least generate the same amount of infected off-springs on average in the process of spreading. This indicates that given the same population heterogeneity set up, we can use the SL model to obtain PE and ES that upper-bound the ML results with less computation.

Collect, we find that the projected single-layer network is not able to accurately predict PE and ES for the corresponding multi-layer network when ignoring the edge type and aggregating the layers, and when matching the R_0. Thus the single-layer projection of a multi-layer network can not capture the essential properties (e.g. PE and ES) of a spreading process. However, PE and ES of the single-layer network model may provide bounds for the multi-layer network: the degree projection of the multi-layer network may provide a lower bound for the considered multi-layer network; while if match R_0, the single-layer network model will provide an upper-bound for the multi-layer network model.

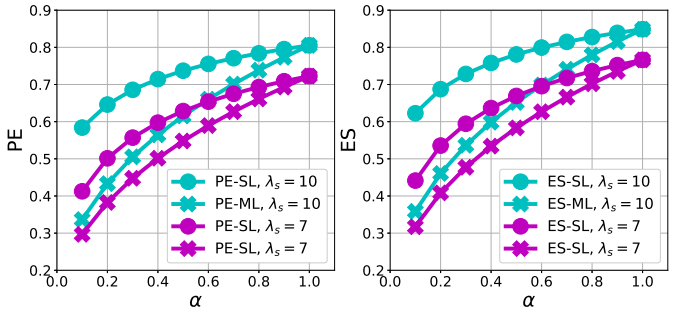


Fig. 5. *Comparison of PE and ES for the multi-layer network model (abbrev. ML) and the single-layer projection (abbrev. SL) by matching* R_0.

VI. Layer-Oriented Spreading Control with Limited Cost

Utilizing the layer-dependent population heterogeneity model, we develop and discuss three layer-oriented mask allocation strategies given limited expected costs of masks. We name the three strategies Policy 1, Policy 2, and Policy 3, respectively. As mentioned, the multi-layer contact network $\mathbb{H} = \mathbb{C} \coprod \mathbb{S}$ is composed of the first layer $\mathbb{C}$ and the second layer $\mathbb{S}$, where nodes in the layer-$\mathbb{C}$ participate in layer-$\mathbb{S}$ with probability α. In this analysis, we assume there are two types of masks: mask and no-mask, labeled as mask type-1 and mask type-2. The outward and inward efficiencies are denoted by $\boldsymbol{\epsilon}_{\text{out}} = [\epsilon_{\text{out},1}, \epsilon_{\text{out},2}]$, and $\boldsymbol{\epsilon}_{\text{in}} = [\epsilon_{\text{in},1}, \epsilon_{\text{in},2}]$, where $0 < \epsilon_{\text{out},1}, \epsilon_{\text{in},1} \leq 1$ and $\epsilon_{\text{out},2} = \epsilon_{\text{in},2} = 0$. Similar to Sec. IV, the entire population is split into four non-overlapping categories: masked on $\mathbb{C}$ and masked on $\mathbb{S}$, not-masked on $\mathbb{C}$ and masked on on $\mathbb{S}$, masked on $\mathbb{C}$ and not-masked on $\mathbb{S}$, and

not-masked on $\mathbb{C}$ and not-masked on $\mathbb{S}$. Policy 1 provides a baseline policy where each node in $\mathcal{N}$ has consistent mask-wearing behavior across the two layers of the contact network. Policy 2 explores a case where nodes do not wear masks over the layer-$\mathbb{S}$. Policy 3 requires masks to be mandated over layer-$\mathbb{S}$. More specifically, let $\boldsymbol{m}^p = [m^p_{11}, m^p_{21}, m^p_{12}, m^p_{22}]$ denote the proportion of the above 4 categories for Policy p, where $p = 1, 2, 3$. Policy 1 has $\boldsymbol{m}^1 = [m^1_{11}, 0, 0, m^1_{22}]$ where $m^1_{11} + m^1_{22} = 1$; Policy 2 has $\boldsymbol{m}^2 = [0, m^2_{12}, 0, m^2_{22}]$, $m^2_{12} + m^2_{22} = 1$; Policy 3 has $\boldsymbol{m}^3 = [m^3_{11}, 0, m^3_{21}, 0]$ and $m^3_{11} + m^3_{21} = 1$.

To quantify the expected total cost of policy p spending on masks for a fair comparison among the three policies, we propose the below formula:

$$\mathbb{E}[\text{cost}^p] = \sum_{i,j=1}^{M} \mathbb{E}[\text{cost}_{ij}] \cdot m^p_{ij} \tag{23}$$

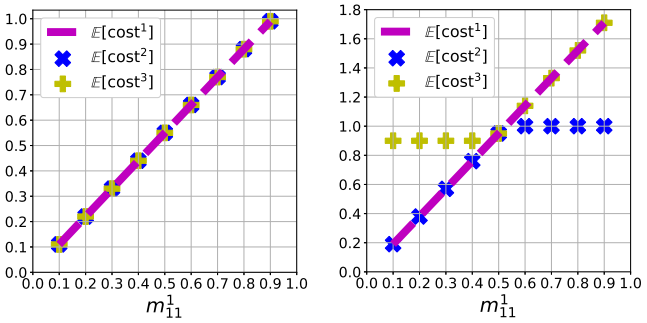
where $i, j = 1, ..., M$ and $p = 1, 2, 3$. $\mathbb{E}[\text{cost}_{ij}]$ denotes the expected cost for a type-ij individual, which is independent of the spreading process and the policies. In the considered case where $M = 2$, we have $\mathbb{E}[\text{cost}_{11}] = 1+\alpha$, $\mathbb{E}[\text{cost}_{12}] = 1$, $\mathbb{E}[\text{cost}_{21}] = \alpha$, and $\mathbb{E}[\text{cost}_{22}] = 0$.

With these in hand, we investigate two problems in turn: (A) Given a budget of C for the cost of masks, i.e., let $\mathbb{E}[\text{cost}^1] = \mathbb{E}[\text{cost}^2] = \mathbb{E}[\text{cost}^3] = C$, which policy yields the *best* mitigation effect (i.e., the lowest PE of an epidemic)? (B) To prevent the epidemic from happening, which policy yields the *lowest* expected cost?

A. PE with same expected cost

Next, we compare the effectiveness of spread control for policies (measured by PE), given the expected costs. In what follows, for simplicity, we generate the networks using Poisson degree distributions. Fig. 6 shows the results of matching the expected costs for Policy 1, 2 and 3 when increasing m^1_{11} from 0.1 to 0.9 when $\alpha = 0.1$ and $\alpha = 0.9$. Given m^1_{11} (x-axis), we have $\boldsymbol{m}^1 = [m^1_{11}, 1 - m^1_{11}, 0, 0]$. $C = \mathbb{E}[\text{cost}^1]$ can be obtained via (23). With $\mathbb{E}[\text{cost}^2] = \mathbb{E}[\text{cost}^3] = C$, reversely, we can solve for m^2_{12} and m^3_{11}. Note it is not guaranteed to have valid solutions ranging $[0, 1]$ for m^2_{12} and m^3_{11}. If the solutions have negative values, we replace them with 0s, and similarly, if they are larger than 1, we replace them with 1s. As shown in Fig. 6(a), when $\alpha = 0.1$, given m^1_{11} and thus C, we can match $\mathbb{E}[\text{cost}^2] = \mathbb{E}[\text{cost}^3] = C$ varying m^1_{11}. However, in Fig. 6(b), when $\alpha = 0.9$, $\mathbb{E}[\text{cost}^3]$ does not match C if $m^1_{11} < 0.5$ because it is lower-bounded by α. Similarly, when $m^1_{11} > 0.5$, $\mathbb{E}[\text{cost}^2]$ reaches the upper bound 1 and can not match C. Therefore we can see (23) is able to characterize and distinguish different policies.

Now we compare PE^1, PE^2, and PE^3 considering the different matching conditions for the expected costs as shown in Fig. 6(a) and 6(b). To explore the effectiveness of layer-dependent mask assignment policies, we further take into account the *transmission power* of each layer. Inspired by Eq. (14), we roughly estimate the *transmission power* of each layer by the product of the first moment of the degree distribution that generates the contact network layer and the baseline transmissibility of that layer, With Poisson(λ_c) generating layer-$\mathbb{C}$ and Poisson(λ_s) generating layer-$\mathbb{S}$ in this analysis, we consider three cases: (i) $T_c\lambda_c = T_s\lambda_s$; (ii) $T_c\lambda_c = 10T_s\lambda_s$; (iii) $10T_c\lambda_c = T_s\lambda_s$. The results are shown in Fig. 7.



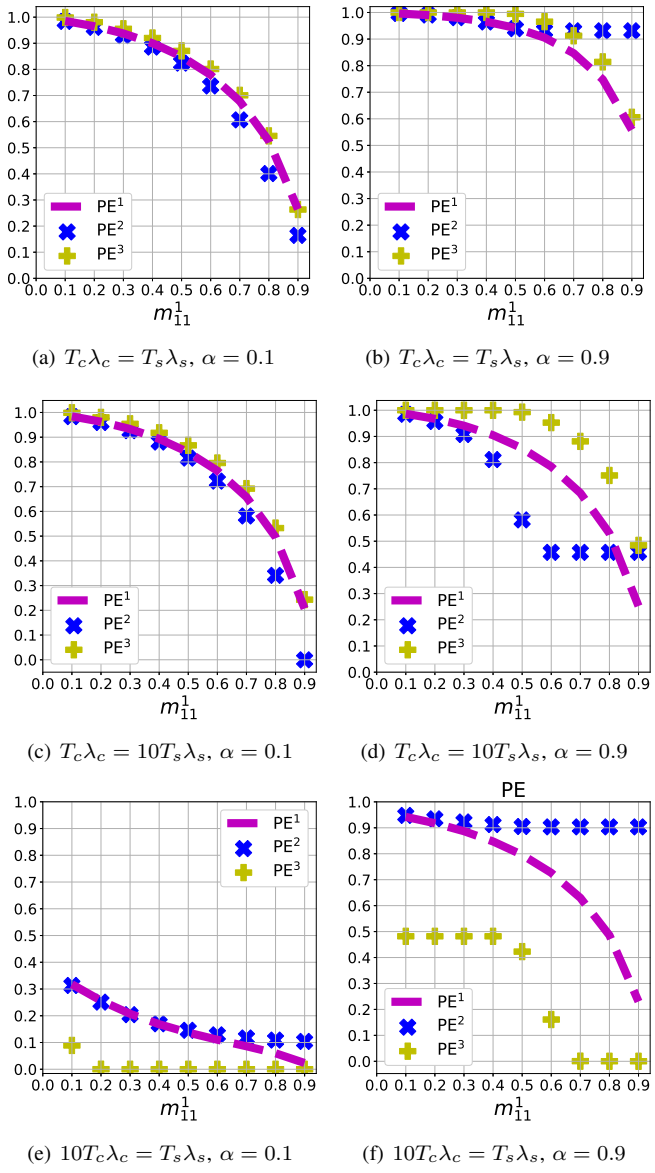
(a) $\alpha = 0.1$ (b) $\alpha = 0.9$

Fig. 6. $\mathbb{E}[cost^1]$ *(dashed line),* $\mathbb{E}[cost^2]$ *(cross) and* $\mathbb{E}[cost^3]$ *(plus) with varying* m^1_{11} *when* $\alpha = 0.1$ *(a), and* $\alpha = 0.9$ *(b).*

In Fig. 7, we can see that the transmission power is a significant factor in determining the best control policy. It is more effective for a layer-oriented spread control policy to direct more masks to the layer that has higher transmission power. When the transmission power of layer-$\mathbb{C}$ is larger than layer-$\mathbb{S}$, as shown in Fig. 7(c) and 7(d), Policy 2 that directs as many masks to layer-$\mathbb{C}$ given C has the lowest PE for most cases. Only at $m^1_{11} = 0.9$ when $\alpha = 0.9$, $\text{PE}^1 < \text{PE}^2$ due to $\mathbb{E}[\text{cost}^2]$ has reached upper bound of 1. When $0.5 < m^1_{11} < 0.9$ in Fig. 7(d), even reached the upper bound, with a lower expected cost than Policy 1 and 3, Policy 2 yields the lowest PE among all. Similarly, in Fig. 7(e) and 7(f), when layer-$\mathbb{S}$ has a higher transmission power, Policy 3 that mandates masks on layer-$\mathbb{S}$ provides the best mitigation effect. When layer-$\mathbb{C}$ and layer-$\mathbb{S}$ share the same transmission power, as shown in Fig.7(a) and Fig.7(b), there is no single policy that outperforms others as α increases. In the specific case of Fig. 7(a), Policy 2 shows slightly lower PE than the other two policies potentially due to the small participation rate of layer-$\mathbb{S}$ and equal transmission power of the two layers. Therefore the contribution to the spreading process of layer-$\mathbb{S}$ is not as significant as layer-$\mathbb{C}$, which can also be seen from the comparison with Fig. 7(e). However, in Fig. 7(b) where $\alpha = 0.9$, when the two layers are more coupled, due to the same transmission power of the two layers, the baseline Policy 1, which favors neither layer, yields the lowest PE. Moreover, Fig. 7 shows that, when the transmission power of layers are different, increasing α increases the discrepancy of PE between the best policy and others. This indicates that given budget, when layer-$\mathbb{S}$ has a higher participation rate, it is more necessary to consider layer-oriented spreading control based on the transmission power of layers.

B. Expected cost on Epidemic Boundaries

Next, we discuss a different problem from that in Sec. VI-A, where we are trying to find the optimal resource allocation strategy with a given budget. In this problem, our primary goal



(a) $T_c\lambda_c = T_s\lambda_s$, $\alpha = 0.1$ (b) $T_c\lambda_c = T_s\lambda_s$, $\alpha = 0.9$

(c) $T_c\lambda_c = 10T_s\lambda_s$, $\alpha = 0.1$ (d) $T_c\lambda_c = 10T_s\lambda_s$, $\alpha = 0.9$

(e) $10T_c\lambda_c = T_s\lambda_s$, $\alpha = 0.1$ (f) $10T_c\lambda_c = T_s\lambda_s$, $\alpha = 0.9$

Fig. 7. *PE*^{1} *(dashed line), PE*^{2} *(cross) and PE*^{3} *(plus) with varying* m_{11}^1. *Throughout* $T_c = T_s = 0.5$, *and* $\boldsymbol{\epsilon}_{out} = \boldsymbol{\epsilon}_{in} = [0.8, 0]$. *For (a) and (b),* $\lambda_c = \lambda_s = 20$. *For (c) and (d),* $\lambda_c = 20, \lambda_s = 2$. *For (e) and (f),* $\lambda_c = 2, \lambda_s = 20$.

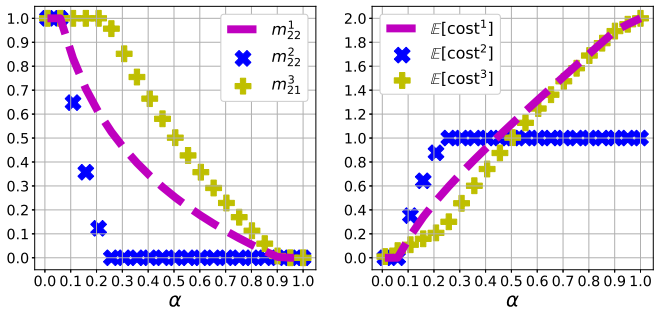
is to find a policy that prevents the epidemic from happening at the lowest cost. In particular, we need to find the expected costs for each policy at the corresponding *epidemic boundary* in the parameter space identified by the epidemic threshold $\rho(\boldsymbol{J}) = 1$, meaning that epidemics will *not* take place with probability one. Fig. 8 shows the boundary of $m - \alpha$ plane for Policy 1, 2 and 3 when the transmission power of layer-$\mathbb{C}$ is equal to, greater than and less than that of layer-$\mathbb{S}$.

On the LHS of each panel in Fig. 8, for each policy $p, p = 1, 2, 3$, the curves separate the areas where epidemics can take place (north-east of the curves) from the areas where they can not (south-west of the curves). Here we explore the trade-off between α and m_{ij}^p on the epidemic boundary where $\rho(\boldsymbol{J}) = 1$ while fixing other parameters. For clarity, we are only showing $\alpha - m_{22}^1$ plane for Policy 1, $\alpha - m_{22}^2$ plane for Policy 2, and $\alpha - m_{21}^3$ plane for Policy 3. Equivalently, the y-axis of LHS panels can also be viewed as the maximum fraction of node types that have less expected cost for a policy. For example, in the $\alpha - m_{22}^1$ plane in Fig. 8(a), for Policy 1, consider two types of nodes type-11 and type-22 where $m_{11}^1 + m_{22}^1 = 1$, and $\mathbb{E}[\text{cost}_{11}] = 0 > \mathbb{E}[\text{cost}_{22}] = 1 + \alpha$. As α increases, the maximum fraction of nodes that can wear no mask on both layers decreases to prevent the epidemic from happening. In fact, in all $\alpha - m_{ij}^p$ planes in Fig. 8, regardless of the transmission power of the layers, increasing α decreases the maximum fraction of less-cost node types for a policy that is needed for epidemics to happen.

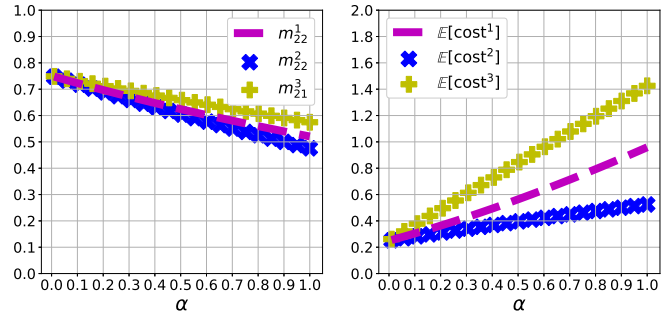
In Fig. 8(c), when the transmission power of the secondary layer $\mathbb{S}$ is higher, it is also more *economical* (i.e., lower expected cost) to control layer-$\mathbb{S}$. However, in Fig. 8(b), when the transmission power of layer-$\mathbb{C}$ is higher, controlling layer-$\mathbb{S}$ takes the least cost. These two trends echo our conclusions in Sec. VI-A that given a total cost, it is more effective to assign more masks to the layer that has higher transmission power. Most interestingly, in Fig. 8(a), when the transmission power of the two layers is the same, it is most economical to prioritize controlling the spreading on the layer-$\mathbb{S}$ to prevent the epidemic, as shown by results of Policy 3. First, on the LHS, Policy 3 allows more fractions of less-cost node types in the population. Meanwhile, on the RHS, Policy 3 also spends the least expected costs compared to other policies. This is potentially due to the fact that the second layer provides extra transmission pathways and thus increases the connectivity of the contact network, which makes it critical to block the transmission over the second layer to prevent the epidemic from happening.

VII. Conclusions

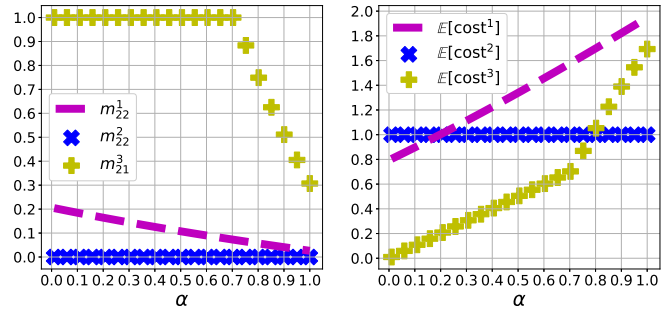
In this work, we provide a comprehensive analysis of the spreading process in multi-layer networks with layer-dependent population heterogeneity. We present analytical solutions for three key epidemiological measures: probability of emergence (PE), epidemic threshold, and expected epidemic size (ES). Our solutions disentangle the impact of multi-layer network structure, transmission dynamics, and population heterogeneity distribution on the final state of the spreading process. We validate our analytical results with extensive simulations. Comparing multi-layer networks to their single-layer projections, we find significant differences in the dynamics, emphasizing the need for separate analysis of multi-layer structures. Additionally, different projection methods may lead to bounds on key epidemiological quantities. We then explore layer-dependent population heterogeneity by studying three layer-oriented mitigation control policies. We propose a metric to quantify the expected cost of mask allocation, which helps characterize different policies. We identify the transmission power of each layer and the participation rate of nodes in the secondary layer as crucial factors for developing effective and economical mitigation strategies. Our findings provide insights into the spreading process in multi-layer complex networks with population heterogeneity. They can help the development of mitigation and control strategies for disease spread and information diffusion.



(a) $T_c\lambda_c = T_s\lambda_s$



(b) $T_c\lambda_c = 10T_s\lambda_s$



(c) $10T_c\lambda_c = T_s\lambda_s$

Fig. 8. *Epidemic boundaries of $\alpha - m_{ij}^p$ planes (LHS) and the corresponding $\mathbb{E}[cost^p]$ (RHS) on the epidemic boundaries. Policy 1, 2 and 3 correspond to dashed lines, cross and plus, respectively. Each curve in $\alpha - m_{ij}^p$ planes is the epidemic boundary identified by $\rho(\boldsymbol{J}) = 1$ for a policy when considering the trade-off between α and the fraction of the nodes whose node types have less cost for that policy. The north and east of each curve in $\alpha - m_{ij}^p$ planes specify the region for which epidemics are possible, while the south and west parts of each curve stand for the region where epidemics can not occur. Each policy's corresponding expected costs on the epidemic boundary are provided on its RHS. Throughout $\boldsymbol{\epsilon}_{\text{out}} = \boldsymbol{\epsilon}_{in} = [0.8, 0]$. For (a), $T_c = 0.2, T_s = 1$, $\lambda_c = 4, \lambda_s = 0.8$. For (b), $T_c = 0.05, T_s = 0.5, \lambda_c = 22, \lambda_s = 0.22$. For (c), $T_c = 0.2, T_s = 1, \lambda_c = 0.7, \lambda_s = 1.4$.*

REFERENCES

[1] M. Tambuscio, G. Ruffo, A. Flammini, and F. Menczer, "Fact-checking Effect on Viral Hoaxes: A Model of Misinformation Spread in Social Networks," in *Proceedings of the 24th International Conference on World Wide Web*, WWW '15 Companion, 2015.

[2] H. W. Hethcote, "The Mathematics of Infectious Diseases," *SIAM Review*, vol. 42, no. 4, 2000. Publisher: Society for Industrial and Applied Mathematics.

[3] J. P. Gleeson and D. J. Cahalane, "Seed size strongly affects cascades on random networks," *Physical Review E*, vol. 75, no. 5, 2007.

[4] M. Newman, "Epidemics on networks," in *Networks* (M. Newman, ed.), 2018.

[5] M. Li, X. Wang, K. Gao, and S. Zhang, "A Survey on Information Diffusion in Online Social Networks: Models and Methods," *Information*, vol. 8, no. 4, 2017.

[6] A. Allard, P.-A. Noël, L. Dubé, and B. Pourbohloul, "Heterogeneous Bond Percolation on Multitype Networks with an Application to Epidemic Dynamics," *Physical review. E, Statistical, nonlinear, and soft matter physics*, vol. 79, 2009.

[7] O. Yağan, D. Qian, J. Zhang, and D. Cochran, "Conjoining Speeds up Information Diffusion in Overlaying Social-Physical Networks," *IEEE Journal on Selected Areas in Communications*, vol. 31, no. 6, 2013.

[8] M. E. J. Newman, "Spread of epidemic disease on networks," *Physical Review E*, vol. 66, no. 1, 2002.

[9] P. Grassberger, "On the critical behavior of the general epidemic process and dynamical percolation," *Mathematical Biosciences*, vol. 63, no. 2, 1983.

[10] Y. Zhuang and O. Yağan, "Information propagation in clustered multilayer networks," *IEEE Transactions on Network Science and Engineering*, vol. 3, no. 4, pp. 211–224, 2016.

[11] M. E. Newman, "Random graphs with clustering," *Physical review letters*, vol. 103, no. 5, p. 058701, 2009.

[12] J. C. Miller, "Percolation and epidemics in random clustered networks," *Physical Review E*, vol. 80, no. 2, p. 020901, 2009.

[13] C. Buono, L. G. Alvarez-Zuzek, P. A. Macri, and L. A. Braunstein, "Epidemics in Partially Overlapped Multiplex Networks," *PLOS ONE*, vol. 9, no. 3, 2014.

[14] A. Hackett, D. Cellai, S. Gómez, A. Arenas, and J. Gleeson, "Bond Percolation on Multiplex Networks," *Physical Review X*, vol. 6, p. 021002, Apr. 2016.

[15] Y. Tian, A. Sridhar, C. W. Wu, S. A. Levin, K. M. Carley, H. V. Poor, and O. Yagan, "The Role of Masks in Mitigating Viral Spread on Networks," Jan. 2023.

[16] X. Chen, G. Zhu, L. Zhang, Y. Fang, L. Guo, and X. Chen, "Age-Stratified COVID-19 Spread Analysis and Vaccination: A Multitype Random Network Approach," *IEEE transactions on network science and engineering*, vol. 8, no. 2, 2021.

[17] Y. Tian, A. Sridhar, O. Yağan, and H. V. Poor, "Analysis of the Impact of Mask-wearing in Viral Spread: Implications for COVID-19," in *2021 American Control Conference (ACC)*, pp. 3132–3137, May 2021. ISSN: 2378-5861.

[18] D.-S. Lee and M. Zhu, "Epidemic Spreading in a Social Network with Facial Masks wearing Individuals," *IEEE Transactions on Computational Social Systems*, vol. 8, no. 6, 2021.

[19] G. Bonifazi, B. Breve, S. Cirillo, E. Corradini, and L. Virgili, "Investigating the COVID-19 vaccine discussions on Twitter through a multilayer network-based approach," *Information Processing & Management*, vol. 59, no. 6, 2022.

[20] G. J. Milne, J. K. Kelso, H. A. Kelly, S. T. Huband, and J. McVernon, "A Small Community Model for the Transmission of Infectious Diseases," *Plos One*, vol. 3, no. 12, 2008.

[21] C. Bongiorno and L. Zino, "A multi-layer network model to assess school opening policies during a vaccination campaign," *Applied Network Science*, vol. 7, no. 1, 2022.

[22] Y. Tian and O. Yagan, "Spreading processes with population heterogeneity over multi-layer networks," Nov. 2022.

[23] "2022 school mask guide," 2022. https://www.cbsnews.com/pittsburgh/news/2022-school-mask-guide/.

[24] M. Molloy, "A Critical Point For Random Graphs With A Given Degree Sequence,"

[25] M. E. J. Newman, S. H. Strogatz, and D. J. Watts, "Random graphs with arbitrary degree distributions and their applications," *Physical Review E*, vol. 64, no. 2, 2001.

[26] J. W. Chun and M. J. Lee, "When does individuals' willingness to speak out increase on social media? Perceived social support and perceived power/control," *Computers in Human Behavior*, vol. 74, pp. 120–129, Sept. 2017.

[27] J. R. Silvester, "Determinants of block matrices," *The Mathematical Gazette*, vol. 84, no. 501, p. 460–467, 2000.

[28] R. Eletreby, Y. Zhuang, K. M. Carley, O. Yağan, and H. V. Poor, "The effects of evolutionary adaptations on spreading processes in complex networks," *Proceedings of the National Academy of Sciences*, vol. 117, no. 11, 2020.

[29] B. Söderberg, "Properties of random graphs with hidden color," *Physical Review. E, Statistical, Nonlinear, and Soft Matter Physics*, vol. 68, no. 2 Pt 2, 2003.

[30] K. Athreya and P. Ney, *Branching Processes*, vol. 196. 1972.

[31] E. A. Leicht and R. M. D'Souza, "Percolation on interacting networks," July 2009. arXiv:0907.0894 [cond-mat].