Impact of behavioral heterogeneity on epidemic outcome and its mapping into effective network topologies

Fabio Mazza,[1] Gabriele Ricci,[2] Francesca Colaiori,[3,4] Stefano Guarino,[2] Sandro Meloni,[2,5,6] and Fabio Saracco[6,7,8]

[1]*Dipartimento di Elettronica, Informazione e Bioingegneria, Politecnico di Milano, Milano, Italy*
[2]*Istituto per le Applicazioni del Calcolo "Mauro Picone," Consiglio Nazionale delle Ricerche, Roma, Italy*
[3]*Istituto dei Sistemi Complessi, Consiglio Nazionale delle Ricerche, Roma, Italy*
[4]*Dipartimento di Fisica, Universitá di Roma La Sapienza, Roma, Italy*
[5]*Institute for Cross-Disciplinary Physics and Complex Systems (IFISC), CSIC-UIB, Palma de Mallorca, Spain*
[6]*Centro Studi e Ricerche "Enrico Fermi," Roma, Italy*
[7]*IMT School For Advanced Studies Lucca, Lucca, Italy*
[8]*Istituto per le Applicazioni del Calcolo "Mauro Picone," Consiglio Nazionale delle Ricerche, Sesto Fiorentino, Italy*

(Received 19 September 2025; accepted 18 December 2025; published 8 January 2026)

Human behavior plays a critical role in shaping epidemic trajectories. During health crises, people respond in diverse ways in terms of self-protection and adherence to recommended measures, largely reflecting differences in how individuals assess risk. This behavioral variability induces effective heterogeneity into key epidemic parameters, such as infectivity and susceptibility. We introduce a minimal extension of the susceptible-infected-removed (SIR) model, denoted HeSIR, that captures these effects through a simple bimodal scheme, where individuals may have higher- or lower-transmission-related traits. We derive a closed-form expression for the epidemic threshold in terms of the model parameters, and the network's degree distribution and homophily, defined as the tendency of like-risk individuals to preferentially interact. We identify a resurgence regime just beyond the classical threshold, where the number of infected individuals may initially decline before surging into large-scale transmission. Through simulations on homogeneous and heterogeneous network topologies we corroborate the analytical results and highlight how variations in susceptibility and infectivity influence the epidemic dynamics. We further show that, under suitable assumptions, the HeSIR model maps onto a standard SIR process on an appropriately modified contact network, providing a unified interpretation in terms of structural connectivity. Our findings quantify the effect of heterogeneous behavioral responses, especially in the presence of homophily, and caution against underestimating epidemic potential in fragmented populations, which may undermine timely containment efforts. The results also extend to heterogeneity arising from biological or other nonbehavioral sources.

DOI: 10.1103/xps9-qqvp

I. INTRODUCTION

The presence of heterogeneity at multiple levels—biological, behavioral, and social—plays a crucial role in shaping epidemic spreading dynamics. Early works in network science demonstrated how structural differences, such as degree variability in contact networks, can substantially alter both the epidemic threshold and outbreak sizes [1–4]. In particular, variations in the number and type of social interactions across individuals significantly influence the overall spreading potential of a pathogen [5–7].

Beyond structural heterogeneity, individual-level biological variations—reflected in key epidemic parameters such as susceptibility, infectivity, and infectious period—also play a critical role. Studies have shown that heterogeneity in susceptibility to infection and infectiousness can markedly affect both the epidemic threshold and the final outbreak size [8–21]. Building on these insights, Miller introduced a modeling framework capable of incorporating general distributions for these parameters and characterized the epidemic threshold as a function of their heterogeneity [22].

In addition to structural and biological heterogeneity, the COVID-19 pandemic has underscored the pivotal role of human behavior in epidemic trajectories. Behavioral variability can arise from economic and social factors, leading to differences in compliance with control measures, as discussed in recent works on epidemic fatigue and adherence [23,24], and studies on economic disparities in adherence behavior [25,26]. Another significant source of variability stems from information exposure and political affiliation, which have been shown to influence individual risk behaviors during epidemics [27].

Within this context, it is essential to model how different behaviors—potentially driven by information or misinformation—affect disease spread. Since the adoption of such behaviors often depends on homophily among individuals with similar beliefs, investigating the influence of homophily on epidemic dynamics becomes particularly important.

In this work, we develop a generalized susceptible-infected-removed (SIR) model (HeSIR) that incorporates individual behavioral heterogeneity in its simplest form through a mixture of two delta functions, corresponding to

low- and high-compliance individuals. We expect this minimal representation to capture the essential behavior of more general bimodal distributions, where the population consists of two groups with distinct risk-related traits. We thus consider two behavioral classes, low-risk (L) and high-risk (H) individuals, where high-risk individuals display enhanced susceptibility and infectivity, quantified by factors $\alpha_S > 1$ and $\alpha_I > 1$. These differences in behavioral compliance might reflect exposure to misinformation [27,28]. Modeling compliance as independent heterogeneity in infectivity and susceptibility, rather than as a single "caution" parameter, enables us to capture the asymmetric effects of nonpharmaceutical interventions such as the use of face masks or social distancing. For instance, mask wearing by a susceptible individual reduces their infection probability (lowering susceptibility), while mask wearing by an infected individual primarily limits onward transmission (reducing infectivity). Empirical and modeling studies during COVID-19 support this distinction, showing that interventions that mainly protect the adopter (reducing susceptibility) can be more effective in limiting epidemic prevalence than those that primarily protect contacts (reducing infectivity) with comparable efficacy [29].

Using a heterogeneous mean-field (HMF) ap proximation, we derive an analytical expression for the epidemic threshold that holds for networks with arbitrary degree distributions and a tunable level of homophily, defined as the tendency of like-risk individuals to preferentially interact. This result generalizes classical threshold conditions derived for homogeneous networks and uncovers nontrivial dependencies on both behavioral heterogeneity and network structure.

To validate our analytical findings, we perform extensive simulations on random networks generated via the degree-corrected stochastic block model (DCSBM) [2,30], also considering the special cases of the stochastic block model (SBM), configuration model (CM), and Erdős-Rényi (ER) model, obtained in the absence of degree heterogeneity, homophily, or both. In each case, we estimate the epidemic threshold by locating the peak of epidemic variability, a method known to accurately identify critical points in finite-size systems [31]. The excellent agreement between simulation and theory across different network topologies and mixing patterns confirms the robustness of our generalized threshold formula.

Our simple model exhibits a rich dynamical behavior, including the emergence of a "resurgence zone," a phenomenology that we have already described and analyzed in a previous work [32]. In a well-mixed population, this corresponds to an interval between two theoretical thresholds, existing only when both $\alpha_S > 1$ and $\alpha_I > 1$, in which the infected population initially declines before exploding into a full outbreak. Interestingly, we find that heterogeneous network structure alone can produce a similar resurgence phenomenon even in the standard SIR limit ($\alpha_S = \alpha_I = 1$), suggesting a unifying mechanism through which structural and epidemiological heterogeneity both sustain hidden transmission potential.

Finally, we show that, in expectation, the HeSIR model is equivalent to a standard SIR process on a directed SBM with a bimodal out-degree distribution. In this mapping, the parameters α_S and α_I translate into modulations of effective contact frequency rather than per-contact transmission risk, revealing a conceptual duality: heterogeneity in epidemic parameters can be recast as heterogeneity in contact structure without altering threshold conditions.

In this work, we thus provide a unified framework for understanding how multilevel heterogeneity—including biological, structural, and behavioral factors—shapes epidemic thresholds and dynamics, with important implications for risk assessment and intervention design.

II. HETEROGENEOUS EPIDEMIC SPREADING MODEL

Let I_H, I_L, S_H, S_L, and R denote the fractions of high-risk infected, low-risk infected, high-risk susceptible, low-risk susceptible, and recovered individuals, respectively. The total population is conserved: $I_H + I_L + S_H + S_L + R = 1$. The mean-field equations for the HeSIR model read

$$\begin{aligned}
\dot{S}_H &= -\alpha_S \lambda S_H I_{\text{eff}}, \\
\dot{S}_L &= -\lambda S_L I_{\text{eff}}, \\
\dot{I}_H &= \alpha_S \lambda S_H I_{\text{eff}} - \gamma I_H, \\
\dot{I}_L &= \lambda S_L I_{\text{eff}} - \gamma I_L, \\
\dot{R} &= \gamma I.
\end{aligned} \tag{1}$$

$I = I_H + I_L$ denotes the total prevalence of infected individuals, whereas $I_{\text{eff}} = I_L + \alpha_I I_H$ is the "effective" fraction of infected individuals, due to the increased infectivity of the I_H class. We set the timescale by fixing the susceptibility of low-risk individuals to 1, and consequently that of high-risk individuals to α_S. In the more general case, these susceptibilities could be defined as σ and $\alpha_S\sigma$, respectively. Individuals in the state S_L thus become infected and flow to I_L at a rate λI_{eff} that is increased by a factor α_S for individuals in the state S_H, and $\lambda = \beta\langle k\rangle$ is the transmission rate, given by the individual transmission rate β times the average contact capacity of the nodes $\langle k\rangle$. From both I_L and I_H, individuals then flow to R at a rate γ by spontaneous recovery.

A. Heterogeneous percolation threshold on a network with homophily

To analyze the model's behavior in a networked population, we consider a degree-corrected SBM (DCSBM) [33] with two communities: $N_L = (1-p)N$ low-risk and $N_H = pN$ high-risk individuals. Let $h \in [0, 1]$ be a homophily parameter interpolating between random mixing ($h = 0$) and full assortativity ($h = 1$). The matrix of conditional probabilities that a randomly chosen neighbor of a node in group i belongs to group j is given by

$$\begin{pmatrix} \pi_{LL} & \pi_{LH} \\ \pi_{HL} & \pi_{HH} \end{pmatrix} = (1-h)\begin{pmatrix} 1-p & p \\ 1-p & p \end{pmatrix} + h\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}. \tag{2}$$

This matrix is defined for $p \in (0, 1)$.

Let x_L and x_H be the probabilities that a vertex of type L or H is not connected to the giant component through a randomly chosen edge. Considering a generic degree distribution, we denote with q_k the corresponding excess degree distribution and with $g_1(x) = \sum_k q_k x^k$ the generating function for the excess degree distribution. We can determine the size of the

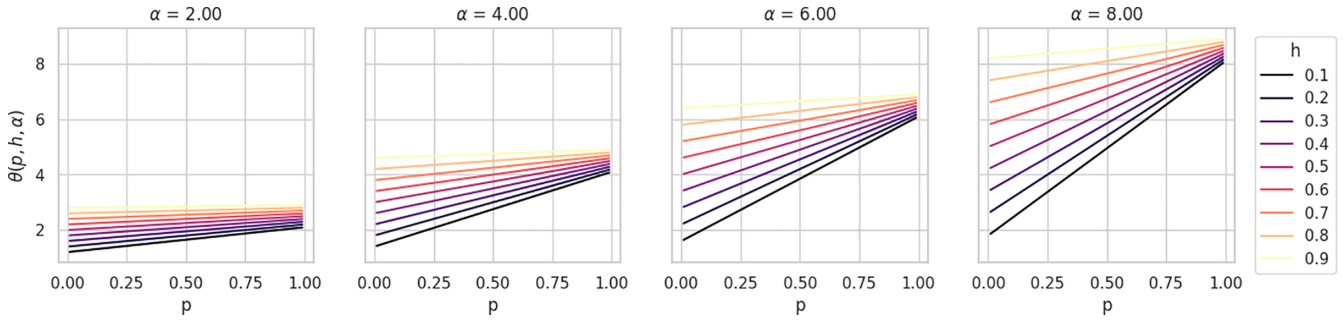


FIG. 1. The value of $\theta(p, h, \alpha)$ for different values of p, h, and α. $\theta(p, h, \alpha)$ can be interpreted as the ratio between the threshold in the baseline SIR case and the threshold in the HeSIR model. As expected, $\theta(p, h, \alpha)$ increases with p, h, and α. For large h (respectively, p), the dependence on p (respectively, h) weakens.

giant component by finding the fixed point of the equation,

$$\begin{aligned} x_L &= \pi_{LL}[1 - \phi_{LL} + \phi_{LL} g_1(x_L)] \\ &\quad + \pi_{LH}[1 - \phi_{LH} + \phi_{LH} g_1(x_H)], \\ x_H &= \pi_{HL}[1 - \phi_{HL} + \phi_{HL} g_1(x_L)] \\ &\quad + \pi_{HH}[1 - \phi_{HH} + \phi_{HH} g_1(x_H)], \end{aligned}$$

where the activation probabilities ϕ_{ij} (transmission probability across an edge from group i to group j) are given by $\phi_{LL} = 1 - e^{-\beta/\gamma}$, $\phi_{LH} = 1 - e^{-\alpha_I \beta/\gamma}$, $\phi_{HL} = 1 - e^{-\alpha_S \beta/\gamma}$, and $\phi_{HH} = 1 - e^{-\alpha_S \alpha_I \beta/\gamma}$. The percolation threshold is found by solving

$$\begin{aligned} &(\pi_{LH}\pi_{HL}\phi_{LH}\phi_{HL} - \pi_{HH}\pi_{LL}\phi_{HH}\phi_{LL}) \dot{g}_1^2(1) \\ &\quad + (\pi_{LL}\phi_{LL} + \pi_{HH}\phi_{HH}) \dot{g}_1(1) - 1 \qquad = 0. \end{aligned} \tag{3}$$

Equation (3) admits no closed-form solution, but can be solved numerically. Further, if we assume $\beta/\gamma \ll 1$, we can linearize (3) with respect to β/γ and ignore the first term of the equation. Denoting $\alpha = \alpha_I \alpha_S$ and $\theta(p, h, \alpha) = \pi_{LL} + \alpha \pi_{HH} = (1 - h)(1 - p + \alpha p) + h(\alpha + 1)$, we obtain

$$\left(\frac{\beta}{\gamma}\right)_c = \frac{\langle k \rangle}{\langle k^2 \rangle - \langle k \rangle} \frac{1}{\theta(p, h, \alpha)}, \tag{4}$$

where we used $\dot{g}_1(1) = \frac{\langle k^2 \rangle - \langle k \rangle}{\langle k \rangle}$.

In epidemiological terms, the percolation threshold $(\frac{\beta}{\gamma})_c$ corresponds to the critical value of the transmission-to-recovery ratio at which the system undergoes a transition from disease extinction to epidemic outbreak. For $\frac{\beta}{\gamma} > (\frac{\beta}{\gamma})_c$, the infection can spread within the population, whereas for $\frac{\beta}{\gamma} < (\frac{\beta}{\gamma})_c$ the outbreak is expected to eventually die out.

Note that for $\theta = 1$ Eq. (4) reduces to the epidemic threshold for the standard SIR in a CM. For small β/γ, the threshold depends on α_S and α_I only through their product α. For a network with no homophily ($h = 0$), θ simplifies to $\theta(p, 0, \alpha) = 1 - p + \alpha p$. The dependence of θ on its arguments is shown in Fig. 1.

A detailed analytical derivation of the epidemic threshold is reported in Appendix A.

B. Early dynamics and effective threshold

To analyze the early-time behavior, we assume a small initial infection uniformly distributed across the population: $I_L(0) = (1 - p)\epsilon$, $I_H(0) = p\epsilon$, $S_L(0) \approx 1 - p$, $S_H(0) \approx p$. Setting $R(0) = 0$, we compare the conditions under which $\dot{I}(0) = 0$ and $\dot{I}_{\text{eff}}(0) = 0$, that translate to

$$\left(\frac{\beta}{\gamma}\right)_I = \frac{1}{\langle k \rangle (1 - p + \alpha_S p)(1 - p + \alpha_I p)} \frac{1}{1 - \epsilon}, \tag{5}$$

$$\left(\frac{\beta}{\gamma}\right)_{I_{\text{eff}}} = \frac{1}{\langle k \rangle (1 - p + \alpha_S \alpha_I p)} \frac{1}{1 - \epsilon}. \tag{6}$$

In the limit $\epsilon \to 0$ and $\langle k \rangle - 1 \approx \langle k \rangle$, (6) matches our epidemic threshold (4) for a homogeneous network without homophily. On the other hand, again in the limit $\epsilon \to 0$ and $\langle k \rangle - 1 \approx \langle k \rangle$, (5) coincides with the estimate found in [22] in the special case of $\langle k \rangle$-regular graphs. The latter is obtained from bond percolation assuming that all edges have the same average occupation probability:

$$\phi = (1 - p)^2 \phi_{LL} + p(1 - p)\phi_{LH} + p(1 - p)\phi_{HL} + p^2 \phi_{HH}. \tag{7}$$

The two approaches—averaging between each group separately or throughout the network—lead to distinct thresholds only when both $\alpha_S > 1$ and $\alpha_I > 1$. This underlines the importance of accounting for both heterogeneous susceptibility and infectivity, although (4) depends on α_S and α_I only through their product. The analysis of early dynamics shows that the combined effect of α_S and α_I makes it possible for the total prevalence $I(t)$ to initially decrease, while the effective infectious pool $I_{\text{eff}}(t)$ grows enough to eventually trigger a resurgence.

III. SIMULATIONS

To validate the analytical predictions of our model, we performed extensive stochastic simulations of the the HeSIR dynamics on DCSBM networks. This family of networks generalizes the CM, SBM, and ER graphs by tuning the homophily parameter h and the degree distribution. We simulate HeSIR epidemics with an optimized Gillespie algorithm [34] in networks with $N = 5 \times 10^4$ nodes and average degree $\langle k \rangle = 20$, with homogeneous and heterogeneous degree distributions (Pareto degree distributions with exponent $a \in \{2.0, 2.5\}$), and different levels of homophily h and high-risk population fraction p_H. In all simulations, we take $I(0) = 10^{-3}$ (i.e., 50 seed nodes when $N = 5 \times 10^4$, taken uniformly at random) and $R(0) = 0$.

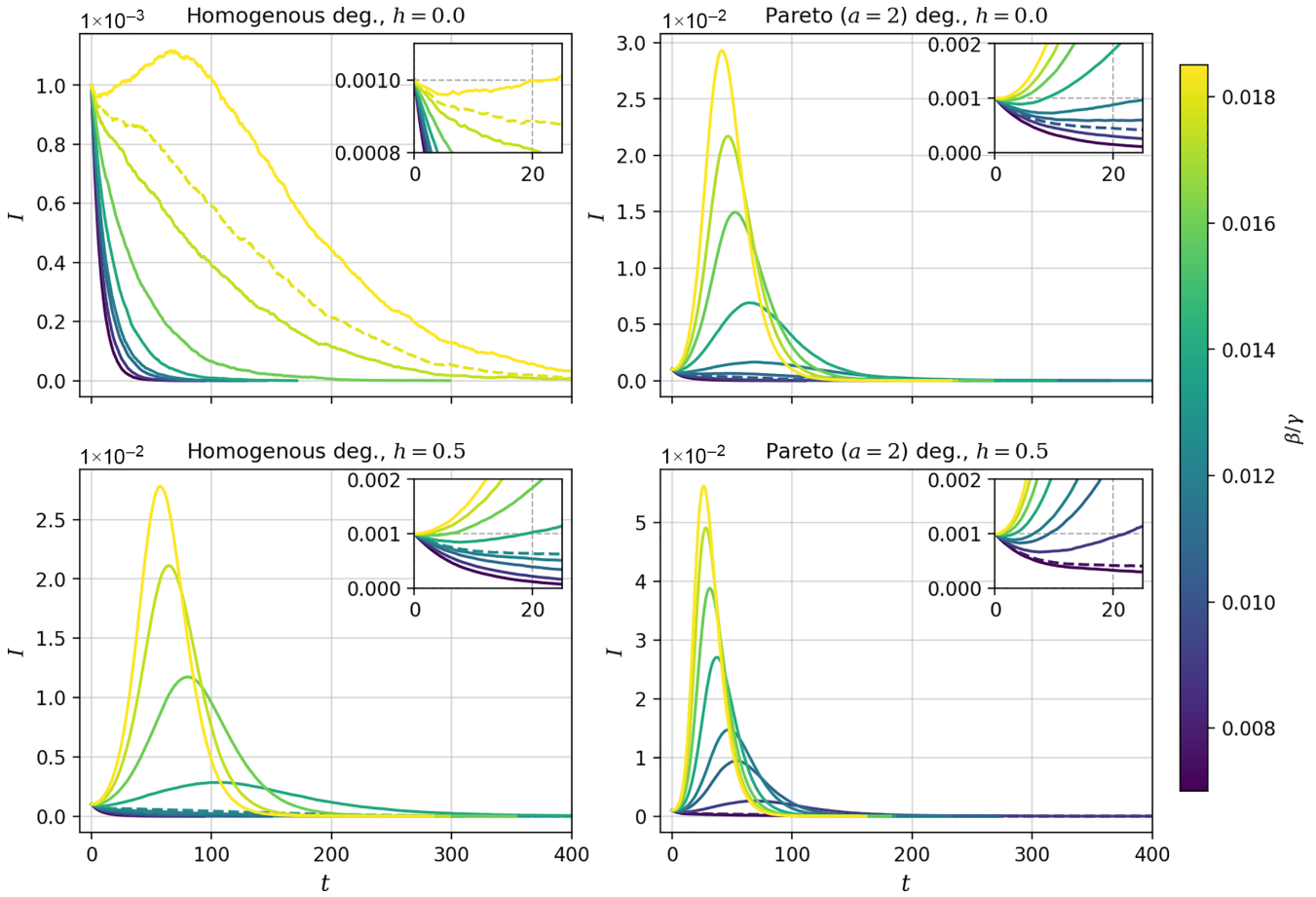


FIG. 2. Curves of the mean number of infected under different configurations of degree heterogeneity and homophily, with $p_H = 0.4$, $\alpha_I = 2$, $\alpha_S = 3$. Results averaged over 400 simulations per setting, for networks of $N = 5 \times 10^4$ nodes with mean degree $\langle k \rangle = 20$, with 50 seed nodes chosen uniformly at random. The dashed lines represent the critical values of β/γ maximizing the epidemic variability V_{AR}. The insets provide a close-up of the outbreak's early dynamics, highlighting the resurgence zone observed in some supercritical regimes: an initial decline in the infected population followed by renewed growth toward a full outbreak.

The mixing structure is determined by the edge-frequency matrix:

$$\begin{pmatrix} p_{LL} & p_{LH} \\ p_{HL} & p_{HH} \end{pmatrix} \approx \frac{\langle k \rangle}{N} \begin{pmatrix} \frac{(1-h)(1-p)+h}{1-p} & 1-h \\ 1-h & \frac{(1-h)p+h}{p} \end{pmatrix}, \quad (8)$$

which for large N maps (2) to the expected frequency of connections between groups for a given h and p.

A. Epidemic curves across mixing and heterogeneity

We analyzed the temporal evolution of the epidemic for different model parameters. Figure 2 shows representative epidemic curves for the fraction of infected individuals over time, comparing the homogeneous and heterogeneous settings, both with and without homophily.

As seen in the figure, the qualitative shape of the epidemic curve is consistent across all settings. The presence of degree heterogeneity leads to earlier peaks due to superspreading individuals. By creating clusters of high-risk individuals associated to acute spreading, homophily facilitates the onset and significantly increases peak prevalence.

B. Epidemic threshold: Theory versus simulation

To assess the accuracy of the analytical epidemic threshold derived in Eq. (4), we compare three different indicators:

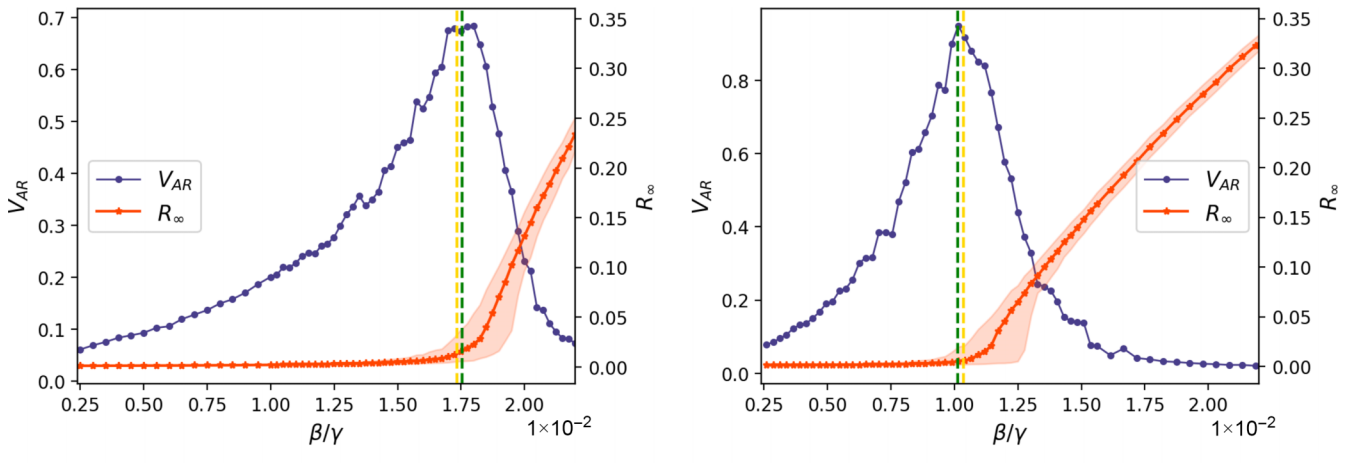
(1) Theoretical threshold $(\beta/\gamma)_c$ from heterogeneous percolation.

(2) The value of β/γ at which the final epidemic size R_∞ exhibits a sharp transition.

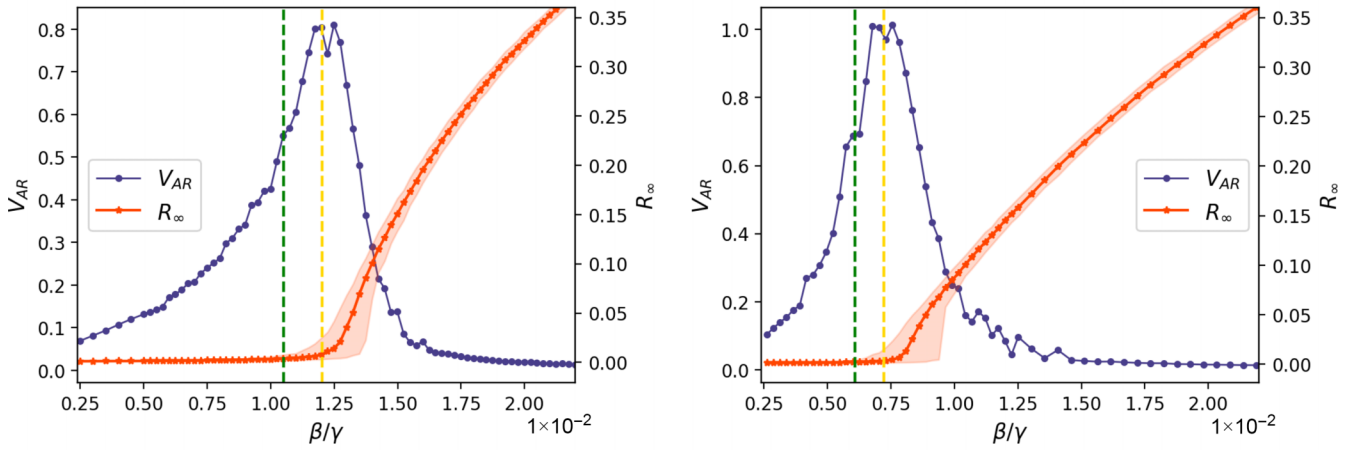
(3) The value maximizing the epidemic variability $V_{\mathrm{AR}} = \sqrt{\langle R_\infty^2 \rangle - \langle R_\infty \rangle^2}/\langle R_\infty \rangle$.

Figure 3 illustrates this comparison for various network configurations, with the actual threshold, computed numerically, identified by the yellow dashed line, and the closed-form approximation, given by (4), shown by the green dashed line.

Overall, the numerical estimates for the threshold confirm the analytical predictions for all the cases considered. The position of the phase transition in the attack rate and the peak of V_{AR} closely match the percolation-based estimates. The small discrepancies observed near the transition point, particularly in highly heterogeneous networks, can be attributed to finite-size effects and stochastic fluctuations near criticality.



(a) Homogeneous degrees, $h = 0$ (b) Pareto degrees ($a = 2$), $h = 0$



(c) Homogeneous degrees, $h = 0.5$ (d) Pareto degrees ($a = 2$), $h = 0.5$

FIG. 3. Epidemic threshold estimates. Yellow dashed line: analytical threshold, computed numerically; green dashed line: approximation of the analytical threshold, as given by (4). The R_∞ curves report the median value, with shaded areas showing the central 80% of simulated outcomes. Results are averaged over 400 simulations per setting, for networks of $N = 5 \times 10^4$ nodes and mean degree $\langle k \rangle = 20$, with 50 seed nodes chosen uniformly at random. Other parameters set as $p_H = 0.4$, $\alpha_I = 2$, and $\alpha_S = 3$. Additional results in Appendix B confirm the robustness of the analysis across a range of values for α_S, α_I, p_H, N, and $\langle k \rangle$.

IV. MAPPING THE HESIR MODEL ONTO A SIR MODEL ON AN EFFECTIVE NETWORK

We now show that the heterogeneous transmission dynamics of the HeSIR model can be reformulated as a standard SIR process on an appropriately modified network. Since individuals of type H differ in both susceptibility ($\alpha_S \neq 1$) and infectiousness ($\alpha_I \neq 1$), it is natural to ask whether the resulting dynamics could instead be viewed as a homogeneous SIR process on a block-structured contact network in which H and L individuals occupy distinct blocks, such that the heterogeneity in susceptibility and infectiousness is effectively transferred to heterogeneity in the connectivity patterns between and within blocks. We demonstrate below that such a representation is indeed possible under certain assumptions.

A related idea—interpreting heterogeneous transmission through an effective directed network under an annealed approximation—appears in works by Clancy [35,36], although in a different setting and for a different purpose: there, the focus is on susceptible-infected-susceptible (SIS) dynamics and on how heterogeneity in susceptibility and infectiousness influences the persistence time from quasistationary endemicity to extinction.

A. General intuition

Let $P_i \in \{L, H\}$ denote the class of node i. In the HeSIR model, the probability that a susceptible node i is infected by an infectious node j is proportional to

$$p_{ij}^{\text{HeSIR}}\, \beta_{P_j P_i},$$

where p_{ij}^{HeSIR} is the connection probability and

$$\beta_{P_j P_i} = \beta[1 - (1 - \alpha_S)\delta_{P_i,H}][1 - (1 - \alpha_I)\delta_{P_j,H}], \quad (9)$$

so that

$$\begin{aligned} \beta_{LL} &= \beta, \\ \beta_{LH} &= \alpha_S \beta, \\ \beta_{HL} &= \alpha_I \beta, \end{aligned}$$

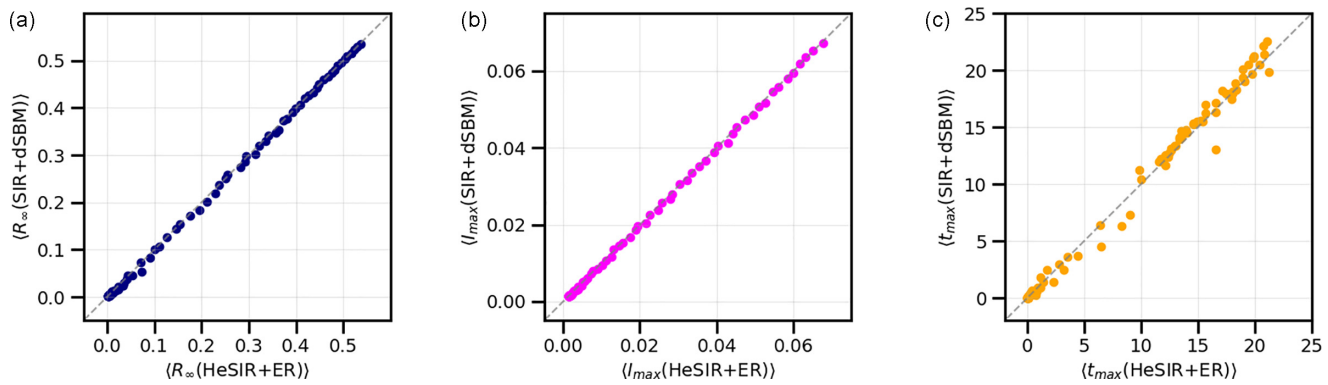


FIG. 4. Model equivalence on homogeneous networks. For $\beta/\gamma \in [0.01, 0.04]$, the average values of R_∞ (a), the frequency of infected individuals at the peak $I_{\max}$ (b), and the time of the peak $t_{\max}$ (c) are shown for both the HeSIR and the equivalent SIR model, exhibiting perfect agreement. Averages are computed over 200 realizations. Parameters are set to $\alpha_I = 2$, $\alpha_S = 3$, and $p_H = 0.4$.

$$\beta_{HH} = \alpha_S \alpha_I \beta.$$

We can now recast the dynamics of the HeSIR model as an equivalent SIR model on an effective network where the heterogeneous transmission is reabsorbed into the edge structure. This amounts to rescaling edge probabilities:

$$p_{ij}^{\mathrm{SIR}}\,\beta = p_{ij}^{\mathrm{HeSIR}}\,\beta_{P_jP_i} \quad \Rightarrow \quad p_{ij}^{\mathrm{SIR}} = p_{ij}^{\mathrm{HeSIR}}\,\frac{\beta_{P_jP_i}}{\beta}.$$

That is, the SIR model on the effective graph has a homogeneous infection rate β, while encoding heterogeneity within the connection structure.

Before showing some examples of this equivalence, let us add some remarks. First, note that $\beta_{P_iP_j}$ factorizes into two node-specific terms; see Eq. (9). Therefore, the equivalence works properly when p_{ij}^{HeSIR} can be expressed as the product of quantities per node, i.e., $p_{ij}^{\mathrm{HeSIR}} = x_i x_j$. In this way, the node-dependent term in $\beta_{P_iP_j}$ can be absorbed in the relative x of p_{ij}^{SIR}. Although this restricts the class of admissible network models (excluding, for example, maximum-entropy ones [37]), the most studied ones, such as Erdős-Rényi random graphs, stochastic block models, Chung-Lu configuration models [38], degree-corrected block models [33], and fitness models [39], are included in this category. Second, a natural question is what is the worth of such an equivalence. Besides the intellectual interest, the main point is that it is beneficial to translate nontrivial results from the SIR model into the HeSIR models. Therefore, theoretical results, as well as computational tools, developed in the SIR framework, can be safely used for the HeSIR, using this equivalence, with limited efforts. In this sense, consider that going, instead, from HeSIR to SIR is much more cumbersome, as when $\alpha_S \neq \alpha_I$, the equivalent contact network for the SIR model is a directed one, i.e., something nonphysical, as standard contacts do not show a preferential direction. In a sense, a heterogeneity in the behavior does not directly correspond to heterogeneous patterns of contact, but to something more involved. In summary, while the mapping holds exactly only under specific structural assumptions, it provides a powerful conceptual bridge between parameter heterogeneity and structural heterogeneity.

B. Homogeneous networks: dSBM equivalence

Consider the case where the interaction network of the HeSIR model is an ER graph with connection probability $p_{ij}^{\mathrm{HeSIR}} = p_{\mathrm{ER}}$. The corresponding effective SIR network is defined by

$$p_{ij}^{\mathrm{SIR}} = p_{\mathrm{ER}}[1-(1-\alpha_S)\delta_{P_i,H}][1-(1-\alpha_I)\delta_{P_j,H}],$$

i.e., a directed SBM (dSBM) where the edge probability depends on both the source and target group. When $\alpha_I \neq \alpha_S$, the effective network is asymmetric: the probability of an edge from H to L differs from the reverse. Figure 4 shows the equivalence between the two models via a comparison of R_∞, the frequency of infected individuals at the peak $I_{\max}$, and the time of the peak $t_{\max}$ across various values of β/γ.

C. Heterogeneous networks: dDCSBM equivalence

When the HeSIR model is defined over a CM, such as a Chung-Lu network, the equivalent SIR model corresponds to a directed degree-corrected stochastic block model (dDCSBM). In this model, nodes have both in- and out-degree parameters that depend on their class, encoding differences in both susceptibility and infectivity.

Let k_i be the degree of node i in the CM. We define effective in and out degrees as

$$k_i^{\mathrm{in}} = f(P_i)k_i, \quad k_i^{\mathrm{out}} = g(P_i)k_i,$$

where f and g are stretching functions that depend exclusively on the class of the node. Note that, for consistency, $f(P) = g(P) = 1$ if $\alpha_S = \alpha_I = 1$. Equating the transmission terms between HeSIR and SIR on the effective network leads to

$$\frac{\beta\, k_i k_j}{2m}\,\frac{\beta_{P_jP_i}}{\beta} = \beta\, k_i^{\mathrm{out}} k_j^{\mathrm{in}}\,\frac{m_{P_iP_j}}{m_{P_i}^{\mathrm{out}} m_{P_j}^{\mathrm{in}}}, \tag{10}$$

where $m = \sum_i k_i/2$ is the total number of links in the CM; $m_{P_iP_j}$, $m_{P_i}^{\mathrm{out}}$, and $m_{P_j}^{\mathrm{in}}$ are, respectively, the number of links from the block P_i to the block P_j, the number of outgoing links from the block P_i, and the number of ingoing links in the block P_j, in the dDCSBM. Equation (10) yields the following set of

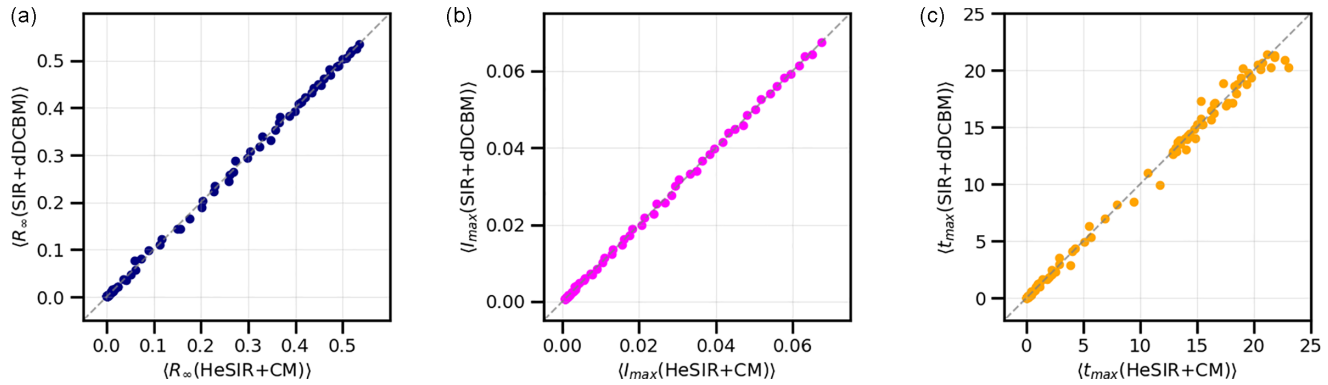


FIG. 5. Model equivalence on heterogeneous networks. For $\beta/\gamma \in [0.01, 0.04]$, the average values of R_∞ (a), fraction of infected individuals at the peak $I_{\max}$ (b), and the time of the peak $t_{\max}$ (c) are shown for both the HeSIR and the equivalent SIR model, exhibiting perfect agreement. Averages are computed over 200 realizations. Parameters are set to $\alpha_I = 2$, $\alpha_S = 3$, and $p_H = 0.4$. Node degrees in the CM network follow a Pareto distribution with exponent $a = 2$.

constraints:

$$\frac{1}{2m} = f(L)g(L)\frac{m_{LL}}{m_L^{\text{out}}m_L^{\text{in}}},$$
$$\frac{\alpha_S}{2m} = f(H)g(L)\frac{m_{LH}}{m_L^{\text{out}}m_H^{\text{in}}},$$
$$\frac{\alpha_I}{2m} = f(L)g(H)\frac{m_{HL}}{m_H^{\text{out}}m_L^{\text{in}}},$$
$$\frac{\alpha_S\alpha_I}{2m} = f(H)g(H)\frac{m_{HH}}{m_H^{\text{out}}m_H^{\text{in}}}.$$

If $\langle m_L\rangle$ and $\langle m_H\rangle$ are, respectively, the expected values under CM of the total number of links incident on L and H, then

$$m_P^{\text{in}} = f(P)\langle m_P\rangle, \quad m_P^{\text{out}} = g(P)\langle m_P\rangle.$$

Therefore, the edge counts read

$$m_{LL} = \frac{\langle m_L\rangle^2}{2m},$$
$$m_{LH} = \alpha_S\frac{\langle m_L\rangle\langle m_H\rangle}{2m},$$
$$m_{HL} = \alpha_I\frac{\langle m_L\rangle\langle m_H\rangle}{2m},$$
$$m_{HH} = \alpha_S\alpha_I\frac{\langle m_H\rangle^2}{2m}.$$

Solving for the stretching functions, we obtain

$$\frac{f(H)}{f(L)} = \alpha_S, \quad f(L) = \frac{\langle m_L\rangle + \alpha_I\langle m_H\rangle}{2m},$$
$$\frac{g(H)}{g(L)} = \alpha_I, \quad g(L) = \frac{\langle m_L\rangle + \alpha_S\langle m_H\rangle}{2m}.$$

Again, when $\alpha_S \neq \alpha_I$, the resulting dDCSBM is asymmetric, reflecting asymmetric transmission dynamics. The equivalence shows that the impact of behavioral heterogeneity in the HeSIR model can be fully encoded in the topology of an effective network for SIR dynamics. Figure 5 shows the equivalence between the two models via a comparison of R_∞, $I_{\max}$, and $t_{\max}$ across various values of β/γ.

V. CONCLUSION

This work introduces the HeSIR model that extends the classical SIR model to incorporate heterogeneity in both susceptibility and infectivity. By modeling these traits as bimodally distributed across the population, capturing, for instance, compliant versus noncompliant behavioral types, we uncover how behavioral asymmetries significantly affect epidemic dynamics. We derive closed-form analytical expressions for the epidemic threshold under homogeneous and heterogeneous network topologies. These estimates provide insight into the interplay between α_S, α_I, population composition, and the network mixing structure. In particular, we demonstrate that the epidemic threshold depends not merely on the average transmissibility, but on a weighted effective quantity that reflects the joint distribution of susceptibility and infectivity across contacts. Our theoretical predictions are corroborated by extensive stochastic simulations, showing excellent agreement across a wide range of settings, including networks with degree heterogeneity and varying levels of homophily. Our findings reveal that the presence of heterogeneity in both susceptibility and infectivity is necessary to observe nontrivial effects such as the emergence of a resurgence zone, where a second epidemic transition appears beyond the classical threshold.

Our findings reveal that the presence of heterogeneity in both susceptibility and infectivity is necessary to observe nontrivial effects such as the emergence of a resurgence zone. When both $\alpha_S > 1$ and $\alpha_I > 1$, in fact, infections can initially decline yet subsequently grow if $(\frac{\beta}{\gamma})_{I_{\text{eff}}} < \frac{\beta}{\gamma} < (\frac{\beta}{\gamma})_I$. Only in the presence of this dual heterogeneity the threshold obtained from the classical percolation problem does not correspond to the initial growth of infections within the mean-field approximation, but rather to the initial increase of the *effective* infections $I_{\text{eff}} = I_L + \alpha_I I_H$. This phenomenon is further amplified in homophilic networks and, remarkably, can also arise in purely structural models such as degree-heterogeneous networks, even in the absence of behavioral heterogeneity. These results highlight the importance of individual-level behavior and network structure in shaping epidemic outcomes.

Beyond theoretical contributions, the model has clear implications for epidemic control. In particular, interventions targeting individuals with low compliance (e.g., through testing, communication, or isolation) may be essential to prevent epidemic resurgence, as these individuals disproportionately drive transmission under HeSIR-like dynamics. Conversely, in networks with strong degree heterogeneity, monitoring and protecting high-degree nodes remains a priority regardless of behavioral assumptions. Finally, by mapping the HeSIR dynamics onto equivalent SIR models on appropriately constructed networks (e.g., directed SBM or DCSBM), we unify behavioral and structural heterogeneity into a common framework. This equivalence provides a powerful tool for extending standard epidemic theory to more realistic scenarios. Future work may explore more finely resolved trait distributions, time-dependent behavior, empirical calibration using contact or genomic data, and the extension of these ideas to other epidemic models or coupled spreading processes.

ACKNOWLEDGMENTS

All authors acknowledge support from the project "CODE – Coupling Opinion Dynamics with Epidemics," funded under PNRR Mission 4 "Education and Research" - Component C2 - Investment 1.1 - Next Generation EU "Fund for National Research Program and Projects of Significant National Interest" PRIN 2022 PNRR, Grant Code No. P2022AKRZ9, CUP No. B53D23026080001. S.M. also acknowledges support from the Agencia Estatal de Investigación and Fondo Europeo de Desarrollo Regional (FEDER, UE) under Project APASOS (Grant No. PID2021-122256NB-C22) and the María de Maeztu program, Project No. CEX2021-001164-M, funded by Grant No. MCIN/AEI/10.13039/501100011033.

DATA AVAILABILITY

The data that support the findings of this article are openly available [40].

APPENDIX A: EPIDEMIC THRESHOLD CALCULATION

As commonly done in the literature, the epidemic threshold of the HeSIR model can be determined based on bond percolation. The activation probability ϕ_{ij} for an edge connecting group i to group j is

$$\phi_{LL} = 1 - e^{-\beta/\gamma},\ \phi_{LH} = 1 - e^{-\alpha_I \beta/\gamma},\ \phi_{HL} = 1 - e^{-\alpha_S \beta/\gamma},$$
$$\phi_{HH} = 1 - e^{-\alpha_S \alpha_I \beta/\gamma}.$$

If x_i denotes the probability that a vertex of type i is *not* connected to the giant component through a randomly chosen edge, we can formulate the fixed point system of equations,

$$\begin{aligned} x_L &= \pi_{LL}[1 - \phi_{LL} + \phi_{LL} g_1(x_L)] + \pi_{LH}[1 - \phi_{LH} \\ &\quad + \phi_{LH} g_1(x_H)], \\ x_H &= \pi_{HL}[1 - \phi_{HL} + \phi_{HL} g_1(x_L)] + \pi_{HH}[1 - \phi_{HH} \\ &\quad + \phi_{HH} g_1(x_H)], \end{aligned} \tag{A1}$$

where $g_1(x) := \sum_k q_k x^k$ denotes the generating function of the network's excess degree distribution q_k. System (A1) always admits the trivial solution $(x_L, x_H) = (1, 1)$, corresponding to the absence of a giant percolation cluster. We look for conditions for a second solution to exist in $[0, 1) \times [0, 1)$.

System (A1) can be rewritten as

$$\begin{aligned} x_H &= f_H(x_L), \\ x_L &= f_L(x_H), \end{aligned} \tag{A2}$$

where

$$f_H(x_L) = g_1^{-1}\left(\frac{x_L - \pi_{LL}[1 - \phi_{LL} + \phi_{LL} g_1(x_L)] - \pi_{LH}(1 - \phi_{LH})}{\pi_{LH}\phi_{LH}}\right),$$

and

$$f_L(x_H) = g_1^{-1}\left(\frac{x_H - \pi_{HH}[1 - \phi_{HH} + \phi_{HH} g_1(x_H)] - \pi_{HL}(1 - \phi_{HL})}{\pi_{HL}\phi_{HL}}\right).$$

After algebraic manipulation, one sees that

$$f_L(0) < 0,\ f_H(0) < 0,\ \ddot{f}_L \leqslant 0 \text{ for all } x_H \in [0, 1],$$
$$\ddot{f}_H \leqslant 0 \text{ for all } x_L \in [0, 1].$$

We first consider the case in which $\dot{f}_L(1) < 0$, meaning that f_L has a maximum at some $x_H^{\max} \leqslant 1$ with $f_L(x_H^{\max}) > 1$. In this case, $f_L = 0$ at some $x_H^0 > 0$ and $f_L = 1$ at some $x_H^1 \in (x_H^0, 1)$. Now, let $h(x_H) = f_H[f_L(x_H)] - x_H$. h is continuous, $h(x_H^0) = f_H(0) - x_H^0 < 0$, and $h(x_H^1) = f_H(1) - x_H^1 > 0$, so for the intermediate value theorem there exists $\tilde{x}_H \in (x_H^0, x_H^1)$ such that $h(\tilde{x}_H) = 0$; that is, $f_H[f_L(\tilde{x}_H)] = \tilde{x}_H$. If we call $\tilde{x}_L = f_L(\tilde{x}_H)$, we have $f_H(\tilde{x}_L) = \tilde{x}_H$. In other words, $x_H = f_H(x_L)$ and $x_L = f_L(x_H)$ meet at $(x_L, x_H) = (\tilde{x}_L, \tilde{x}_H) \in [0, 1) \times [0, 1)$.

Similarly, if $\dot{f}_H(1) < 0$, then f_H has a maximum at some $x_L^{\max} \leqslant 1$ with $f_H(x_L^{\max}) > 1$. In this case, $f_H = 0$ at some $x_L^0 > 0$ and $f_H = 1$ at some $x_L^1 \in (x_L^0, 1)$. Now, let $h(x_L) = f_L[f_H(x_L)] - x_L$. h is continuous, $h(x_L^0) = f_L(0) - x_L^0 < 0$, and $h(x_L^1) = f_L(1) - x_L^1 > 0$, so for the IVT there exists $\tilde{x}_L \in (x_L^0, x_L^1)$ such that $h(\tilde{x}_L) = 0$; that is, $f_L[f_H(\tilde{x}_L)] = \tilde{x}_L$. If we call $\tilde{x}_H = f_H(\tilde{x}_L)$, we have $f_L(\tilde{x}_H) = \tilde{x}_L$. In other words, $x_L = f_L(x_H)$ and $x_H = f_H(x_L)$ meet at $(x_L, x_H) = (\tilde{x}_L, \tilde{x}_H) \in [0, 1) \times [0, 1)$.

Finally, let us consider the case in which both $\dot{f}_L(1) \geqslant 0$ and $\dot{f}_H(1) \geqslant 0$. We can invert f_L in $[0, 1]$ and rewrite system (A2) as

$$\begin{aligned} x_H &= f_H(x_L), \\ x_H &= f_L^{-1}(x_L). \end{aligned} \tag{A3}$$

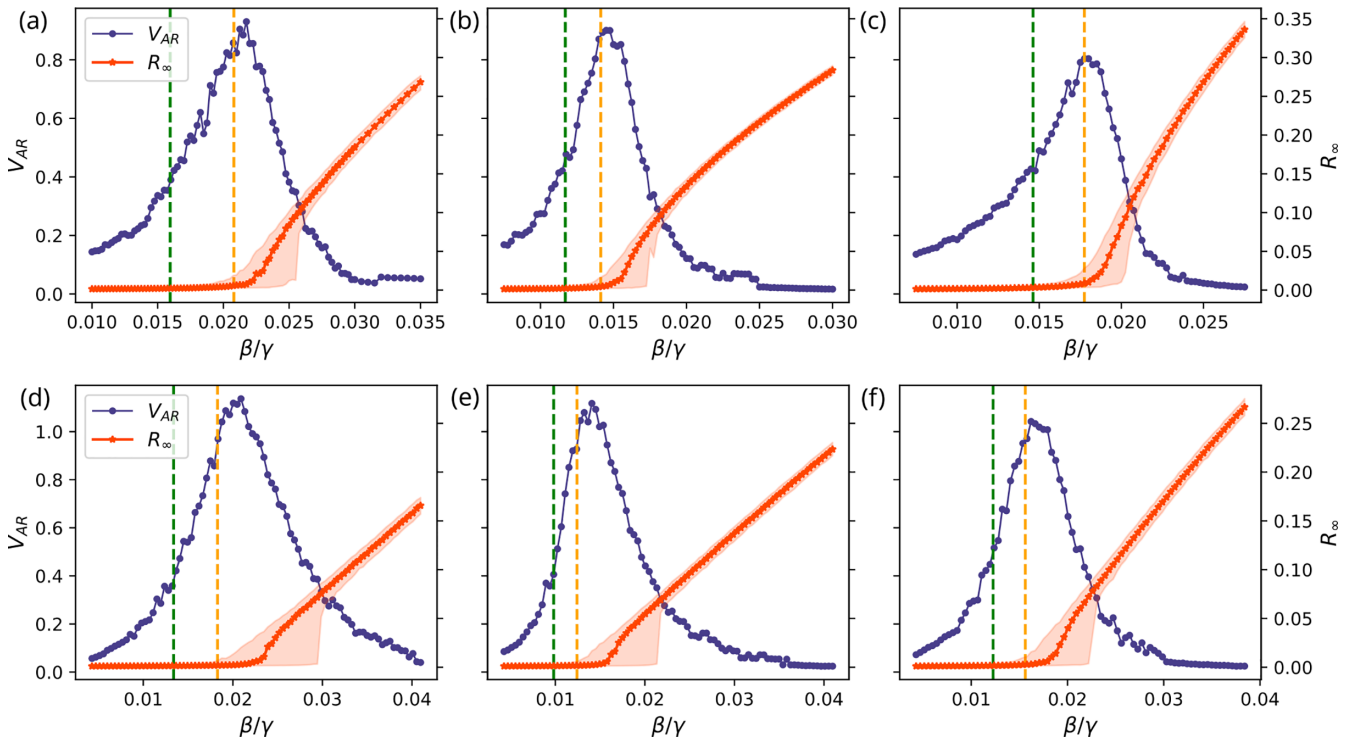


FIG. 6. Epidemic threshold estimates for different configuration settings. Yellow dashed lines: analytical threshold, computed numerically; green dashed lines: approximation of the analytical threshold, as given by (4). The R_∞ curves report the median value, with shaded areas showing the central 80% of simulated outcomes. Results are obtained from 500 simulations per setting, with networks of size $N = 5 \times 10^4$ and 50 seed nodes chosen uniformly at random. Results in the top panels (a)–(c) refer to a homogeneous degree network, with $h = 0.5$, $\langle k \rangle = 20$. The bottom panels (d)–(f), instead, refer to Pareto distributed degrees with exponent $a = 2$, with $h = 0.5$, $\langle k \rangle = 12$. Each column explores a different combination of α_I, α_S, and p_H: $\alpha_I = 2$, $\alpha_S = 2$, $p_H = 0.2$ in panels (a), (d); $\alpha_I = 2$, $\alpha_S = 3$, $p_H = 0.2$ in panels (b), (e); $\alpha_I = 2$, $\alpha_S = 2$, $p_H = 0.4$ in panels (c), (f).

The system admits a solution in $[0, 1) \times [0, 1)$ if and only if $\dot{f}_L^{-1}(1) > \dot{f}_H(1)$ (with $\dot{f}_L^{-1}(1) = +\infty$ if $\dot{f}_L(1) = 0$). In fact,

(1) On the one hand, $\dot{f}_L^{-1}(1) > \dot{f}_H(1)$ implies that some $\hat{x}_L$ sufficiently close to 1 exists such that $f_H(\hat{x}_L) > f_L^{-1}(\hat{x}_L)$. Now, let $l(x_L) = f_H(x_L) - f_L^{-1}(x_L)$. Since $l(0) < 0$ and $l(\hat{x}_L) > 0$, for the IVT there exists $\tilde{x}_L \in (0, \hat{x}_L)$ such that $l(\tilde{x}_L) = 0$, meaning that f_H and f_L^{-1} meet at $\tilde{x}_L < 1$.

(2) On the other hand, if f_H and f_L^{-1} meet at $\tilde{x}_L < 1$, for the mean value theorem there must exist $x_L', x_L'' \in [\tilde{x}_L, 1]$ such that $\dot{f}_H(x_L') = \dot{f}_L^{-1}(x_L'')$. Since $\dot{f}_H$ is nonincreasing ($\ddot{f}_H \leqslant 0$) and $\dot{f}_L^{-1}$ is nondecreasing ($\ddot{f}_L^{-1} \geqslant 0$), this implies $\dot{f}_H(1) < \dot{f}_L^{-1}(1)$.

In summary, recalling that $\dot{f}_L^{-1}(1) = \frac{1}{\dot{f}_L(1)}$, we expect to observe a giant percolation cluster if any of the following conditions holds:

(1) $\dot{f}_L(1) < 0$,
(2) $\dot{f}_H(1) < 0$,
(3) $\dot{f}_L(1) \geqslant 0$, $\dot{f}_H(1) \geqslant 0$, and $\dot{f}_L(1)\dot{f}_H(1) < 1$.

It is easy to check that the last condition is the most restrictive, so the epidemic threshold is determined by the solution of

$$\dot{f}_L(1)\dot{f}_H(1) - 1 = 0,$$

which can be rewritten as

$$(\pi_{LH}\pi_{HL}\phi_{LH}\phi_{HL} - \pi_{HH}\pi_{LL}\phi_{HH}\phi_{LL})\dot{g}_1^2(1) + (\pi_{LL}\phi_{LL} + \pi_{HH}\phi_{HH})\dot{g}_1(1) - 1 = 0. \tag{A4}$$

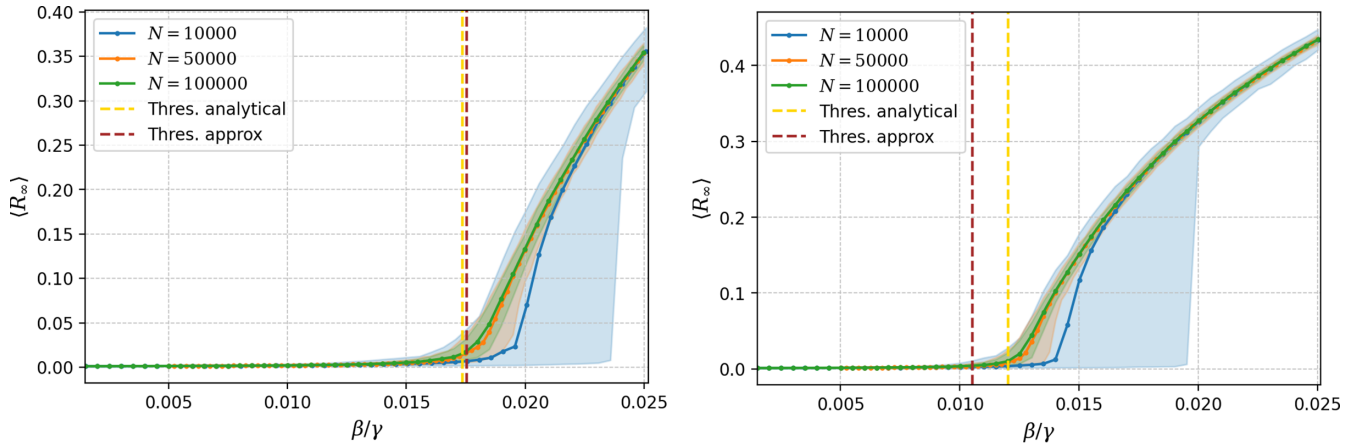
If we assume $\beta/\gamma \ll 1$, we can linearize (A4) with respect to β/γ and ignore the first term of the equation. Denoting $\alpha = \alpha_I\alpha_S$ and $\theta(p, h, \alpha) = \pi_{LL} + \alpha\pi_{HH} = (1 - h)(1 - p + \alpha p) + h(\alpha + 1)$, we obtain

$$\left(\frac{\beta}{\gamma}\right)_c = \frac{1}{(\pi_{LL} + \alpha\pi_{HH})\dot{g}_1(1)} = \frac{\langle k \rangle}{\langle k^2 \rangle - \langle k \rangle}\frac{1}{\theta(p, h, \alpha)}, \tag{A5}$$

where the last equality follows from $\dot{g}_1(1) = \frac{\langle k^2 \rangle - \langle k \rangle}{\langle k \rangle}$.

APPENDIX B: COMPARISON OF EPIDEMIC THRESHOLDS WITH SIMULATIONS

In Fig. 6 we present comparisons between numerical estimates of the epidemic threshold (using V_{AR}) and the theoretical predictions [Eqs. (3) and (4) of the main text] across different combinations of α_S, α_I, p_H, and $\langle k \rangle$. We further examine simulations for varying population sizes under homogeneous degree distributions (Fig. 7) in order to assess finite-size effects. Taken together, these analyses indicate that the findings remain consistent across all parameter configurations and population sizes examined.



(a) $h = 0$ (b) $h = 0.5$

FIG. 7. R_∞ curves for different system sizes, for networks with homogeneous degree distribution ($\langle k \rangle = 20$). The lines show the median value over 400 simulations, and the shaded areas show the central 80% of simulated outcomes. In all simulations, $I(0) = 10^{-3}$ and the seed nodes are chosen uniformly at random. Yellow dashed lines: analytical threshold, computed numerically; brown dashed lines: approximation of the analytical threshold, as given by (4). Other parameters set as $\alpha_I = 2$, $\alpha_S = 3$, $p_H = 0.4$.

[1] R. Pastor-Satorras, C. Castellano, P. Van Mieghem, and A. Vespignani, Epidemic processes in complex networks, Rev. Mod. Phys. **87**, 925 (2015).

[2] M. E. Newman, The structure and function of complex networks, SIAM Rev. **45**, 167 (2003).

[3] A. Barrat, M. Barthelemy, and A. Vespignani, *Dynamical Processes on Complex Networks* (Cambridge University Press, Cambridge, 2008)

[4] M. Newman, *Networks*, 2nd ed. (Oxford University Press, Oxford, 2018).

[5] R. Pastor-Satorras and A. Vespignani, Epidemic spreading in scale-free networks, Phys. Rev. Lett. **86**, 3200 (2001).

[6] M. E. J. Newman, Spread of epidemic disease on networks, Phys. Rev. E **66**, 016128 (2002).

[7] E. Volz, SIR dynamics in random networks with heterogeneous connectivity, J. Math. Biol. **56**, 293 (2008).

[8] J. O. Lloyd-Smith, S. J. Schreiber, P. E. Kopp, and W. M. Getz, Superspreading and the effect of individual variation on disease emergence, Nature (London) **438**, 355 (2005).

[9] W. Cota, D. Soriano-Paños, A. Arenas, S. C. Ferreira, and J. Gómez-Gardeñes, Infectious disease dynamics in metapopulations with heterogeneous transmission and recurrent mobility, New J. Phys. **23**, 073019 (2021).

[10] G. F. de Arruda, G. Petri, F. A. Rodrigues, and Y. Moreno, Impact of the distribution of recovery rates on disease spreading in complex networks, Phys. Rev. Res. **2**, 013046 (2020).

[11] Z. Feng and H. R. Thieme, Endemic models with arbitrarily distributed periods of infection I: Fundamental properties of the model, SIAM J. Appl. Math. **61**, 803 (2000).

[12] D. Clancy, Sir epidemic models with general infectious period distribution, Stat. Probab. Lett. **85** (*Feb*), 1 (2014).

[13] W. Gou and Z. Jin, How heterogeneous susceptibility and recovery rates affect the spread of epidemics on networks, Infect. Dis. Modell. **2**, 353 (2017).

[14] P. Rodrigues, A. Margheri, C. Rebelo, and M. G. M. Gomes, Heterogeneity in susceptibility to infection can explain high reinfection rates, J. Theor. Biol. **259**, 280 (2009).

[15] F. M. Neri, F. J. Pérez-Reche, S. N. Taraskin, and C. A. Gilligan, Heterogeneity in susceptible–infected–removed (SIR) epidemics on lattices, J. R. Soc. Interface. **8**, 201 (2011).

[16] R.-X. Ming, J. Liu, W. K. Cheung, and X. Wan, Stochastic modelling of infectious diseases for heterogeneous populations, Infect. Dis. Poverty **5**, 107 (2016).

[17] O. Krylova and D. J. D. Earn, Effects of the infectious period distribution on predicted transitions in childhood disease dynamics, J. R. Soc. Interface. **10**, 20130098 (2013).

[18] S. Bonaccorsi and S. Ottaviano, Epidemics on networks with heterogeneous population and stochastic infection rates, Math. Biosci. **279**, 43 (2016).

[19] A. Montalbán, R. M. Corder, and M. G. M. Gomes, Herd immunity under individual variation and reinfection, J. Math. Biol. **85**, 2 (2022).

[20] A. Darbon, D. Colombi, E. Valdano, L. Savini, A. Giovannini, and V. Colizza, Disease persistence on temporal contact networks accounting for heterogeneous infectious periods, R. Soc. Open Sci. **6**, 181404 (2019).

[21] A. L. Lloyd, Destabilization of epidemic models with the inclusion of realistic distributions of infectious periods, Proc. R. Soc. London, Ser. B **268**, 985 (2001).

[22] J. C. Miller, Epidemic size and probability in populations with heterogeneous infectivity and susceptibility, Phys. Rev. E **76**, 010101 (2007).

[23] G. De Meijere, V. Colizza, E. Valdano, and C. Castellano, Effect of delayed awareness and fatigue on the efficacy of self-isolation in epidemic control, Phys. Rev. E **104**, 044316 (2021).

[24] L. Di Domenico, C. E. Sabbatini, P.-Y. Boëlle, C. Poletto, P. Crépey, J. Paireau, S. Cauchemez, F. Beck, H. Noel, D. Lévy-Bruhl, and V. Colizza, Adherence and sustainability of

interventions informing optimal control against the covid-19 pandemic, Commun. Med. **1**, 57 (2021).

[25] P. Valgañón, A. F. Useche, D. Soriano-Paños, G. Ghoshal, and J. Gómez-Gardeñes, Quantifying the heterogeneous impact of lockdown policies on different socioeconomic classes during the first covid-19 wave in Colombia, Sci. Rep. **13**, 16481 (2023).

[26] P. Valgañón, U. Lería, D. Soriano-Paños, and J. Gómez-Gardeñes, Socioeconomic determinants of stay-at-home policies during the first covid-19 wave, Front. Public Health **11**, 1193100 (2023).

[27] M. R. DeVerna, F. Pierri, Y.-Y. Ahn, S. Fortunato, A. Flammini, and F. Menczer, Modeling the amplification of epidemic spread by individuals exposed to misinformation on social media, npj Complexity **2**, 11 (2025).

[28] M. T. Beca-Martínez, M. Romay-Barja, M. Falcón-Romero, C. Rodríguez-Blázquez, A. Benito-Llanes, and M. J. Forjaz, Compliance with the main preventive measures of COVID-19 in Spain: The role of knowledge, attitudes, practices, and risk perception, Transboundary Emerging Dis. **69**, e871 (2022).

[29] R. Pastor-Satorras and C. Castellano, The advantage of self-protecting interventions in mitigating epidemic circulation at the community level, Sci. Rep. **12**, 15950 (2022).

[30] P. W. Holland, K. B. Laskey, and S. Leinhardt, Stochastic blockmodels: First steps, Soc. Networks **5**, 109 (1983).

[31] C. Castellano and R. Pastor-Satorras, Thresholds for epidemic spreading in networks, Phys. Rev. Lett. **105**, 218701 (2010).

[32] F. Mazza, F. Colaiori, S. Guarino, S. Meloni, M. Brambilla, C. Piccardi, F. Pierri, and F. Saracco, The impact of heterogeneity on epidemics: Insights from a modified sir model, *International Conference on Complex Networks and Their Applications* (Springer, Berlin, 2024), pp. 66–75.

[33] B. Karrer and M. E. J. Newman, Stochastic blockmodels and community structure in networks, Phys. Rev. E **83**, 016107 (2011).

[34] M. A. Gibson and J. Bruck, Efficient exact stochastic simulation of chemical systems with many species and many channels, J. Phys. Chem. A **104**, 1876 (2000).

[35] D. Clancy, Persistence time of SIS infections in heterogeneous populations and networks, J. Math. Biol. **77**, 545 (2018).

[36] D. Clancy, Precise estimates of persistence time for SIS infections in heterogeneous populations, Bull. Math. Biol. **80**, 2871 (2018).

[37] G. Cimini, T. Squartini, F. Saracco, D. Garlaschelli, A. Gabrielli, and G. Caldarelli, The statistical physics of real-world networks, Nat. Rev. Phys. **1**(1), 58 (2019)

[38] F. Chung and L. Lu, Connected components in random graphs with given expected degree sequences, Ann. Combinatorics **6**, 125 (2002).

[39] G. Caldarelli, A. Capocci, P. De Los Rios, and M. A. Muñoz, Scale-free networks from varying vertex intrinsic fitness, Phys. Rev. Lett. **89**, 258702 (2002).

[40] F. Mazza, S. Guarino, and F. Saracco, Data for "impact of behavioral heterogeneity on epidemic outcome and its mapping into effective network topologies", Zenodo Data (2025), V1, 10.5281/zenodo.18017301.