The Impact of Heterogeneity on Epidemics: Insights from a Modified SIR Model

Fabio Mazza[1(✉)], Francesca Colaiori[2,3], Stefano Guarino[4], Sandro Meloni[4], Marco Brambilla[1], Carlo Piccardi[1], Francesco Pierri[1], and Fabio Saracco[4,5,6]

[1] Dipartimento di Elettronica, Informazione e Bioingegneria, Politecnico di Milano, Milan, Italy
fabio.mazza@polimi.it
[2] Istituto dei Sistemi Complessi, Consiglio Nazionale delle Ricerche, Roma, Italy
[3] Dipartimento di Fisica, Università di Roma La Sapienza, Roma, Italy
[4] Istituto per le Applicazioni del Calcolo "Mauro Picone", Consiglio Nazionale delle Ricerche, Roma, Italy
[5] Centro Studi e Ricerche "Enrico Fermi", Roma, Italy
[6] IMT School For Advanced Studies Lucca, Lucca, Italy

Abstract. Human behavior is a key determinant of epidemic outcomes. During health crises, variations in people's responses to control measures, often driven by different levels of risk perception, lead to variability in epidemic parameters such as infectiousness and susceptibility. We introduce a model within the Susceptible-Infected-Removed (SIR) class that accounts for these heterogeneities. We find that there is a region in the space of parameters just above the epidemic threshold, where trajectories showing an initial decline in the number of Infected can suddenly reverse and give rise to widespread transmission. Such heterogeneity can lead to an underestimation of transmission potential and delayed recognition of epidemic resurgence, thereby severely compromising efforts for a timely response. We examine this phenomenon in the mean-field scenario and then simulate the dynamics on homogeneous and heterogeneous contact networks, confirming that this phenomenology persists beyond mean field. Our model also encompasses cases where the heterogeneity originates from biological or other factors.

Keywords: complex networks · human behavior · epidemics · SIR

1 Introduction

The effect of population heterogeneity in the evolution of an epidemic is widely studied in the scientific literature. Previous work considers heterogeneity in connectivity [14], behavior [5,13], susceptibility [12] or infectivity [3,8,10,11]. The combined effect of heterogeneous susceptibility and infectivity has been studied in [7] by Miller, focusing on the outbreak size and probability for epidemic evolving on a network. Among other findings, Miller estimates the epidemic threshold

H. Cherifi et al. (Eds.): COMPLEX NETWORKS 2024, SCI 1190, pp. 66–75, 2025.
https://doi.org/10.1007/978-3-031-82439-5_6

imposing that the average transmission rate equals the reciprocal of the average excess degree of the network. In this paper, we show that, under specific conditions, the actual epidemic threshold can differ from the estimate given in [7]. In more detail, we introduce an epidemic model in the Susceptible-Infected-Removed (SIR) class, denoted HeSIR (*Heterogeneous* SIR), that incorporates heterogeneity by assuming a bimodal distribution of susceptibility and infectivity in the population. In the following we attribute heterogeneity to individual behavior.

Such a variability may more generally arise from a range of factors. For instance, it could be due to some biological characteristics of the hosts, or to the lack of compliance with the safety guidelines during a pandemic [4]. The effect of the latter was clearly observed during the unprecedented global spread of SARS-CoV2 [2,9]. Therefore, in the following we label individuals according to their behavior: Ordinary (O) for those who comply with the guidelines, and Misbehaving (M) for those who do not. Given that misbehaving individuals are at a higher risk of transmitting and contracting the disease, we assign to them values of infectiousness and susceptibility increased by a factor $a > 1$ and $b > 1$, respectively. Our approach is similar to the modeling framework introduced in [4], where the lack of compliance with preventive measures is attributed to misinformation.

We estimate the epidemic threshold in a random regular graph (RRG), finding the conditions for the existence of a giant percolation cluster when the link occupation probability corresponds to the total transmission probability for that link.

Direct integration of the mean-field equations and simulations on an RRG show the existence of a region corresponding to a "resurgence zone", where an initial fall in the size of infected population is followed by a sudden growth that leads to an outbreak. In the standard SIR model, the condition $\dot{I}(0) > 0$, where I is the fraction of the infected population, is necessary and sufficient for an outbreak to occur. We show that, in our model, the necessary and sufficient condition corresponds to $\dot{I}_{\text{eff}}(0) > 0$, where $I_{\text{eff}} = aI_M + I_O$ is the effective fraction of the infected population. Standard behavior is recovered as soon as $a = 1$ or $b = 1$ (or both), meaning that this aspect of the epidemic can only be captured by models in which susceptibility and infectivity *both* vary.

Finally, to further support our findings, we consider the evolution of our model on two simple networks characterized, respectively, by homophily and degree heterogeneity. The term homophily is generally used to refer to the tendency of "similar" nodes to be connected. In this case, by similar we intend having the same risk attitude, and, hence, same susceptibility and infectivity. Simulations show that homophilic populations (where O (M) individuals are more likely to be connected with other O (M) individuals) correspond to a wider resurgence zone. A similar phenomenon occurs in heterogeneous networks. In that case, we find that the resurgence zone persists in the limit $a = b = 1$, which corresponds to the standard SIR model, showing that topological heterogeneity plays a role similar to heterogeneity in the epidemic parameters. Our

results indicate that heterogeneity in epidemiological parameters can result in an underestimate of transmission potential, which, initially obscured, could trigger a resurgence later. This delay in recognizing the outbreak can, in turn, dramatically compromise an effective and timely epidemic response.

2 Model Description and Analysis

We fix the timescale by setting the susceptibility of ordinary individuals to $\sigma = 1$. The mean–field (MF) equations for the HeSIR model are the following:

$$\begin{cases} \dot{I}_M &= b\beta S_M(aI_M + I_O) - \gamma I_M = b\beta S_M I_{\text{eff}} - \gamma I_M \\ \dot{I}_O &= \beta S_O(aI_M + I_O) - \gamma I_O = \beta S_O I_{\text{eff}} - \gamma I_O \\ \dot{S}_M &= -b\beta S_M(aI_M + I_O) = -b\beta S_M I_{\text{eff}} \\ \dot{S}_O &= -\beta S_O(aI_M + I_O) = -\beta S_O I_{\text{eff}} \\ \dot{R} &= \gamma(I_M + I_O) \end{cases} \tag{1}$$

where I_M (resp. I_O) is the fraction of misbehaving (ordinarily behaving) infected individuals, S_M (S_O) is the fraction of misbehaving (ordinary) susceptible individuals, and R is the fraction of recovered. The total population is conserved, so that $I_M + I_O + S_M + S_O + R = 1$.

To investigate the behaviour of the HeSIR system in the discrete individual-based case, we start from the case where the contacts between the N individuals in the system can be described by a random regular graph (RRG). We resort to the classic equivalence with a percolation problem and find the conditions that lead to the emergence of a giant percolation cluster. Let p be the fraction of misbehaving population and k the number of edges of each node in the graph, and let x_O and x_M be respectively the average probability that an ordinary vertex and a misbehaving vertex are not connected to the giant percolation cluster via a specific one of their edges. Further, for $i, j \in \{O, M\}$, let ϕ_{ij} be the occupation probability for an edge connecting a susceptible vertex of type i with an infected vertex of type j. x_O and x_M satisfy the system

$$\begin{cases} x_O &= (1-p)(1 - \phi_{OO} + \phi_{OO} x_O^{k-1}) + p(1 - \phi_{OM} + \phi_{OM} x_M^{k-1}) \\ x_M &= (1-p)(1 - \phi_{MO} + \phi_{MO} x_O^{k-1}) + p(1 - \phi_{MM} + \phi_{MM} x_M^{k-1}) \end{cases} \tag{2}$$

In fact: the considered edge leads to an ordinary vertex with probability $1-p$, to a misbehaving one with probability p; the node at the other end of the considered edge has other $k-1$ links.

The system admits a trivial solution for $x_O = x_M = 1$, corresponding to the absence of any giant cluster. The epidemic threshold can thus be found imposing that the two equations in (2) define curves being tangent in $(1, 1)$. To do so, we define

$$f(x_O, x_M) = x_O - (1-p)(1 - \phi_{OO} + \phi_{OO} x_O^{k-1}) - p(1 - \phi_{OM} + \phi_{OM} x_M^{k-1}) \tag{3}$$

$$g(x_O, x_M) = x_M - (1-p)(1 - \phi_{MO} + \phi_{MO} x_O^{k-1}) - p(1 - \phi_{MM} + \phi_{MM} x_M^{k-1}) \tag{4}$$

and we impose $\nabla f(1,1) = \nabla g(1,1)$. Since the occupation probability ϕ_{ij} is the total transmission probability for that edge, we use the approximation $\phi_{OO} \approx \beta/\gamma$, $\phi_{OM} \approx a\beta/\gamma$, $\phi_{MO} \approx b\beta/\gamma$ and $\phi_{MM} \approx ab\beta/\gamma$. Finally, we find

$$\left(\frac{\beta}{\gamma}\right)_c = \frac{1}{(k-1)(abp+1-p)} \tag{5}$$

If, instead, we let x be the average probability that *any* random vertex is connected to the giant percolation cluster via a specific one of its edges, the system (2) is replaced by the standard equation $x = 1 - \phi + \phi x^{k-1}$, where the average occupation probability is

$$\phi = (1-p)^2\phi_{OO} + p(1-p)\phi_{OM} + p(1-p)\phi_{MO} + p^2\phi_{MM} \tag{6}$$

In this case, the condition for the existence of a giant percolation cluster is estimated as

$$\left(\frac{\beta}{\gamma}\right)_* = \frac{1}{(k-1)(ap+1-p)(bp+1-p)} \tag{7}$$

which, for k-regular graphs, coincides with the estimate found in [7].

While (5) is the actual epidemic threshold, the epidemic behavior is described by both critical thresholds $(\beta/\gamma)_c$ and $(\beta/\gamma)_*$: for $\beta/\gamma < (\beta/\gamma)_c$ outbreaks are suppressed; for $\beta/\gamma > (\beta/\gamma)_*$ the contagion grows exponentially right from the start; in the intermediate region where $(\beta/\gamma)_c < \beta/\gamma < (\beta/\gamma)_*$ there exists a deceptive phase in which an initial decay is then reverted and followed by an outbreak. This can be proved assuming $R(0) = 0$ and that a small uniformly distributed fraction ϵ of the population is initially infected, so that we have $I_M(0) = p\epsilon$, $I_O(0) = (1-p)\epsilon$, $S_M(0) = p(1-\epsilon)$, $S_O(0) = (1-p)(1-\epsilon)$. In this case, we have

$$\dot{I}(0) = 0 \iff \frac{\beta}{\gamma} = \frac{1}{(ap+1-p)(bp+1-p)(1-\epsilon)} \tag{8}$$

$$\dot{I}_{\text{eff}}(0) = 0 \iff \frac{\beta}{\gamma} = \frac{1}{(abp+1-p)(1-\epsilon)} \tag{9}$$

If ϵ is small, (8) is essentially equivalent to (7), whereas (9) is essentially equivalent to (5). Condition (8) is generally used to separate the two epidemic regimes, the inactive and the active phase. In the HeSIR model, however, an initial decrease in I_O can be reversed once the more active M population becomes involved in the outbreak.

For this reason, (9) determines, up to a factor $1/(1-\epsilon)$, the actual epidemic threshold $(\beta/\gamma)_c$, as confirmed by numerical integration results (Fig. 1). When $p = 0$ (no misbehaving) or $a = b = 1$ (no increase in misbehaving infectivity and susceptibility) the usual SIR dynamics and epidemic threshold are recovered, with $(\beta/\gamma)_c = 1$. When $p > 0$ and $a = 1$ or $b = 1$ (but not both), (8) and (9) are equivalent and $(\beta/\gamma)_c < 1$. Only when $p > 0$ and both $a > 1$ and $b > 1$, (9) sets a weaker and more accurate condition than (8).

Both the numeric solution of system (1) and extensive simulations on a RRG confirm the above described scenario and the existence of three regimes. For the simulations we generated one RRG with $N = 40000$ nodes and degree $k = 800$, and we used an optimized variant of the Gillespie algorithm [6]. The value of k was chosen to be sufficiently large in order to verify the results of the mean field model, and to avoid a saturation effect which could lead to more complex behavior. To empirically distinguish these regimes, we can focus on the number of stationary points of $I(t)$, which are 0 when $\beta/\gamma < (\beta/\gamma)_c$, 2 when $(\beta/\gamma)_c < \beta/\gamma < (\beta/\gamma)_*$, and 1 when $\beta/\gamma > (\beta/\gamma)_*$. Figures 2 and 3 outline the regions for the three regimes, separated by (8) and (9), in the mean field model and the RRG, respectively.

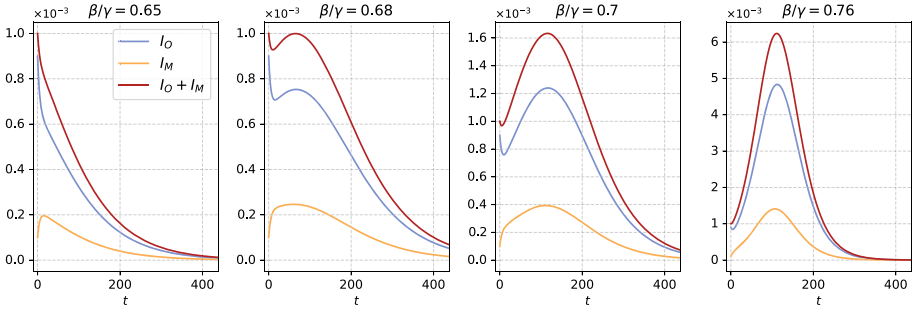

Fig. 1. Fraction of infected individuals in the system, from the integration of the mean field Eqs. (1), for different values of β/γ. The other parameters are set to $\gamma = 0.2$, $a = 2$, $b = 3$ and $p = 0.1$. When β/γ does not satisfy (9), I decreases monotonically to 0 (leftmost plot). On the other hand, when β/γ satisfies (8), I increases right from the start (rightmost plot). Whenever β/γ satisfies (9) but not (8), I initially decreases to then increase and lead to an outbreak (second and third plot).

3 Model Evolution in Heterogeneous Networks

We now investigate whether the phenomenology observed in mean field and on RRG persists on heterogeneous networks. First, we consider the case in which the network is assortative with respect to the vertex type, meaning that the individuals have a preference for connecting with other people with similar behavior (ordinary with ordinary, misbehaving with misbehaving). The level of *homophily* can be controlled by a parameter $h \in [0, 1]$, with the mixing between the two groups defined by the following matrix:

$$(1-h)\begin{pmatrix} 1-p & p \\ 1-p & p \end{pmatrix} + h\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \tag{10}$$

In words, an ordinary (O) vertex connects with another O with probability $(1-p)(1-h)+h$, and with a misbehaving (M) with probability $p(1-h)$; and an

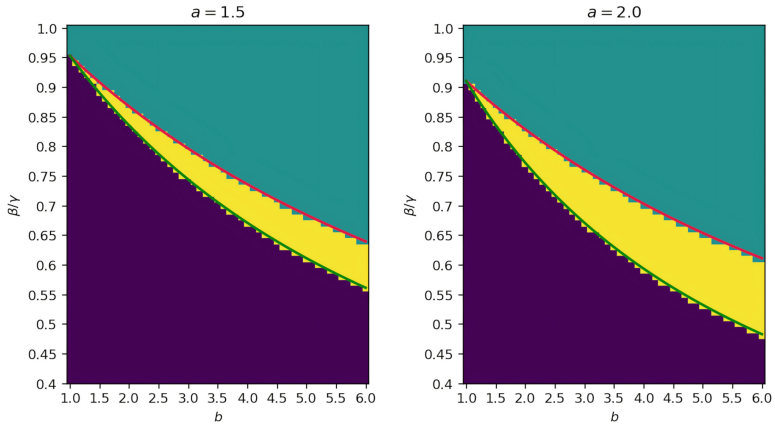

Fig. 2. The number of stationary points (s.p.) in $I(t)$ from numerical integration of the mean field model with different values of β/γ and b (purple: 0 s.p., cyan: 1 s.p., yellow: 2 s.p.). The red curve corresponds to Eq. (7), the green one to Eq. (5). In the yellow region, there is a resurgence of the global fraction of infected population. The initial infected population is $\epsilon = 10^{-3}$, divided among O and M proportionally to the fraction of misbehaving $p = 0.1$, and $\gamma = 0.25$. (Color figure online)

M vertex connects with an O with probability $(1-p)(1-h)$, and with another M with probability $p(1-h)+h$. When $h = 0$ a standard network with no homophily is obtained, while $h = 1$ means that ordinary and misbehaving nodes form two separate connected components, each with a Poisson-like degree distribution as the whole network.

We perform numerical simulations of the HeSIR model on Stochastic Block Model (SBM) graphs of $N = 40000$ individuals, constructed with the mixing matrix above (Eq. 10) with average degree $k = 400$. From the results shown in Figs. 4 and 5, it is clear that when increasing network homophily the region of epidemic resurgence (when 2 stationary points are found in $I(t)$) becomes larger, covering a broader range of β/γ values. In the same figures, we also show the empirical estimate for the epidemic threshold obtained maximizing the epidemic variability V_{AR}, defined as

$$V_{\mathrm{AR}} = \frac{\langle R_\infty^2 \rangle - \langle R_\infty \rangle^2}{\langle R_\infty \rangle}$$

where R_∞ is the attack rate of the epidemic simulations. In both Figs. 4 and 5, this empirical threshold corresponds with the lower boundary of the resurgence zone.

Finally, in order to investigate the effect of the heterogeneity of the degree distribution, we also explore the HeSIR model on Barabàsi-Albert (BA) random graphs. The simulations have been run on a single graph of $N = 40000$ and with each node adding $k = 4$ links with preferential attachment [1]. Figure 6 shows that the resurgence zone is much larger then in the previous cases. As done for

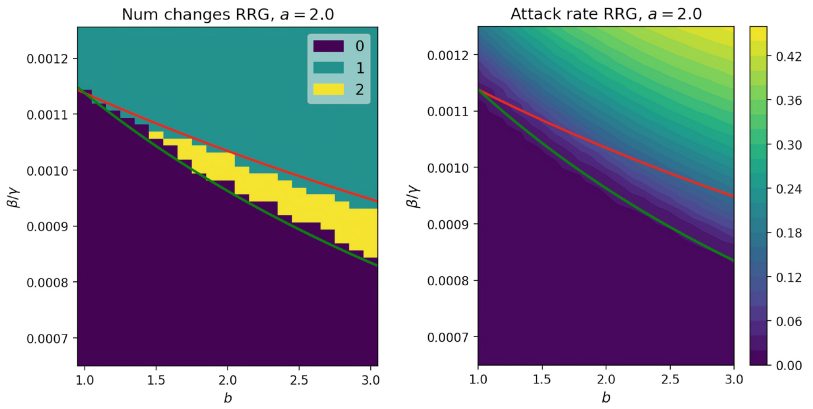

Fig. 3. Results for simulations on RRG graph. Left: number of stationary points of $I(t)$. In the yellow intermediate region $I(t)$ starts decreasing at $t = 0$, then later it increases. Right: Median attack rate (R_∞). The red line corresponds to Eq. (7), the green one to Eq. (5), rescaled by a factor $d - 1$. Other parameters are $p = 0.1$, $\gamma = 0.2$.

the SBM, we also show the empirical estimate obtained maximizing the epidemic variability V_{AR}. Again, the empirical epidemic threshold corresponds with the lower boundary of the resurgence zone.

The phenomenon of epidemic resurgence described above can also be found on BA graphs under the SIR model. Its presence in this case can be explained by the heterogeneity of the degree: since most nodes have very low degree, it

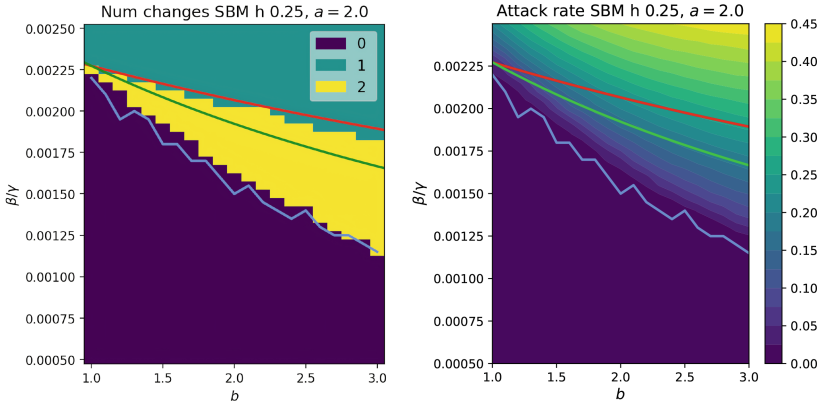

Fig. 4. Results for SBM with $h = 0.25$. Left: number of stationary points of $I(t)$. Right: Median attack rate . The red line corresponds to Eq. (7), the green one to Eq. (5), both rescaled by the average degree of the graph. The light blue lines show the β/γ values corresponding to the peak variability of the attack rate V_{AR}. Other parameters are $p = 0.1$, $\gamma = 0.2$.

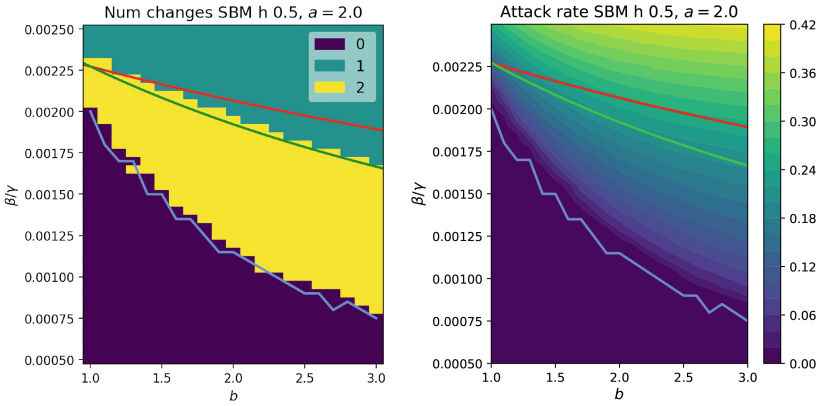

Fig. 5. Results for SBM with $h = 0.5$. Left: number of stationary points of $I(t)$. Right: Median attack rate (R_∞). The red line corresponds to Eq. (7), the green one to Eq. (5), both rescaled by the average degree of the graph. The light blue lines show the β/γ values corresponding to the peak variability of the attack rate V_{AR}. Other parameters are $p = 0.1$, $\gamma = 0.2$.

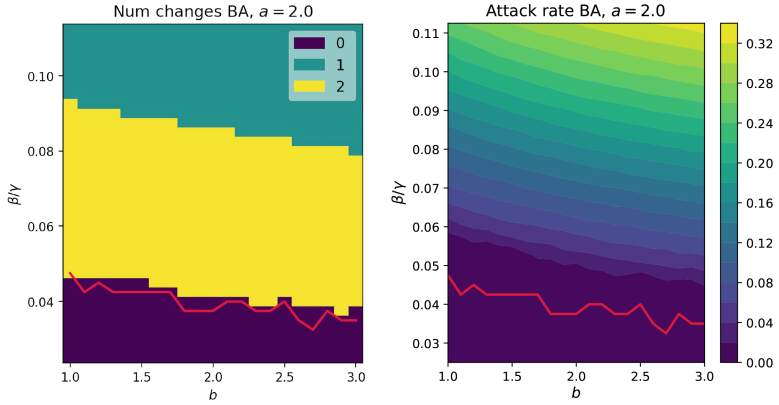

Fig. 6. Results for the BA graph. Left: number of stationary points of $I(t)$. Right: Median attack rate (R_∞). The red lines show the β/γ values corresponding to the peak variability of the attack rate V_{AR}. Other parameters are $p = 0.1$, $\gamma = 0.2$.

is likely that by choosing the epidemic seeds at random, most of them cannot infect many other nodes, and only when a sufficient number of nodes with higher degree are infected, then an epidemic outbreak is possible (see Fig. 7).

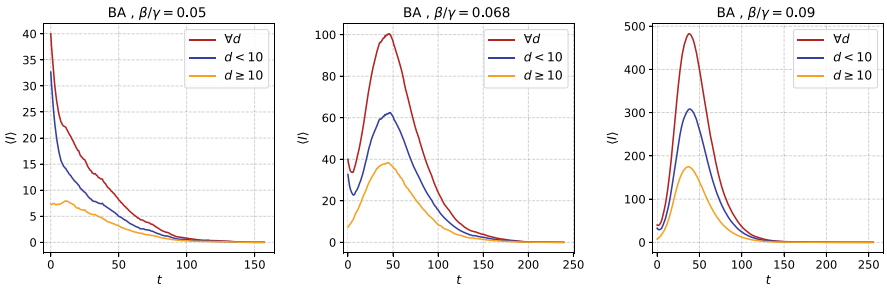

Fig. 7. Average total number of infecteds for $a = 1.0$ and $b = 1.0$ (i.e., ordinary SIR model) on the BA graph, with the average done on a) all nodes ($\forall d$), b) only nodes with degree $d < 10$, c) only nodes with $d \geq 10$. Other parameters are $p = 0.1$, $\gamma = 0.2$.

4 Conclusion

Our study highlights the critical impact of heterogeneity in susceptibility and infectivity on epidemic dynamics, particularly in relation to the epidemic threshold and the emergence of a resurgence zone. By extending the standard SIR model to incorporate a bimodal distribution of these parameters, we demonstrate that the necessary and sufficient conditions for an outbreak shift from a simple consideration of the infected population to the effective infected population, I_{eff}. We remark that the HeSIR model, defined with human behavior in mind, also describes cases where the heterogeneity in transmission originates from biological or other factors.

Our findings reveal that, at least in homogeneous networks, both susceptibility and infectivity must vary for this more complex behavior to emerge, with homophilic networks further widening the resurgence zone. This underscores the importance of considering population heterogeneity when modeling epidemics, as it plays a pivotal role in shaping epidemic trajectories. Disregarding this factor could lead to an underestimation of transmission potential, compromising the intervention in a mitigation scenario. In such a scenario, it would be important to control infections among the non-compliant population, which is the driving force behind the resurgence phenomenon in our findings. This could be accomplished through targeted testing, or by encouraging non-compliant individuals to adhere to safety guidelines. On the other hand, heterogenous structures create the conditions for the resurgence zone to emerge even for the standard SIR model. This result indicates that nodes with highly heterogeneous degree should also be closely monitored.

The HeSIR model presented in this work could be extended in the future to include a higher diversity in susceptibility and infectivity. We also leave to future work the comparison of the outcomes of our model against collected data on epidemic trajectories and the further investigation of the effect of degree heterogeneity.

Acknowledgements. 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 P2022AKRZ9, CUP B53D23026080001.

References

1. Albert, R., Barabási, A.L.: Statistical mechanics of complex networks. Rev. Mod. Phys. **74**(1), 47–97 (2002). https://doi.org/10.1103/RevModPhys.74.47. https://link.aps.org/doi/10.1103/RevModPhys.74.47
2. Beca-Martínez, M.T., et al.: Compliance with the main preventive measures of COVID-19 in Spa in: the role of knowledge, attitudes, practices, and risk perception. Transbound. Emerg. Dis. **69**(4), e871–e882 (2022). https://doi.org/10.1111/tbed.14364. https://onlinelibrary.wiley.com/doi/abs/10.1111/tbed.14364
3. Bonaccorsi, S., Ottaviano, S.: Epidemics on networks with heterogeneous population and stochastic infection rates. Math. Biosci. **279**, 43–52 (2016)
4. DeVerna, M.R., Pierri, F., Ahn, Y.Y., Fortunato, S., Flammini, A., Menczer, F.: Modeling the amplification of epidemic spread by misinformed populations. arXiv preprint arXiv:2402.11351 (2024)
5. Funk, S., Salathé, M., Jansen, V.A.: Modelling the influence of human behaviour on the spread of infectious diseases: a review. J. R. Soc. Interface **7**(50), 1247–1256 (2010)
6. Gibson, M.A., Bruck, J.: Efficient exact stochastic simulation of chemical systems with many species and many channels. J. Phys. Chem. A **104**(9), 1876–1889 (2000). https://doi.org/10.1021/jp993732q
7. Miller, J.C.: Epidemic size and probability in populations with heterogeneous infectivity and susceptibility. Phys. Rev. E, Stat. Nonlinear Soft Matter Phys. **76**(1), 010101 (2007). https://doi.org/10.1103/PhysRevE.76.010101
8. Ming, R.X., Liu, J., Cheung, W.K., Wan, X.: Stochastic modelling of infectious diseases for heterogeneous populations. Infect. Dis. Poverty **5**, 1–11 (2016)
9. Mitze, T., Kosfeld, R., Rode, J., Wälde, K.: Face masks considerably reduce COVID-19 cases in Germany. Proc. Natl. Acad. Sci. **117**(51), 32,293–32,301 (2020). https://doi.org/10.1073/pnas.2015954117. https://www.pnas.org/doi/full/10.1073/pnas.2015954117
10. Neri, F.M., Pérez-Reche, F.J., Taraskin, S.N., Gilligan, C.A.: Heterogeneity in susceptible-infected-removed (SIR) epidemics on lattices. J. R. Soc. Interface **8**(55), 201–209 (2011)
11. Qu, B., Wang, H.: SIS epidemic spreading with heterogeneous infection rates. IEEE Trans. Netw. Sci. Eng. **4**(3), 177–186 (2017)
12. Rodrigues, P., Margheri, A., Rebelo, C., Gomes, M.G.M.: Heterogeneity in susceptibility to infection can explain high reinfection rates. J. Theor. Biol. **259**(2), 280–290 (2009)
13. Vilone, D., Vriens, E., Andrighetto, G.: The effect of heterogeneous distributions of social norms on the spread of infectious diseases. J. Phys. Complexity **5**(2), 025012 (2024)
14. Volz, E.: SIR dynamics in random networks with heterogeneous connectivity. J. Math. Biol. **56**, 293–310 (2008)