Optimizing the mitigation of epidemic spreading through targeted adoption of contact tracing apps

Aleix Bassolas,[1,2,*] Andrea Santoro,[1,3,*] Sandro Sousa,[1,†] Silvia Rognone,[1] and Vincenzo Nicosia[1,‡]

[1]*School of Mathematical Sciences, Queen Mary University of London, London E1 4NS, United Kingdom*
[2]*Departament d'Enginyeria Informàtica i Matemàtiques, Universitat Rovira i Virgili, 43007 Tarragona, Spain*
[3]*Institute of Bioengineering, Center for Neuroprosthetics, EPFL, 1202 Geneva, Switzerland*

(Received 8 July 2021; accepted 24 March 2022; published 2 May 2022)

The ongoing COVID-19 pandemic is the first epidemic in human history in which digital contact tracing has been deployed at a global scale. Tracking and quarantining all the contacts of individuals who test positive for a virus can help slow down an epidemic, but the impact of contact tracing is severely limited by the generally low adoption of contact-tracing apps in the population. We derive here an analytical expression for the effectiveness of contact-tracing app installation strategies in a susceptible-infected-recovered (SIR) model on a given contact graph. We propose a decentralized heuristic to improve the effectiveness of contact tracing under fixed adoption rates, which targets a set of individuals to install contact-tracing apps and can be easily implemented. Simulations on a large number of real-world contact networks confirm that this heuristic represents a feasible alternative to the current state of the art.

DOI: 10.1103/PhysRevResearch.4.023092

I. INTRODUCTION

Since the first human infection toward the end of 2019, the spread of the SARS-COV-2 virus has caused an unprecedented shock around the world with serious repercussions in all aspects of our social and economic activities [1–3] and a number of casualties that has already passed the two million figure and is unfortunately due to rise further in the near future [4]. The initial efforts to curb the spread of the disease focused on nonpharmaceutical interventions including travel bans, lockdowns, and curfews [5]. These measures are able to drastically reduce the opportunities of contact between infected and susceptible people and thus the spread of a virus [6–10] but also have non-negligible effects on economy and social life [11–13]. After the first wave of infections in February–May 2020 and thanks to a better understanding of the specific transmission dynamics of SARS-COV-2 [14–19], many countries have implemented some sort of "test-trace-treat" system based on digital contact tracing (CT) [20,21]. Some of these systems consist of deploying CT apps on mobile phones which allow to identify and isolate individuals who have been in contact with infected ones, thus disrupting secondary infections paths as early as possible. With CT in place, many countries have been able to partially reopen several sectors of their economy and to diminish the damage of prolonged disruptions [22–29].

An effective digital CT strategy should aim at maximizing the probability of detecting contacts between infected and susceptible individuals, and it would completely eradicate contagion in the ideal case where CT apps are installed by the totality of a population [30–39]. However, throughout the SARS-COV-2 pandemic, the percentage of the population with CT apps installed has remained quite low, between 5 and 20% in most countries [40], resulting in a dramatically decreased efficiency of contact tracing.

Here we focus on the problem of determining the set of nodes which should install CT apps in order to optimize the effect of contact tracing, i.e., to maximally slow down spreading and reduce the incidence of a disease under the assumption that the rate of CT app adoption is fixed. We provide an analytic derivation to quantify the decrease of the basic reproduction number caused by a generic CT installation strategy, and we show that uniform random installation—which is the strategy implicitly adopted by governments when people are simply asked to install a CT app—has the worst performance of all. We find that relatively simple targeting strategies based on the structure of the contact network are significantly more efficient in reducing the number of secondary infections at low adoption rates in both synthetic and real-world systems.

II. RESULTS

In Fig. 1 we report a sketch of a fictitious contact network, where some individuals are infected (pink), some other are susceptible (black), and some have a CT app installed (indicated by the mobile icon). An almost perfect lockdown as the one enforced in many countries during the spring of 2020, in which only essential mobility is allowed (i.e., with the only exception of individuals sharing the same household or related to basic services), would remove most of the links in that graph. In this specific case, the infected individuals will eventually be unable to find any susceptible person to pass

*These authors contributed equally to this paper.
†Present address: NEtwoRks, Data, and Society (NERDS), IT University of Copenhagen, 2300 Copenhagen, Denmark.
‡v.nicosia@qmul.ac.uk.

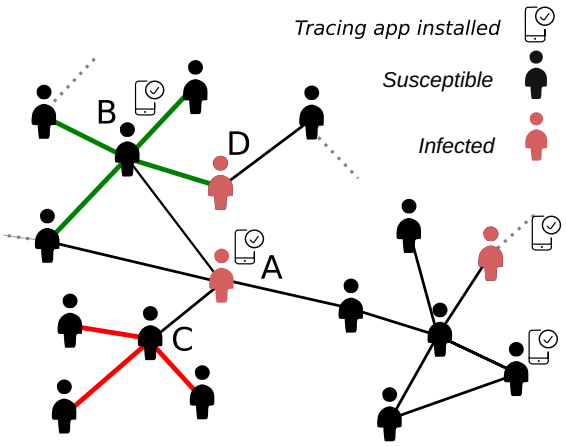

FIG. 1. Effect of contact-tracing apps on secondary infections on a contact network. CT apps can only detect potential contagious contacts if both the infected and the susceptible individual have a CT app installed. In this case, A can in principle infect any of his direct neighbors, including B who has a CT app installed as well, and the CT app cannot do anything to avoid this. However, if the CT system detects a contact between A and B and then A tests positive, then B can be contacted and put into quarantine, thus disrupting all the potential infection paths to the direct contacts of B (solid green lines). If C has a contact with A instead, C has no way of knowing whether the contact resulted in an infection or not and will not be notified when A tests positive, and C will not go into quarantine. In this case, all of C's neighbors are at risk of catching the disease (solid red lines). Similarly, the fact that B has a CT app installed cannot be a safeguard from being infected by D (who does not have a CT app installed) and passing the infection to neighbors while having no symptoms.

the disease on to. When only CT is in place, instead, some infections are still unavoidable either due to a limited app adoption rate or to a delay in the notification of test results and in isolation of subjects exposed to those who are infected. As made evident by Fig. 1, maximizing the impact of contact tracing corresponds to maximizing the probability that the potential transmission of the disease between two individuals is detected, since only contacts among people with CT apps installed can be detected and traced back. This intuitively corresponds to maximizing the number of edges among the individuals with CT app installed, i.e., the density of the subgraph induced by the nodes with CT apps under the constraint that only a fraction r of the population will have the CT app installed.

A. Reduction of R_0 in a SIR + CT dynamics

We consider here a susceptible-infected-recovered (SIR) + CT model, that is a classical SIR model on a static contact graph [10], with the addition of ideal CT. This means that any susceptible node with an installed CT app is quarantined (recovered) as soon as one of their contacts with a CT app installed gets infected. As a consequence, our stylized model is mainly designed to compare the effectiveness of different CT strategies in a simple setting rather than capturing precisely the complexity of real-world scenarios. Indeed, depending on the type of disease we aim to study, several key factors may and should be taken into account in the model, including transmissibility (which might depend on the days passed since the infection), intrinsic characteristics of the carrier, or the timing of infection. Moreover, certain diseases might be transmitted by an infected individual before showing any kind of symptoms (like in asymptomatic or preasymptomatic cases), while others are only transmitted once symptoms appear [41–44].

The parameters of the SIR + CT model are the probability β that an infected individual passes the disease to each of their susceptible neighbors and the probability μ that an infected individual is removed (due to either recovery or death). That is, at each time step, infected nodes can infect any neighbors with probability β, while infected nodes recover with probability μ. When a node i equipped with a CT app gets infected, then all neighbors j having a CT app installed get instantaneously recovered. This dynamic mimics an ideal test-trace-isolate system where nodes equipped with CT apps isolate immediately on receiving the notification of a contact with a positive individual.

We call the contact graph $G(V, E)$, with $N = |V|$ nodes and $K = |E|$ edges, and we denote by $G'(V', E')$ the subgraph of G induced by CT app installations, i.e., such that V' is the set of nodes in G with CT apps and E' is the set of edges among nodes in V'. We quantify the effect of the installation of CT apps in a certain subset V' of nodes by computing the reduction of the basic reproduction number R_0, that is, the expected number of secondary infections caused by a single contagion event. Let us assume that the generic node ℓ is infected and has passed the disease to neighbor i. The expected number R_i of secondary infections caused by i while it remains infected depends on whether i is in V' and on how many of its k_i neighbors are in V' as well. In particular, if $i \notin V'$, $R_i = \frac{\beta}{\mu}(k_i - 1)$ as in the classical SIR (we have to remove ℓ from the count, hence the $k_i - 1$) [10]. If $i \in V'$, instead there are two possible cases: (1) If $\ell \notin V'$, the contact between ℓ and i remains undetected, and i will infect on average $R_i = \frac{\beta}{\mu}(k_i - 1)$ more nodes as in the classical SIR. (2) If instead $\ell \in V'$, then the contact with i gets detected by the CT system and i goes into self-isolation immediately, thus avoiding any secondary infection. If we denote by k'_i the degree of node i in G', the expected number of infections caused by the infection of i is equal to

$$R_i = \frac{\beta}{\mu}(k_i - 1)\frac{k_i - k'_i}{k_i},$$

and the expected number of secondary infections caused by each node infected by ℓ is given by

$$R_\ell = \frac{\beta}{\mu}\frac{1}{k_\ell}\sum_i a_{\ell i}(k_i - 1)\frac{k_i - k'_i}{k_i},$$

where $a_{\ell i}$ are the entries of the adjacency matrix of the contact graph G. By averaging R_ℓ over all the nodes of G we obtain the value of the basic reproduction number in the presence of CT (see Appendix C for details of the derivation),

$$R'_0 = R_0 - \frac{1}{N}\frac{\beta}{\mu}\sum_\ell \frac{1}{k_\ell}\sum_i a_{\ell i}\frac{k'_i}{k_i}(k_i - 1) \qquad (1)$$

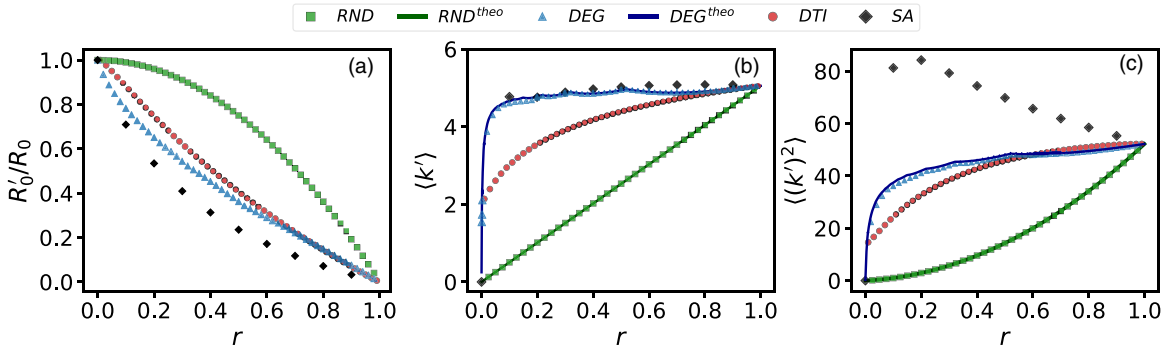

FIG. 2. Effect of different CT strategies on the induced subgraph as a function of app adoption rate r. (a) The ratio R_0'/R_0 as a function of the CT adoption rate r for the RND, DEG, DTI, and SA strategies calculated using Eq. (1). (b) Average degree of G' as a function of the adoption rate r for each of the strategies. The theoretical predictions for the RND and DEG strategies (solid lines) were obtained using Eqs. (B3) and (B13), respectively. (c) Second moment of the degree distribution of G' as a function of the adoption rate r for each of the strategies. The plots correspond to an ensemble of configuration model graphs with degree distribution $P(k) \sim k^{-3}$ and $N = 10^4$ nodes. Results averaged over 100 realizations.

where R_0 is the basic reproduction number of the classical SIR dynamics on G [10]. As made clear by Eq. (1), we can minimize the value of R_0' by using a generic optimization algorithm to compute

$$\max_{G'} \quad \mathcal{F}(G') = \sum_{\ell} \frac{1}{k_\ell} \sum_{i} a_{\ell i} \frac{k_i'}{k_i}(k_i - 1), \tag{2}$$

over the ensemble of possible choices of G'. Notice that if the entire population installs CT apps (i.e., if $k_i' = k_i \; \forall i \in V$) we trivially get $R_0' = 0$ (see Appendix C and Supplemental Note 1 for details [45]). This formulation is valid for networks with any number of nodes N with or without degree-degree correlations and not necessarily being treelike.

B. CT targeting strategies

If we assume that we can install the CT app only to a fraction $r \in [0, 1]$ of "willing" individuals, Eq. (2) states that a good CT installation strategy should include in G' nodes having a high degree in G (so that the ratio $\frac{(k_i-1)}{k_i}$ is as large as possible) and, at the same time, a high number of connections to other nodes in G' (i.e., so that k_i' is as large as possible). The most basic strategy to select a fraction r of the N individuals to install CT apps consists in asking the population to install a CT app on their mobile phones under the assumption that each individual will comply with probability equal to r, irrespective of any of their specific social or behavioral characteristics. In this case, the total number of installations will be distributed according to a binomial with mean equal to rN. In the following, we call this strategy "uniform random installation" (RND).

A second strategy consists in explicitly targeting all the potential *super-spreaders* [46,47]. In practice, we ask the rN individuals with the largest number of contacts (links) in G to install the app, assuming that they will all comply with probability 1. This strategy is indeed utopistic since it requires full knowledge of the contact network and full compliance by the selected nodes. In the following we call this strategy "degree-based installation" (DEG).

Here we propose and study a constructive strategy to maximize Eq. (2) that does not require detailed global information on G and thus lends itself easily to a distributed implementation. We start from a CT set that contains only the node with the largest degree in G. Then, at each subsequent step t, we add to the CT set one of the neighbors i of any of the nodes in V' with probability proportional to the total number of neighbors of that node that are already in V' (see Appendix A for details). This creates a "social pressure" on individuals with no CT app installed which is proportional to the number of their contacts already in V'. We call this strategy "distributed targeting installation" (DTI). Notice that the choice of the initial node is not that important, and qualitatively similar results are obtained when the first node is chosen uniformly at random rather than according to its degree.

In Fig. 2(a) we plot the ratio $\frac{R_0'}{R_0}$ as a function of the CT adoption rate r for the RND, DEG, and DTI strategies on an ensemble of configuration model graphs with power-law degree distributions. As a reference, we also report the results obtained by optimizing Eq. (2) by means of simulated annealing (SA). It is worth noting that RND is the worst-performing strategy overall, characterized by a much slower decrease of R_0' with r. Conversely, DEG is close to the theoretical limit established by SA and produces a noticeable decrease of R_0 already for quite small values of r. Remarkably, the performance of DTI is quite close to that of DEG, although DTI is not using any global information about the structure of G. In Figs. 2(b) and 2(c) we show how the first and second moments of the degree distribution $\widetilde{P}(k')$ of the subgraph G' vary with r for each of the four strategies. More details on the derivation of the full degree distribution of G' in RND and DEG are reported in Appendix B, while Supplemental Fig. S1 [45] shows the perfect agreement between the empirical and the analytical degree distributions for these two strategies.

It is worth noting that under the DEG strategy, $\langle k' \rangle$ increases very sharply with r and is already quite similar to the value of $\langle k \rangle$ in G for very small values of r. On the other hand, in RND, $\langle k' \rangle$ increases only linearly with r (see Appendix B for details) while the performance of DTI is in between those two. However, these plots make it clear that the sheer density of G' is not the only important ingredient for CT app installation. Indeed, SA can attain consistently lower values of $\frac{R_0'}{R_0}$ than DEG, although the values of $\langle k' \rangle$ produced

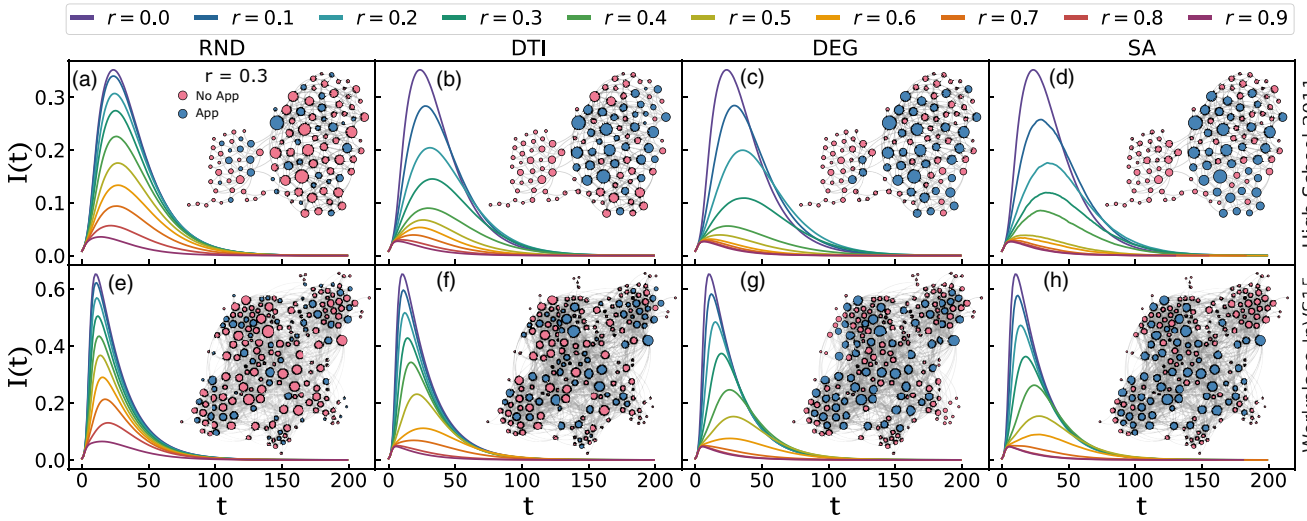

FIG. 3. Impact of CT strategy and adoption rate on the epidemic peak of SIR + CT in real-world networks. The evolution of the disease in a SIR + CT model (here for $\beta = 0.1$, $\mu = 0.05$) depends heavily on the adoption rate r and on the strategy used to select which individuals will have a CT app installed. We show here the results on two real-world social networks, namely, the high-resolution face-to-face contact data recorded in a high school (a–d) and a workplace (e–h), respectively. We have applied a threshold to both contact networks: Keeping only contacts larger than 240 links for the high school and 10% of the links with the largest weight for the workplace (See Appendix D for details). The high school graph contains 117 nodes and 332 edges whereas the workplace has 208 nodes and 1656 edges. At adoption rate $r \gtrsim 0.4$, the RND strategy (a, e) displays a higher percentage of infected compared with DTI (b, f), DEG (c, g), and SA (d, h). These differences are likely linked to the structure of the subgraph induced by each CT strategy. Typical examples of those graphs for each strategy and $r = 0.3$ are shown in the insets, where the size of each node is proportional to its degree and nodes with CT app installed are indicated in blue.

by SA are almost identical to those provided by DEG [see Fig. 2(b)].

C. SIR + CT in real-world graphs

In Fig. 3 we show the ratio of infected nodes $I(t)$ for the four strategies with different values of r on two real-world contact networks, the network of friendship in a high school (top panels) and at a workplace (bottom panels), respectively [48]. Here each simulation is performed by initially infecting one seed node chosen uniformly at random. In the inset of each panel we report a typical composition of V' for each strategy. It is evident that, even at low adoption rates, the DEG and DTI strategy can heavily mitigate the incidence of the disease better than RND. This is most probably due to the fact that DEG and DTI are targeting different sets of nodes than RND and in general end up selecting nodes with high degree which results in a higher edge density in G'.

The dynamics of SIR + CT in the two systems exhibit some noticeable qualitative differences when distinct strategies are adopted with respect to the height of the infection peak (the maximum incidence of the disease), the actual position of the peak (the time at which it occurs), and the overall duration of the epidemic. Interestingly, the position of the peak shifts to the right (delays) at small values of r for the DEG, DTI, and SA strategies. Conversely, the peak starts to recede (it is anticipated) with respect to the baseline when r becomes larger than a certain threshold, which depends on the particular structure of the contact network. While low values of r lead to a delay in the dynamics—the peak shifts to the right—large enough values of r effectively break the network into a number of disconnected components, resulting in a considerable disruption of the spreading—the peak shifts to the left.

Notice that the results shown in Fig. 3 correspond to $R_0 = 2$. In general, R_0 is disease dependent and its value deeply affects the evolution of an epidemic. However, in the proposed model, the actual value of R_0 does not alter the qualitative behavior of the dynamics since we assumed that there is no delay in the notification of contacts with infected individuals. Indeed, as can be seen in Eqs. (C3) and (C6) in Appendix C, μ and β only change the slope of the dynamics but not its overall qualitative behavior. In other words, a higher value of R_0 would make the dynamics faster but not qualitatively different. Conversely, if we consider a model with delay in the reporting of cases, as the one discussed in the Supplemental Information [45], R_0 could have a stronger impact since the disease could spread faster than the notification of contacts with infected individuals [33,41]. To better understand these qualitative differences, we look at three key properties of the epidemic curve, namely, the total number of individuals recovered R_∞, the maximum number of individuals infected across the duration of the epidemic $I(t_{\text{peak}})$, and the time to reach the infection peak t_{peak}. In particular, we compute the relative performance of each strategy s (DEG, DTI, or SA) with respect to RND using the quantities,

$$\Delta R^s_\infty = 1 - \frac{R^s_\infty}{R^{\text{RND}}_\infty}$$

$$\Delta I^s(t_{\text{peak}}) = 1 - \frac{I^s_{t_{\text{peak}}}}{I^{\text{RND}}(t_{\text{peak}})}$$

$$\Delta t^s_{\text{peak}} = 1 - \frac{t^s_{\text{peak}}}{t^{\text{RND}}_{\text{peak}}}. \tag{3}$$

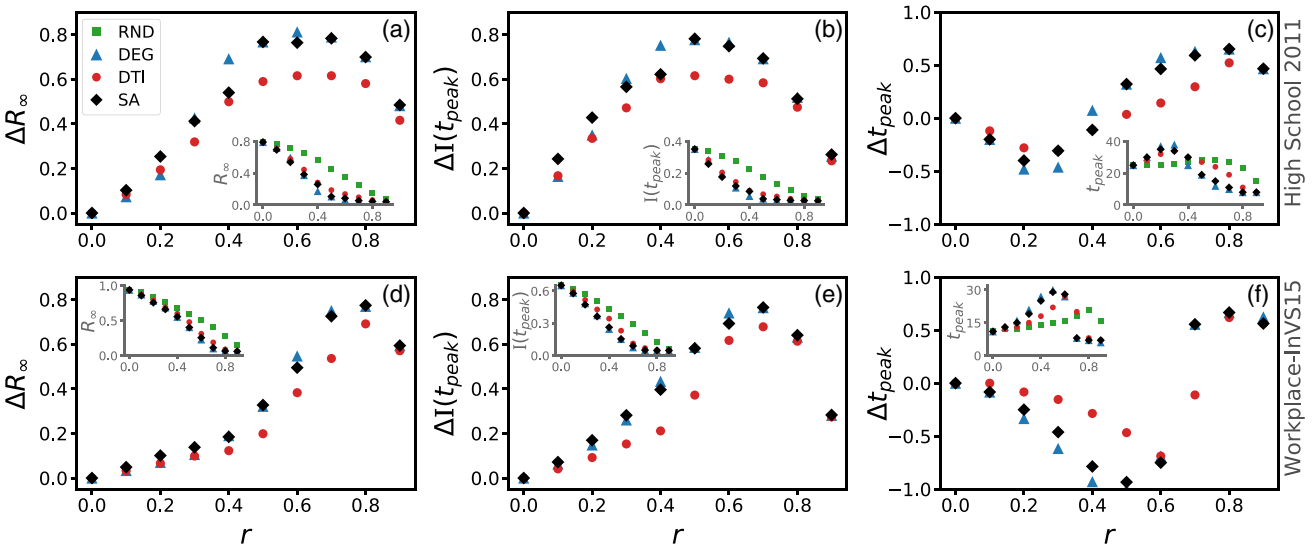

FIG. 4. Comparison of epidemic indicators under different CT strategies. Relative decrease with respect to random installations of the total number of recovered ΔR^s_∞, height of the infection peak $\Delta I^s(t_{\rm peak})$, and position of the peak $\Delta t^s_{\rm peak}$ [see Eq. (3)] for DTI-, DEG-, and SA-targeted installation in the same contact networks shown in Fig. 3. The inset of each panel reports the plot of the raw variable R_∞ (panels a and d), $I(t_{\rm peak})$ (panels b and e), and $t_{\rm peak}$ (panels c and f), respectively.

The results are shown in Fig. 4. We found that in the high school network an adoption rate of $r = 0.2$ can decrease the number of infected individuals at the peak by as much as 40% [Fig. 4(b)] for all three strategies. At the same adoption rate, the total number of infected individuals decreases by 20% [Fig. 4(a)] and the peak is delayed by approximately 40% [Fig. 4(c)]. For $r = 0.5$, we observe a substantially stronger mitigation where the peak is reduced by up to 70% and the total number of infected is reduced by up to 80%. The effectiveness of targeting strategies is somehow less pronounced in the workplace graph at similar adoption rates (bottom panels). Interestingly, in both networks the DTI strategy performs similarly to DEG and SA, as can be observed in the insets in Fig. 4, which show the raw values of R_∞, $I(t_{\rm peak})$ and $t_{\rm peak}$.

We simulated the SIR + CT dynamics in 84 unique real-world contact network data sets filtered by applying two different thresholds for a total of 168 undirected graphs [48–55] (see Appendix D for details). In Figs. 5(a)–5(c) we report the Spearman correlation between the analytical R'_0 and the epidemiological indicators R_∞, $I(t_{\rm peak})$ and $t_{\rm peak}$ for the RND, DEG, and DTI strategies with several adoption rates r, respectively. The correlation with both R_∞ and $I(t_{\rm peak})$ is high for the three strategies confirming the analytical predictions of Eq. (1) despite the small size of the graphs and the presence of degree-degree correlations. Still, as r increases we observe a decrease in the correlations likely due to finite size effects. The correlation between R'_0 and $t_{\rm peak}$ displays a much richer behavior: We start with a significant but negative correlation for small r, which changes sign until it reaches a maximum. We conjecture that the change of sign is related to the movement of the peak: Whereas in the small-r regime lower values of R'_0 contribute to a delay of the peak, for larger values of r we observe a stronger anticipation of the peak. The value $r_{\rm peak}$ at which the correlation peaks depends on the strategy in use; the more efficient it is, the lower the value of $r_{\rm peak}$. The concrete value of $r_{\rm peak}$ seems thus related to the actual structural properties of the graphs.

In a realistic scenario, in which the adoption rate is not fixed but needs to be promoted, we might be more interested in the minimum adoption rate r^* needed on each network to obtain a given reduction of the infection peak with respect to the absence of contact tracing. In Figs. 5(d)–5(f) we report the histograms of the value of r^* in DTI and RND for the 168 networks when we set a reduction in the peak $I(t_{\rm peak})$ of 10, 30, and 50%, respectively. We found that DTI can achieve a reduction of 30% of the peak in 85% of the networks with an adoption rate smaller than 0.3 [panel (e)], while the RND strategy would need an adoption rate of 0.5 to achieve an equivalent reduction.

While the strategies analyzed here require some level of global information and full compliance by individuals to install the app, we have obtained qualitatively similar results with other decentralized strategies based on local information and with a tunable level of compliance (see Supplemental Note 2 and Supplemental Figs. S2–S7 for details [45]). These strategies are inspired by the friendship paradox in social networks [56,57], which has been previously considered to implement efficient vaccination protocols [58–60]. The paradox states that, on average, in any graph each individual is more likely to have fewer friends than its own friends do, which is often expressed by saying that "your friends are more social than you are". We consider this concept to construct node rankings based on a simple and decentralized "voting" system. Furthermore, qualitatively similar results are also obtained when considering a SIR + CT with maximum delay, i.e., where an individual with a CT app installed goes into quarantine only when they become infected (see Supplemental Note 3 and Supplemental Figs. S7–S9 [45]). Although we already know that the expansion of infectious diseases usually depends on the age and socioeconomic status of individuals

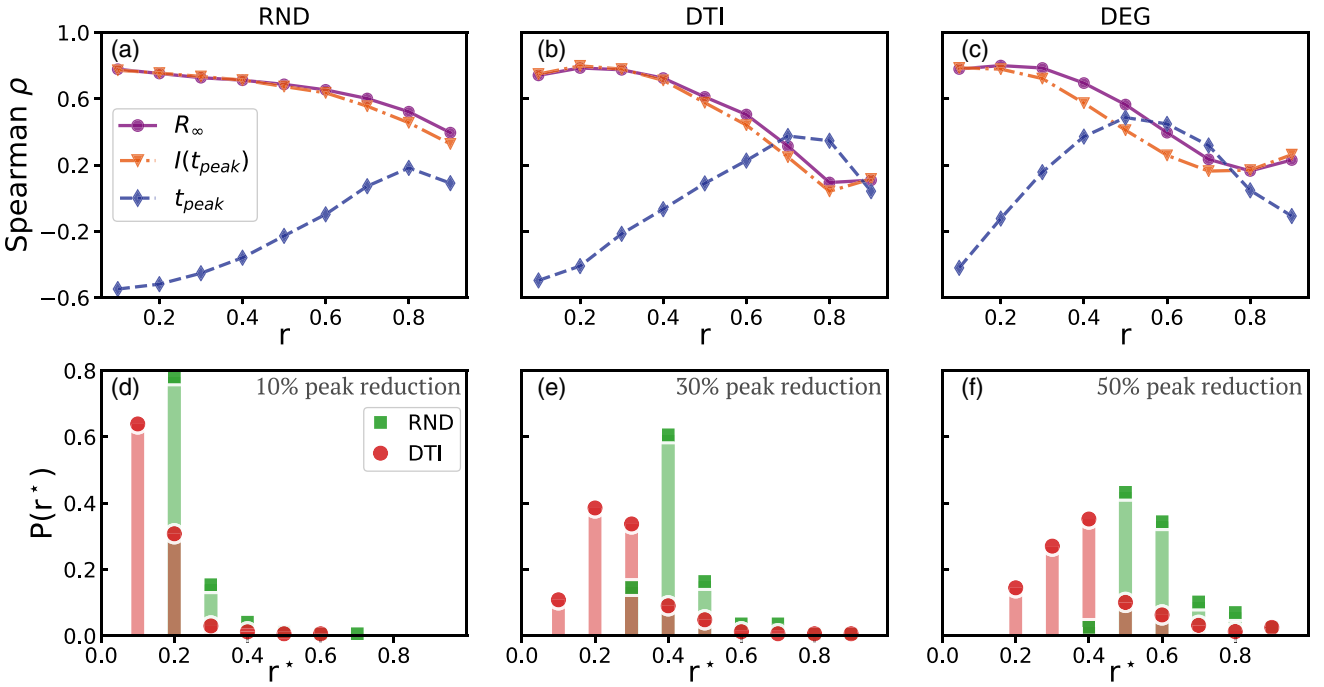

FIG. 5. Correlations with network structural measures and performance of CT strategies to mitigate an epidemic. For the 168 real-world contact networks analyzed, panels (a–c) report as a function of r the Spearman rank correlation between the analytical value of R'_0 [Eq. (1)] and the epidemiological indicators R^s_∞ (purple), $I^s(t_{\rm peak})$ (orange), and $t^s_{\rm peak}$ (blue) for the RND (a), DEG (b), and DTI (c) strategies. Panels (d–f) show the distribution of minimum adoption ratios r^* needed to produce 10% (d), 30% (e), and 50% (f) for the DTI (red) and RND (green) strategies. Overall, DTI largely outperforms RND, which is the strategy currently adopted by many governments.

[61,62], large contact networks containing such information are not yet widely available.

The present work provides a simple theoretical framework to optimize the installation of CT apps in order to contain the spread of an infectious disease in its early stages when vaccines are either not yet available or scarce. Our framework could be easily extended to include other compartments such as presymptomatic or asymptomatic subjects, which have proven to play an important role in the SARS-CoV2 pandemic as well as in other infectious diseases [14,19,63].

III. DISCUSSION

We have shown here that the random CT app installation, which has been widely considered in recent studies on the topic [36–38], is the least effective strategy to mitigate the effects of a pandemic through CT. The theoretical argument presented here, which links the reduction of R_0 to the structure of the subgraph G' induced by CT, holds for any graph under any CT strategy. In particular, Eq. (1) provides a tangible recipe to maximize the effectiveness of a CT app deployment as we have shown for many real-world contact networks.

The reduction of a disease incidence attainable by the DTI strategy is comparable with degree-based targeting, which performs similarly to the optimal targeting obtained through SA (see Supplemental Note 1 for details [45]). A notable advantage in using DTI over DEG is that it does not require any global information about the graph G and it can be implemented in a distributed manner. For instance, one could ask every new individual who installs the CT app to broadcast a message to all their contacts asking them to install it as well. By doing so, each contact with no app installed will be subject to a level of "social pressure" linearly proportional to the number of contacts who already have the CT app installed (in agreement with the heuristic algorithm of which DTI is based), consequently increasing the likelihood that the other person will also install it.

Although several effective vaccines for COVID-19 have recently been made available [64], mass vaccination campaigns are at their initial stages in many countries and might last for several months before a sufficient percentage of the population is vaccinated. Moreover, variants of the SARS-COV-2 may develop vaccine resistance and prolong the duration of the epidemic, creating an unsustainable loop of vaccine updates and vaccination campaigns. Hence reducing the spread of the virus by detecting potentially infected individuals and limiting their contacts–through digital CT—is still essential [65].

The model we considered here includes only some stylized features of a real-world epidemic, but the methodology we presented and the results we obtained pave the way to a deeper understanding of the effectiveness of CT systems. A direct application of these results to a real-world scenario would need to take into account several other ingredients. For instance, perfect compliance with public policy mandates (i.e., installation of CT apps) and the absence of any delay in the notification of individuals who test positive are just two simplistic assumptions that are easily false in a realistic setup. Nevertheless, the derivation proposed here points unmistakably to the central role played by appropriate CT app installation strategies. By following the same simple principles, governments could significantly improve the

effectiveness of CT programs even when more realistic factors are considered.

The contact networks used to model the epidemic spreading are public and have been published in Ref. [48].

ACKNOWLEDGMENTS

A.B. and V.N. acknowledge support from EPSRC New Investigator Award Grant No. EP/S027920/1. A.S. and V.N. acknowledge support from the EPSRC Impact Acceleration Award—Large Award Competition programme. This work made use of the MidPLUS cluster, EPSRC Grant No. EP/K000128/1. This research used Queen Mary's Apocrita HPC facility supported by QMUL Research-IT [66].

All the authors conceived the study. A.B. and A.S. performed the numerical simulations, analysed the results, and prepared the figures. S.S. and S.R. contributed to the methods for the analysis of the results and prepared the figures. V.N. provided methods for the analysis of the results and performed the analytical derivations. All the authors wrote the manuscript and approved it in its final form. A.B., A.S., and S.S. carried out the present research while affiliated at the Queen Mary University of London.

APPENDIX A: DTI STRATEGY

The proposed heuristic constructive algorithm to optimize Eq. (1), denominated DTI, starts with a set $V'(t=0)$ containing the node of largest degree in G. At each step t, we consider the set $S(t)$ of nodes which have at least one neighbor in $V'(t)$; then, a node i is selected at $t+1$ from $S(t)$ and added to $V'(t)$ with probability,

$$P(i;t) = \frac{\sum j \in V'(t) a_{ij}}{\sum_{i\in S(t)} \sum_{j\in V'(t)} a_{ij}}, \tag{A1}$$

i.e., node i is selected linearly proportional to the number of neighbors it has in $V'(t)$.

APPENDIX B: PROPERTIES OF THE SUBGRAPH INDUCED BY APP INSTALLATION

We provide here a sketch of the derivation of the first two moments of the degree distribution of the subgraph G' obtained from a graph G by considering only the nodes which have the contact-tracing app installed and the edges among them. The full derivations are provided in Supplemental Note 1 [45]. In the case of random installation strategy, the probability that a node installs the CT app is uniform across all nodes. As a consequence, the probability that a node with degree k in G has degree k' in G' is given by the binomial distribution

$$P(k'|k) = \binom{k}{k'} r^{k'}(1-r)^{k-k'}. \tag{B1}$$

This means that the expected degree in G' of a node that has degree k in G is just

$$E[k'_i] = rk_i. \tag{B2}$$

The degree distribution of the subgraph G' can be obtained by summing the probability in Eq. (B1) over all possible values of k, from which we obtain

$$\widetilde{P}_{\mathrm{RND}}(k') = \sum_{k=0}^{N-1} P(k) \binom{k}{k'} r^{k'}(1-r)^{k-k'}.$$

Finally, for the first two moments of $\widetilde{P}_{\mathrm{RND}}(k')$ we get

$$\langle k\rangle_{\mathrm{RND}} = \sum_{j=0}^{N-1} j\widetilde{P}_{\mathrm{RND}}(j) = r\langle k\rangle_G \tag{B3}$$

and

$$\langle k^2\rangle_{\mathrm{RND}} = \sum_{j=0}^{N-1} j^2\widetilde{P}_{\mathrm{RND}}(j) = r^2\langle k^2\rangle_G + r(1-r)\langle k\rangle_G. \tag{B4}$$

To compute the degree distribution of G' for degree-based installations, we start from the observation that a node is in G' only if its degree is $k_i \geqslant \widetilde{k}$, where $\widetilde{k}$ is obtained by solving the inequality

$$\sum_{\widetilde{k}}^{N} P(k) \geqslant r, \tag{B5}$$

where $P(k)$ is the degree distribution of G. Now the probability that one of the k_i neighbors of i is in G' is equal to

$$Q_{\widetilde{k}}(i) = \sum_{k=\widetilde{k}}^{N-1} P_{k|k_i}, \tag{B6}$$

where $P_{k|k_i}$ is the conditional probability of finding in G a node with degree k by following one of the edges of a node with degree k_i, chosen uniformly at random. In the special case of graphs with no degree-degree correlations, $P_{k|k_i} = \frac{kP(k)}{\langle k\rangle} = q_k$, so we have

$$Q_{\widetilde{k}}(i) = \sum_{k=\widetilde{k}}^{N-1} q_k = \widetilde{r} \quad \forall i. \tag{B7}$$

In the absence of degree-degree correlations, the probability of any two nodes to be connected does not depend on their degree, by definition. Hence the probability that a node of G' has a degree equal to k' (provided that it has degree equal to k in g) is given again by the binomial distribution

$$P(k'|k) = \binom{k}{k'} \widetilde{r}^{k'}(1-\widetilde{r})^{k-k'}, \quad k \geqslant \widetilde{k}, \tag{B8}$$

while $P(k'|k) = 0$ if $k < \widetilde{k}$. In particular, this means that the expected value $E[k'_i]$ is equal to

$$E[k'_i] = \widetilde{r}k_i. \tag{B9}$$

Notice that $\widetilde{r}$ has the same role that r has in the equations for uniform random installation. With an argument in all similar to that used for random installation, we obtain

$$\widetilde{P}(k') = \sum_{k=\widetilde{k}}^{N-1} P(k) \binom{k}{k'} \widetilde{r}^{k'}(1-\widetilde{r})^{k-k'}. \tag{B10}$$

Here $\widetilde{P}(k')$ represents the probability to find a node of G which has degree k' in the subgraph induced by app installations.

To obtain the actual degree distribution in the induced subgraph, i.e., the probability that one of the nodes of G' has degree k', we must rescale $\widetilde{P}(k')$ to the nodes in G', i.e., we consider the probability distribution

$$\widetilde{P}_{\mathrm{DEG}}(k') = \frac{1}{r}\widetilde{P}(k'). \tag{B11}$$

It is easy to show that $\widetilde{P}_{\mathrm{DEG}}(k')$ is correctly normalized

$$\begin{aligned}\sum_{k'=0}^{rN-1}\widetilde{P}_{\mathrm{DEG}}(k') &= \frac{1}{r}\sum_{k'=0}^{rN-1}\sum_{k=\widetilde{k}}^{N-1}P(k)\binom{k}{k'}\widetilde{r}^{k'}(1-\widetilde{r})^{k-k'}\\ &= \frac{1}{r}\sum_{k=\widetilde{k}}^{N-1}P(k)\sum_{k'=0}^{rN-1}\binom{k}{k'}\widetilde{r}^{k'}(1-\widetilde{r})^{k-k'}\\ &= \frac{1}{r}\sum_{k=\widetilde{k}}^{N-1}P(k) = 1.\end{aligned} \tag{B12}$$

The average degree in the induced graph is obtained as follows

$$\begin{aligned}\langle k\rangle_{\mathrm{DEG}} &= \sum_{k'=0}^{rN-1}k'\widetilde{P}_{\mathrm{DEG}}(k')\\ &= \frac{1}{r}\sum_{k=\widetilde{k}}^{N-1}P(k)\sum_{k'=0}^{rN-1}k'\binom{k}{k'}\widetilde{r}^{k'}(1-\widetilde{r})^{k-k'}\\ &= \frac{1}{r}\sum_{k=\widetilde{k}}^{N-1}P(k)k\widetilde{r} = \frac{\widetilde{r}}{r}\sum_{k=\widetilde{k}}^{N-1}kP(k) = \frac{\widetilde{r}^2}{r}\langle k\rangle,\end{aligned} \tag{B13}$$

where we have used the fact that $\sum_{k=\widetilde{k}}^{N-1}kP(k) = \widetilde{r}\langle k\rangle$ as per the definition of $\widetilde{r}$ in Eq. (B7). Similarly, for the second moment we obtain

$$\begin{aligned}\langle k^2\rangle_{\mathrm{DEG}} &= \sum_{k'=0}^{N-1}k'^2\widetilde{P}_{\mathrm{DEG}}(k')\\ &= \frac{1}{r}\sum_{k=\widetilde{k}}^{N-1}P(k)\sum_{k'=0}^{rN-1}k'^2\binom{k}{k'}\widetilde{r}^{k'}(1-\widetilde{r})^{k-k'}\\ &= \frac{1}{r}\sum_{k=\widetilde{k}}^{N-1}P(k)[k\widetilde{r}+k(k-1)\widetilde{r}^2]\\ &= \frac{\widetilde{r}^2}{r}\left[(1-\widetilde{r})\langle k\rangle + \sum_{k=\widetilde{k}}^{N-1}k^2P(k)\right].\end{aligned} \tag{B14}$$

APPENDIX C: REDUCTION OF R_0 UNDER IDEAL CT

We derive here a general expression for the effective value of the expected number of secondary infections caused by a single infection in a graph with perfect CT under the assumption that a fraction r of the nodes has installed a CT app. Assuming that a generic node ℓ is infected, we want to estimate the number of secondary infections caused by a node i infected by ℓ. The number of neighbors R_i that can be infected by i depends on whether i has the CT app installed and on how many of its neighbors have their app installed. In particular, if i does not have the app, $R_i = \frac{\beta}{\mu}(k_i - 1)$, since we have to remove the neighbor from which i got the disease. If i has a CT app installed, instead, there are two possible cases.

(1) If the node ℓ who infected i has the CT app, then the infection has been "detected" by the app and i goes into self-isolation immediately. If this happens, i will not produce any secondary infection in the graph.

(2) If i got infected by a neighbor without the CT app, then the infection remains undetected, and the expected number of nodes i can infect is given by $R_i = \frac{\beta}{\mu}(k_i - 1)$ as in the classical SIR (we have to remove l from the count; hence the $k_i - 1$).

If we take both possibilities into account, the expected number of infections needs to be multiplied by the probability that the infection of i goes undetected,

$$R_i = \frac{\beta}{\mu}(k_i - 1)\frac{k_i - k_i'}{k_i}, \tag{C1}$$

where the term $\frac{k_i - k_i'}{k_i}$ is the probability that the infection of node i does not get detected by the CT system. When $k_i' = k_i$, the node i and all its neighbors have the CT app installed producing $R_i = 0$. If, instead, node i is not in the CT subgraph, we recover $R_i = \frac{\beta}{\mu}(k_i - 1)$. Hence the expected number of secondary infections caused by a contact of ℓ with one of its neighbors is given by

$$R_\ell = \frac{\beta}{\mu}\frac{1}{k_\ell}\sum_i a_{\ell i}(k_i - 1)\frac{k_i - k_i'}{k_i}, \tag{C2}$$

where $a_{\ell i}$ are the entries of the adjacency matrix of G. By averaging R_ℓ over all the nodes of the graph we get the value of the basic reproduction number with CT,

$$\begin{aligned}R_0' &= \frac{1}{N}\sum_\ell\frac{1}{k_\ell}\sum_i a_{\ell i}R_i\\ &= \frac{\beta}{\mu N}\sum_\ell\frac{1}{k_\ell}\sum_i a_{\ell i}(k_i - 1)\frac{k_i - k_i'}{k_i}\\ &= \frac{\beta}{\mu N}\left[\sum_\ell\frac{1}{k_\ell}\sum_i a_{\ell i}(k_i - 1)\right.\\ &\quad\left. - \sum_\ell\frac{1}{k_\ell}\sum_i a_{\ell i}\frac{k_i'}{k_i}(k_i - 1)\right].\end{aligned} \tag{C3}$$

This equation holds in general for any graph with any CT strategy. Notice that the quantity

$$\frac{1}{N}\sum_\ell\frac{1}{k_\ell}\sum_i a_{\ell i}(k_i - 1) \tag{C4}$$

is the expected excess degree of the neighbors of a randomly sampled node of G. In other words, it is equal to $\langle k_{nn}(i)\rangle - 1$, where $k_{nn}(i)$ is the average degree of the neighbors of node i. The basic reproduction number of the original graph G is equal to

$$R_0 = \frac{\beta}{\mu}\frac{1}{N}\sum_\ell\frac{1}{k_\ell}\sum_i a_{\ell i}(k_i - 1), \tag{C5}$$

that is, the average degree of the neighbors of a randomly selected nodes of G, multiplied by $\frac{\beta}{\mu}$. Hence we can conve-

niently rewrite Eq. (C3) as

$$R_0' = R_0 - \frac{1}{N}\frac{\beta}{\mu}\sum_\ell \frac{1}{k_\ell}\sum_i a_{\ell i}\frac{k_i'}{k_i}(k_i - 1). \tag{C6}$$

In the special case when G has no degree-degree correlations, we have

$$\langle k_{nn}(i)\rangle = \sum_{k'} k' P(k'|k) = \sum_{k'} \frac{k'^2}{\langle k\rangle} P(k') = \frac{\langle k^2\rangle}{\langle k\rangle} \qquad \forall i, \tag{C7}$$

where we have considered that $P(k'|k) = \frac{k'P(k')}{\langle k\rangle} = q_{k'}$ in uncorrelated graphs. Hence we can write

$$R_0 \overset{\text{nc}}{=} \frac{\beta}{\mu}\left[\frac{\langle k^2\rangle}{\langle k\rangle} - 1\right], \tag{C8}$$

and we can rewrite Eq. (C3) as

$$R_0' \overset{\text{nc}}{=} \frac{\beta}{\mu}\left[\frac{\langle k^2\rangle}{\langle k\rangle} - 1 - \frac{1}{N}\sum_\ell \frac{1}{k_\ell}\sum_i a_{\ell i}\frac{k_i'}{k_i}(k_i - 1)\right]. \tag{C9}$$

As expected, the effect of CT is to reduce the basic reproduction number of the original graph. In general, Eq. (C6) [or Eq. (C9) in uncorrelated graphs] provides a recipe to maximize the impact of CT app installation. Indeed, given a certain adoption rate r, we can use any optimization algorithm to maximize the fitness function

$$\max_{G'} \quad \mathcal{F}(G') = \sum_\ell \frac{1}{k_\ell}\sum_i a_{\ell i}\frac{k_i'}{k_i}(k_i - 1) \tag{C10}$$

over the ensemble of of the possible choices of G'. Notice that $\mathcal{F}(G')$ can be decomposed in two terms. The first one is

$$\sum_\ell \frac{1}{k_\ell}\sum_i a_{\ell i} k_i', \tag{C11}$$

that is, the sum of the expected degrees in the induced subgraph G' of the neighbors of nodes in G, while the second one is

$$\sum_\ell \frac{1}{k_\ell}\sum_i a_{\ell i}\frac{k_i'}{k_i}, \tag{C12}$$

i.e., the sum of the average fraction of degree in G' and degree in G of all the neighbors of nodes in G.

APPENDIX D: DATA DESCRIPTION AND SET OF NETWORKS STUDIED

In this work we considered 84 unique contact network data sets constructed from two different types of data: (i) Temporal network data, which provide information regarding the different contacts between individuals and the duration of each interaction. (ii) Static network data, where the contacts have been already aggregated for the whole duration and a corresponding weight is associated with each link. We reconstructed each network considering two distinct filtering thresholds by either time—in seconds for type (i)—or by fraction of links retained—weight values for type (ii)—resulting in 168 unique graphs.

For the networks of type (i), we considered hospital [50] and high school [67] temporal data sets from which we filtered the contacts by applying thresholds of 240 and 360 s, i.e., each temporal snapshot resulted in a distinct network. For an art gallery [68] we used the thresholds of 0 and 20 s. Notice that we selected these threshold values since they provide the largest connected component for each network. The type (ii) networks obtained from the "Sociopatterns" project include the contacts between individuals with a weight that corresponds to either the number of contacts or their duration [48,52,53]. Given that most type (ii) networks are densely connected and a significant proportion of the weights have small values, we filter the networks by keeping the top 25 and 10% links with the largest weights. A list of all networks considered here is reported in Supplemental Material Table S1 [45].

[1] W. McKibbin and R. Fernando, The economic impact of COVID-19, in *Economics in the Time of COVID-19* (CEPR Press, London, 2020), Vol. 45.

[2] W. J. McKibbin and R. Fernando, The global macroeconomic impacts of COVID-19: Seven scenarios Asian Economic Papers **20**, 1 (2021).

[3] E. A. Holman, R. R. Thompson, D. R. Garfin, and R. C. Silver, The unfolding COVID-19 pandemic: A probability-based, nationally representative study of mental health in the united states, Sci. Adv. **6**, eabd5390 (2020).

[4] E. R. Mega, Covid has killed more than one million people. how many more will die? Nature (London) (2020).

[5] N. Perra, Non-pharmaceutical interventions during the COVID-19 pandemic: A review, Phys. Rep. **913**, 1 (2021).

[6] D. Balcan, V. Colizza, B. Gonçalves, H. Hu, J. J. Ramasco, and A. Vespignani, Multiscale mobility networks and the spatial spreading of infectious diseases, Proc. Natl. Acad. Sci. U.S.A. **106**, 21484 (2009).

[7] S. Riley, Large-scale spatial-transmission models of infectious disease, Science **316**, 1298 (2007).

[8] N. M. Ferguson, D. A. T. Cummings, S. Cauchemez, C. Fraser, S. Riley, A. Meeyai, S. Iamsirithaworn, and D. S. Burke, Strategies for containing an emerging influenza pandemic in Southeast Asia, Nature (London) **437**, 209 (2005).

[9] V. Colizza, A. Barrat, M. Barthelemy, A.-J. Valleron, and A. Vespignani, Modeling the worldwide spread of pandemic influenza: Baseline case and containment interventions, PLoS Med. **4**, e13 (2007).

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

[11] S. Flaxman *et al.*, Estimating the effects of non-pharmaceutical interventions on COVID-19 in europe, Nature (London) **584**, 257 (2020).

[12] B. J. Cowling *et al.*, Impact assessment of non-pharmaceutical interventions against coronavirus disease 2019 and influenza in

Hong Kong: An observational study, Lancet Public Health **5**, e279 (2020).

[13] C. M. Peak, L. M. Childs, Y. H. Grad, and C. O. Buckee, Comparing nonpharmaceutical interventions for containing emerging epidemics, Proc. Natl. Acad. Sci. U.S.A. **114**, 4023 (2017).

[14] A. Arenas, W. Cota, J. Gómez-Gardeñes, S. Gómez, C. Granell, J.T. Matamalas, D. Soriano-Paños, B. Steinegger, Modeling the Spatiotemporal Epidemic Spreading of COVID-19 and the Impact of Mobility and Social Distancing Interventions, Phys. Rev. X **10**, 041055 (2020).

[15] S. Chang, E. Pierson, P. W. Koh, J. Gerardin, B. Redbird, D. Grusky, and J. Leskovec, Mobility network models of COVID-19 explain inequities and inform reopening, Nature **589**, 82 (2021).

[16] L. Di Domenico, G. Pullano, C. E. Sabbatini, P.-Y. Boëlle and V. Colizza, Impact of lockdown on COVID-19 epidemic in île-de-france and possible exit strategies, BMC Med. **18**, 240 (2020).

[17] W. J. Wiersinga, A. Rhodes, A. C. Cheng, S. J. Peacock, and H. C. Prescott, Pathophysiology, transmission, diagnosis, and treatment of coronavirus disease 2019 (COVID-19): A review, JAMA **324**, 782 (2020).

[18] N. Haug, L. Geyrhofer, A. Londei, E. Dervic, A. Desvars-Larrive, V. Loreto, B. Pinior, S. Thurner, and P. Klimek, Ranking the effectiveness of worldwide COVID-19 government interventions, Nat. Hum. Behav. **4**, 1303 (2020).

[19] J. Aguilar *et al.*, Impact of urban structure on infectious disease spreading, Sci. Rep. **12**, 3816 (2021).

[20] P. Rodríguez, S. Graña, E. E. Alvarez-León, M. Battaglini, F. J. Darias, M. A. Hernán, R. López, P. Llaneza, M. C. Martín, O. Ramirez-Rubio, A. Romaní, B. Suárez-Rodríguez, J. Sánchez-Monedero, A. Arenas, and L. Lacasa, A population-based controlled experiment assessing the epidemiological impact of digital contact tracing, Nat. Commun. **12**, 587 (2021).

[21] W. J. Bradshaw, E. C. Alley, J. H. Huggins, A. L. Lloyd, and K. M. Esvelt, Bidirectional contact tracing could dramatically improve COVID-19 control, Nat. Commun. **12**, 232 (2021).

[22] A. Aleta *et al.*, Modelling the impact of testing, contact tracing and household quarantine on second waves of COVID-19, Nat. Hum. Behav. **4**, 964 (2020).

[23] J. Panovska-Griffiths, C. C. Kerr, R. M. Stuart, D. Mistry, D. J. Klein, R. M. Viner, and C. Bonell, Determining the optimal strategy for reopening schools, the impact of test and trace interventions, and the risk of occurrence of a second COVID-19 epidemic wave in the UK: A modelling study, The Lancet Child Adolesc. Health **4**, 817 (2020).

[24] J. Cohen and K. Kupferschmidt, Countries test tactics in 'war' against COVID-19 Science **367**, 1287 (2020).

[25] M. Abueg *et al.*, npj Digit. Med. **4**, 49 (2021).

[26] T. Jiang, Y. Zhang, M. Zhang, T. Yu, Y. Chen, C. Lu, J. Zhang, Z. Li, J. Gao, and S. A. Zhou, A survey on contact tracing: The latest advancements and challenges, ACM Trans. Spat. Algorithms Syst. **8**, 1 (2022).

[27] C. C. Kerr, D. Mistry, R. M. Stuart, K. Rosenfeld, G. R. Hart, R. C. Núñez, J. A. Cohen, P. Selvaraj, R. G. Abeysuriya, M. Jastrzębski, L. George, B. Hagedorn, J. Panovska-Griffiths, M. Fagalde, J. Duchin, M. Famulare, and D. J. Klein, Controlling COVID-19 via test-trace-quarantine, Nat. Commun. **12**, 2993 (2021).

[28] A. J. Kucharski *et al.*, Effectiveness of isolation, testing, contact tracing, and physical distancing on reducing transmission of SARS-CoV-2 in different settings: A mathematical modelling study, Lancet Infect. Dis. **20**, 1151 (2020).

[29] L. E. Smith, H. W. Potts, R. Amlôt, N. T. Fear, S. Michie, and G. J. Rubin, Adherence to the test, trace, and isolate system in the UK: Results from 37 nationally representative surveys, BMJ **372**, n608 (2021).

[30] K. T. Eames, and M. J. Keeling, Contact tracing and disease control, Proc. R. Soc. B: Biol. Sci. **270**, 2565 (2003).

[31] D. Klinkenberg, C. Fraser, and H. Heesterbeek, The effectiveness of contact tracing in emerging epidemics, PloS One **1**, e12 (2006).

[32] M. E. Kretzschmar, G. Rozhnova, M. C. Bootsma, M. van Boven, J. H. van de Wijgert, and M. J. Bonten, Impact of delays on effectiveness of contact tracing strategies for COVID-19: A modelling study, Lancet Public Health **5**, e452 (2020).

[33] L. Ferretti, C. Wymant, M. Kendall, L. Zhao, A. Nurtay, L. Abeler-Dörner, M. Parker, D. Bonsall, and C. Fraser, Quantifying SARS-CoV-2 transmission suggests epidemic control with digital contact tracing, Science **368**, eabb6936 (2020).

[34] M. J. Keeling, T. D. Hollingsworth, and J. M. Read, Efficacy of contact tracing for the containment of the 2019 novel coronavirus (COVID-19), J. Epidemiol. Community Health **74**, 861 (2020).

[35] Y. J. Park *et al.*, Contact tracing during coronavirus disease outbreak, South Korea, 2020, Emerg. Infect. Dis. **26**, 2465 (2020).

[36] A. Barrat, C. Cattuto, M. Kivelä, S. Lehmann, and J. Saramäki, Effect of manual and digital contact tracing on COVID-19 outbreaks: A study on empirical contact data, J. R. Soc. Interface. **18**, 20201000 (2021).

[37] A. Reyna-Lara, D. Soriano-Paños, S. Gómez, C. Granell, J. T. Matamalas, B. Steinegger, A. Arenas, and J. Gómez-Gardeñes, Virus spread versus contact tracing: Two competing contagion processes, Phys. Rev. Research **3**, 013163 (2021).

[38] I. Kryven, and C. Stegehuis, Contact tracing in configuration models, J. Phys. Complex. **2**, 025004 (2021).

[39] G. Bianconi, H. Sun, G. Rapisardi, and A. Arenas, Message-passing approach to epidemic tracing and mitigation with apps, Phys. Rev. Research **3**, L012014 (2021).

[40] Statista.com statistics on Contact-tracing app adoption rates, as of July 2020, https://www.statista.com/statistics/1134669/share-populations-adopted-covid-contact-tracing-apps-countries/ (2020) [Online; accessed 2021-01-27].

[41] J. Müller and B. Koopmann, The effect of delay on contact tracing, Mathematical Biosciences **282**, 204 (2016).

[42] M. Cevik, M. Tate, O. Lloyd, A. E. Maraolo, J. Schafers, and A. Ho, SARS-CoV-2, SARS-CoV, and MERS-CoV viral load dynamics, duration of viral shedding, and infectiousness: A systematic review and meta-analysis, Lancet Microbe **2**, e13 (2021).

[43] M. Cevik, K. Kuppalli, J. Kindrachuk, and M. Peiris, Virology, transmission, and pathogenesis of SARS-CoV-2, BMJ **371**, m3862 (2020).

[44] E. A. Meyerowitz, A. Richterman, I. I. Bogoch, N. Low, and M. Cevik, Towards an accurate and systematic characterisation of persistently asymptomatic infection with SARS-CoV-2, Lancet Infect. Dis. **21**, e163 (2020).

[45] See Supplemental Material at http://link.aps.org/supplemental/10.1103/PhysRevResearch.4.023092 for the full analytical

derivation of the degree distribution of the induced graphs for the random and degree-based strategies, and further analyses for a SIR model that accounts for maximum delay in self-isolation.

[46] A. Gómez-Carballa, X. Bello, J. Pardo-Seco, F. Martinón-Torres and A. Salas, Mapping genome variation of SARS-CoV-2 worldwide highlights the impact of COVID-19 super-spreaders, Genome Res. **30**, 1434 (2020).

[47] S. L. Miller, W. W. Nazaroff, J. L. Jimenez, A. Boerstra, G. Buonanno, S. J. Dancer, J. Kurnitski, L. C. Marr, L. Morawska, and C. Noakes, Transmission of SARS-CoV-2 by inhalation of respiratory aerosol in the Skagit Valley Chorale superspreading event, Indoor Air **31**, 314 (2021).

[48] DATASETS SocioPatterns.org, http://www.sociopatterns.org/datasets/ (2016), [Online; accessed 2020-09-30].

[49] M. Starnini, A. Machens, C. Cattuto, A. Barrat, and R. Pastor-Satorras, Immunization strategies for epidemic processes in time-varying contact networks, J. Theor. Biol. **337**, 89 (2013).

[50] P. Vanhems, A. Barrat, C. Cattuto, J.-F. Pinton, N. Khanafer, C. Régis, B. A. Kim, B. Comte, and N. Voirin, Estimating potential infection transmission routes in hospital wards using wearable proximity sensors, PloS One **8**, e73970 (2013).

[51] A. Barrat, C. Cattuto, A. E. Tozzi, P. Vanhems, and N. Voirin, Measuring contact patterns with wearable sensors: Methods, data characteristics and applications to data-driven simulations of infectious diseases, Clin. Microbiol. Infect. **20**, 10 (2014).

[52] M. Génois, C. L. Vestergaard, J. Fournet, A. Panisson, I. Bonmarin, and A. Barrat, Data on face-to-face contacts in an office building suggest a low-cost vaccination strategy based on community linkers, Netw. Sci. **3**, 326 (2015).

[53] M. Génois and A. Barrat, Can co-location be used as a proxy for face-to-face contacts? EPJ Data Sci. **7**, 11 (2018).

[54] L. Gadár, Z. T. Kosztyán, A. Telcs, and J. Abonyi, A multilayer and spatial description of the Erasmus mobility network, Sci. Data **7**, 41 (2020).

[55] P. Sapiezynski, A. Stopczynski, D. D. Lassen, and S. Lehmann, Interaction data from the Copenhagen Networks Study, Sci. Data **6**, 315 (2019).

[56] S. L. Feld, Why your friends have more friends than you do, Am. J. Sociol. **96**, 1464 (1991).

[57] Y.-H. Eom and H.-H. Jo, Generalized friendship paradox in complex networks: The case of scientific collaboration, Sci. Rep. **4**, 4603 (2014).

[58] R. Cohen, S. Havlin, and D. Ben-Avraham, Efficient Immunization Strategies for Computer Networks and Populations, Phys. Rev. Lett. **91**, 247901 (2003).

[59] M. Lelarge, Efficient control of epidemics over random networks, ACM SIGMETRICS Perform. Evaluation Rev. **37**, 1 (2009).

[60] N. A. Christakis and J. H. Fowler, Social network sensors for early detection of contagious outbreaks, PloS One **5**, e12948 (2010).

[61] G. E. Mena, P. P. Martinez, A. S. Mahmud, P. A. Marquet, C. O. Buckee, and M. Santillana, Socioeconomic status determines COVID-19 incidence and related mortality in Santiago, Chile, Science **372**, eabg5298 (2021).

[62] N. G. Davies *et al.*, Age-dependent effects in the transmission and control of COVID-19 epidemics, Nat. Med. **26**, 1205 (2020).

[63] M. Keeling and P. Rohani, *Modeling Infectious Disease in Humans and Animals* (Princeton University Press, Princeton, 2008).

[64] M. Jeyanathan, S. Afkhami, F. Smaill, M. S. Miller, B. D. Lichty, and Z. Xing, Immunological considerations for COVID-19 vaccine strategies, Nat. Rev. Immunol. **20**, 615 (2020).

[65] N. Gozzi, P. Bajardi, and N. Perra, The importance of non-pharmaceutical interventions during the COVID-19 vaccine rollout, PLoS Comput. Biol. **17**, e1009346 (2021).

[66] T. King, S. Butcher, and L. Zalewski, Apocrita - High Performance Computing Cluster for Queen Mary University of London (2017), https://doi.org/10.5281/zenodo.438045.

[67] M. Salathé, M. Kazandjieva, J. W. Lee, P. Levis, M. W. Feldman, and J. H. Jones, A high-resolution human contact network for infectious disease transmission, Proc. Natl. Acad. Sci. U.S.A. **107**, 22020 (2010).

[68] L. Isella, J. Stehlé, A. Barrat, C. Cattuto, J. F. Pinton, and W. Van den Broeck, What's in a crowd? analysis of face-to-face behavioral networks, J. Theor. Biol. **271**, 166 (2011).