INTERFACE

royalsocietypublishing.org/journal/rsif

Research

Cite this article: Cencetti G, Lucchini L, Santin G, Battiston F, Moro E, Pentland A, Lepri B. 2024 Temporal clustering of social interactions trades-off disease spreading and knowledge diffusion. *J. R. Soc. Interface* **21**: 20230471.
https://doi.org/10.1098/rsif.2023.0471

Received: 14 August 2023
Accepted: 23 November 2023

Subject Category:
Life Sciences–Physics interface

Subject Areas:
biocomplexity

Keywords:
complex networks, epidemics, simple and complex contagion, network structures

Author for correspondence:
Giulia Cencetti
e-mail: giuliacencetti@gmail.com

Electronic supplementary material is available online at https://doi.org/10.6084/m9.figshare.c.6961112.

THE ROYAL SOCIETY PUBLISHING

Temporal clustering of social interactions trades-off disease spreading and knowledge diffusion

Giulia Cencetti[1,2], Lorenzo Lucchini[3], Gabriele Santin[1,4], Federico Battiston[5], Esteban Moro[6,7], Alex Pentland[6] and Bruno Lepri[1]

[1]Digital Society Center, Fondazione Bruno Kessler, Trento, Italy
[2]Centre de Physique Théorique, CNRS, Aix-Marseille Univ, Université de Toulon, Marseille, France
[3]DONDENA and BIDSA Research Centres—Bocconi University, Milan, Italy
[4]Department of Environmental Sciences, Informatics and Statistics, University of Venice, Venezia, Italy
[5]Department of Network and Data Science, Central European University, Vienna, Austria
[6]Media Lab, Massachusetts Institute of Technology, Cambridge, MA, USA
[7]Department of Mathematics & GISC, Universidad Carlos III de Madrid, Leganes, Spain

GC, 0000-0002-6946-3666; LL, 0000-0002-2843-2580

Non-pharmaceutical measures such as preventive quarantines, remote working, school and workplace closures, lockdowns, etc. have shown effectiveness from an epidemic control perspective; however, they have also significant negative consequences on social life and relationships, work routines and community engagement. In particular, complex ideas, work and school collaborations, innovative discoveries and resilient norms formation and maintenance, which often require face-to-face interactions of two or more parties to be developed and synergically coordinated, are particularly affected. In this study, we propose an alternative hybrid solution that balances the slowdown of epidemic diffusion with the preservation of face-to-face interactions, that we test simulating a disease and a knowledge spreading simultaneously on a network of contacts. Our approach involves a two-step partitioning of the population. First, we tune the level of node clustering, creating 'social bubbles' with increased contacts within each bubble and fewer outside, while maintaining the average number of contacts in each network. Second, we tune the level of temporal clustering by pairing, for a certain time interval, nodes from specific social bubbles. Our results demonstrate that a hybrid approach can achieve better trade-offs between epidemic control and complex knowledge diffusion. The versatility of our model enables tuning and refining clustering levels to optimally achieve the desired trade-off, based on the potentially changing characteristics of a disease or knowledge diffusion process.

1. Introduction

The recent experience with COVID-19 has made us all aware of epidemics, their possible appearance and their likelihood to overturn our lives all of a sudden. The COVID-19 pandemic, as of June 2023, has caused almost 800 million cases (among which 7 million deaths) around the world.[1] Many containment measures and non-pharmaceutical interventions (NPIs) have been put in place before vaccines became available: lockdowns, preventive quarantines, masks, physical distancing, school and workplace closures, remote working, etc. [1–5]. All these measures have considerably impacted people's lives, social relationships, work and economy [6–10]. Indeed, if these measures can represent possible solutions to reduce disease spreading, the other side of the coin is that they imply a severe slowdown, or even interruption, of all the social exchanges and face-to-face interactions which prove fundamental for

the proper functioning of a society and, specifically, for fruitful collaborative interactions.

Moreover, despite many collaborations, intimate relationships, and, more in general, information and knowledge sharing can nowadays easily travel through the Internet, the limits of only remote interactions have become clear [11–13]. Several studies, indeed, have highlighted the importance of in-person interactions for physical, psychological and social wellbeing [14–17]. In particular, observational, interview-based and questionnaire-based studies have found that in workplaces face-to-face interactions are associated with increased trust and improved communication among employees, efficient problem solving and a positive effect on the overall organization knowledge diffusion, innovation ability and performance [18–21]. Similar results have been also found using wearable sensors to study face-to-face interactions and their effect on productivity, performance and complex tasks' completion [22–24]. For this reason it is crucial, along with the epidemic spreading reduction, to maintain as many physical interactions as possible. These two objectives are not easily pursued simultaneously. In this work, we will try to find a trade-off between them by working on the design of the network of social interactions, a useful exercise to find alternative solutions allowing us to cope with a fast-spreading disease. Tackling this problem within a coupled-dynamics framework can nourish new perspectives for the future management of epidemics and health emergencies.

Since the behaviour in time of a spreading process on a network is heavily conditioned by the network topology, several studies tried to regulate spreading by acting on the network structure [25–28]. In this work, inspired by Block *et al.* [29], we explore the 'social bubbles' strategy, which implies partitioning the society into communities where each individual can physically interact at will inside the bubble but in a controlled amount (or not at all) outside. In this way, people can maintain a normal amount of face-to-face interactions but restrict them to a set of people who interact exclusively in the same group. For example, in a workplace this could imply restricting face-to-face interactions only among the members of the same team or department, while in a school and in a university campus only within a classroom or a dorm. This strategy has been proposed and largely discussed, and many numerical experiments have been performed to assess the effect of social bubbles on real populations [30] and in specific contexts like schools [31–34] and workplaces [35], always showing important advantages in reducing contagions. In order to evaluate the effect of the community structure it is useful to consider synthetic networks where we can tune the modularity, i.e. the strength of division into modules or bubbles, and observe how this affects the spreading [36–38]. The size and density of the synthetic networks are chosen according to real data of physical interactions [39].

In this work, we explore the effect of the social bubbles' reorganization of a network of proximity interactions. The specific goal of this study is to find an optimal topological network structure that minimizes the number of infected individuals (and in particular the number of simultaneously infected individuals) in order to avoid burdening hospital intensive care units (ICUs), and, at the same time, minimizing the social deterioration due to restrictions [40,41]. For this reason, we consider two different and non-interacting spreading processes: one regarding a disease, and one regarding the diffusion of knowledge or of a social behaviour [42–47]. The first one is represented by a simple contagion model, a SIR (susceptible, infected or recovered) compartmental model inspired by Anderson & May [48] and Pastor-Satorras *et al.* [49], while the second one is governed by a complex contagion approach, the threshold model with memory [50]. The two processes take place simultaneously on an artificial population where individuals are connected via a temporal network, i.e. a set of pairwise links that appear and disappear in time. The way these links are distributed among nodes heavily affects the temporal evolution of the two spreading processes. The effect of social bubbles, which in networks is reflected by node clustering, is investigated for different sizes of the groups (e.g. different sizes of teams or departments in a workplace, different sizes of classrooms at school or university, etc.) and different levels of modularity (i.e. the strength of network partition, represented by the connectivity inside each bubble with respect to the admitted contacts between bubbles). Importantly, all these different networks have the same average number of links, such that our analysis is not influenced by the number of connections (a parameter that clearly has a role in fastening all kinds of spreading processes).

We show that it is possible to find a trade-off between minimizing the timescale of knowledge diffusion and the number of simultaneous infected individuals, two competing objectives. The effect of social bubbles (without preventive quarantines) is compared with the effect of quarantines (on a network that is not organized in bubbles). We will see that, even if the quarantines are more effective in containing the number of infected individuals, they do not allow knowledge diffusion until most of the population is recovered. By contrast, with the bubbles strategy it is instead possible to share knowledge in the network since the beginning of the simulation, and simultaneously maintain the number of infected below a critical level (generally higher but comparable with the case of quarantines). The bubbles strategy, therefore, allows social processes to coexist with an epidemic.

Additionally, we also find an optimal value of modularity for information diffusion, revealing a non-monotonic relation between knowledge diffusion and network structure. This is in agreement with the results found by Nematzadeh *et al.* [37] and Peng *et al.* [38] who, studying different models for information diffusion (a linear threshold model and an independent cascade model) on static networks, discovered the existence of an optimal value of modularity able to enhance both local and global spreading.

2. Results

Combining simple and complex contagion allows in general to find strategies that take into account both the epidemic threats and the socio-economical issues deriving from prolonged isolation periods, societal fragmentation and, potentially, segregation. Several works exist studying the interplay between different spreading processes [46,47,51]. These works, however, consider that the two processes mutually interact or one strongly affects the other one, while in our work they are considered two parallel processes (knowledge does not affect disease, and disease affects knowledge only indirectly, via isolations and quarantines). The focus of this work is indeed on the network structure

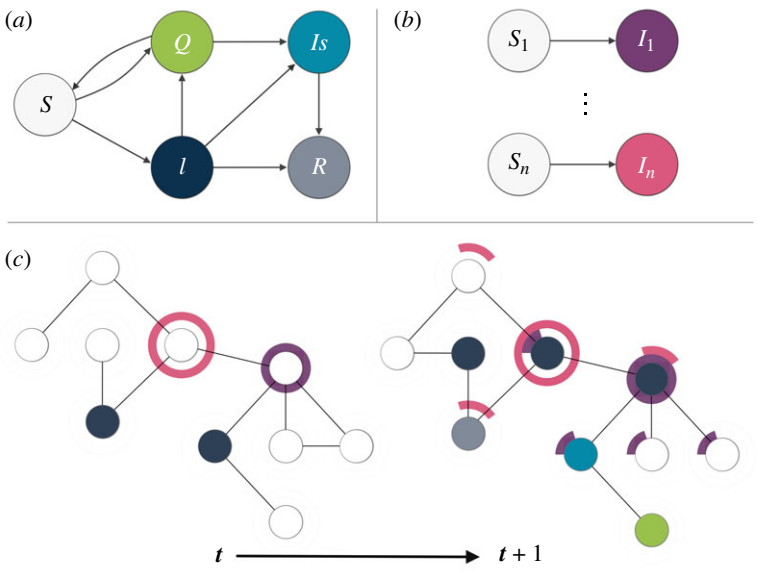

Figure 1. Schematic of the combined model dynamics. (*a*) The epidemic spread is modelled through a modified SIR compartmental model framework. '*S*' are individuals susceptible to the infection, '*Q*' are quarantined individuals, '*I*' are infected individuals, '*Is*' are isolated individuals, and '*R*' are recovered individuals. (*b*) The information dynamics follows a standard multistrain SI compartmental model. With 'S_i' representing agents susceptible to information '*i*', and 'I_i' representing agents who acquired information '*i*'. (*c*) Schematic of the main aspects of the two dynamics over a network of agents, showing the evolution from time *t* to time $t + 1$, and assuming for simplicity to have only two informations. Agents are represented with three concentric circles: the inner circle has a colour corresponding to its current epidemiological compartment (either susceptible, quarantined, infected or isolated—see (*a*)); the two outer circles represent one information each, and are partially coloured according to the fraction of information acquired at the given time.

and how this can be set so as to regulate different spreading processes taking place on it. We hence chose to consider a setting where the interaction between the two processes is minimal, thus avoiding to insert additional effects with the risk of not being able to understand from what they are generated, and hence to complicate the results' interpretation.

The disease starts from one random infected node at time 0. Simple contagion implies that nodes can only be infected if they have an infected neighbour and each contagion event is independent of the other ones. When a connection appears between an infected and a susceptible node, the probability that the susceptible node gets infected is set by $\omega(\tau)$, which depends on the age τ of the infected node's infection, i.e. the time since it has been in turn infected (see Methods for more details). Infected individuals can be identified and isolated, which means that all their connections are cut and they cannot spread disease or knowledge for a fixed interval of time. This happens with a probability per unit time ε_I, one of the parameters of the model. Infected nodes, whether they are isolated or not, eventually become recovered, hence immune. This is the baseline model that we use to simulate the disease spreading and to investigate social bubbles, but we can also additionally consider the existence of quarantines: in this case, once a node is isolated, its last contacts are traced and preventively quarantined with a probability ε_T, which means cutting the node's contacts for a short interval of time without knowing if it is actually infected or not. All infected nodes that are not isolated or quarantined are classified as active infected ones, being free to spread the disease (see Methods for a more detailed description of the disease spreading).

In parallel, knowledge spreads across the network. We consider 20 different pieces of knowledge distributed in the network, they can be thought of as 20 different pieces of information or expertise spread among different teams or departments in a workplace. Initially, each piece of knowledge is possessed by only one random node and we assume that the other nodes need multiple exposures in order to acquire them. This is represented by a threshold model with memory [50]: each susceptible node becomes infected (i.e. it acquires a specific knowledge) only after K interactions with nodes possessing that particular knowledge. Nodes progressively store and accumulate the different pieces of information that they get in touch with and only when the threshold is reached for a particular knowledge it is considered as acquired. This is an SI model: we assume that knowledge cannot be unlearned. See figure 1*c* for a schematic of the two processes: infected nodes (dark blue) infect their neighbours and then become recovered (grey) or isolated (bright blue) and their neighbours can be quarantined (green). In the meanwhile, nodes which possess a particular knowledge pass pieces of it to their neighbours. The neighbours start to collect them and when they obtain the entire set of K pieces that knowledge is acquired. In figure 1 only two different spreading pieces of knowledge are represented, the purple and the pink one, while in the numerical experiments we consider 20 pieces of knowledge.

We analyse several scenarios in which the two processes can interact and several networks of contacts. The temporal networks are generated building each layer with the stochastic block model [52] where the number of nodes is fixed ($N = 680$) and the number of links is fixed on average (400 links). The number of nodes and connections are chosen so as to mimic the proximity interactions' dataset of the Copenhagen Network Study [39]. The population is partitioned in several communities which are strongly connected inside (random temporal connections with a probability p_{intra}) and poorly connected between each other (random temporal connections with a probability p_{inter}). Since the number of links is fixed, the values of p_{intra} and p_{inter} are not independent of each other, and increasing p_{intra} implies decreasing p_{inter} and

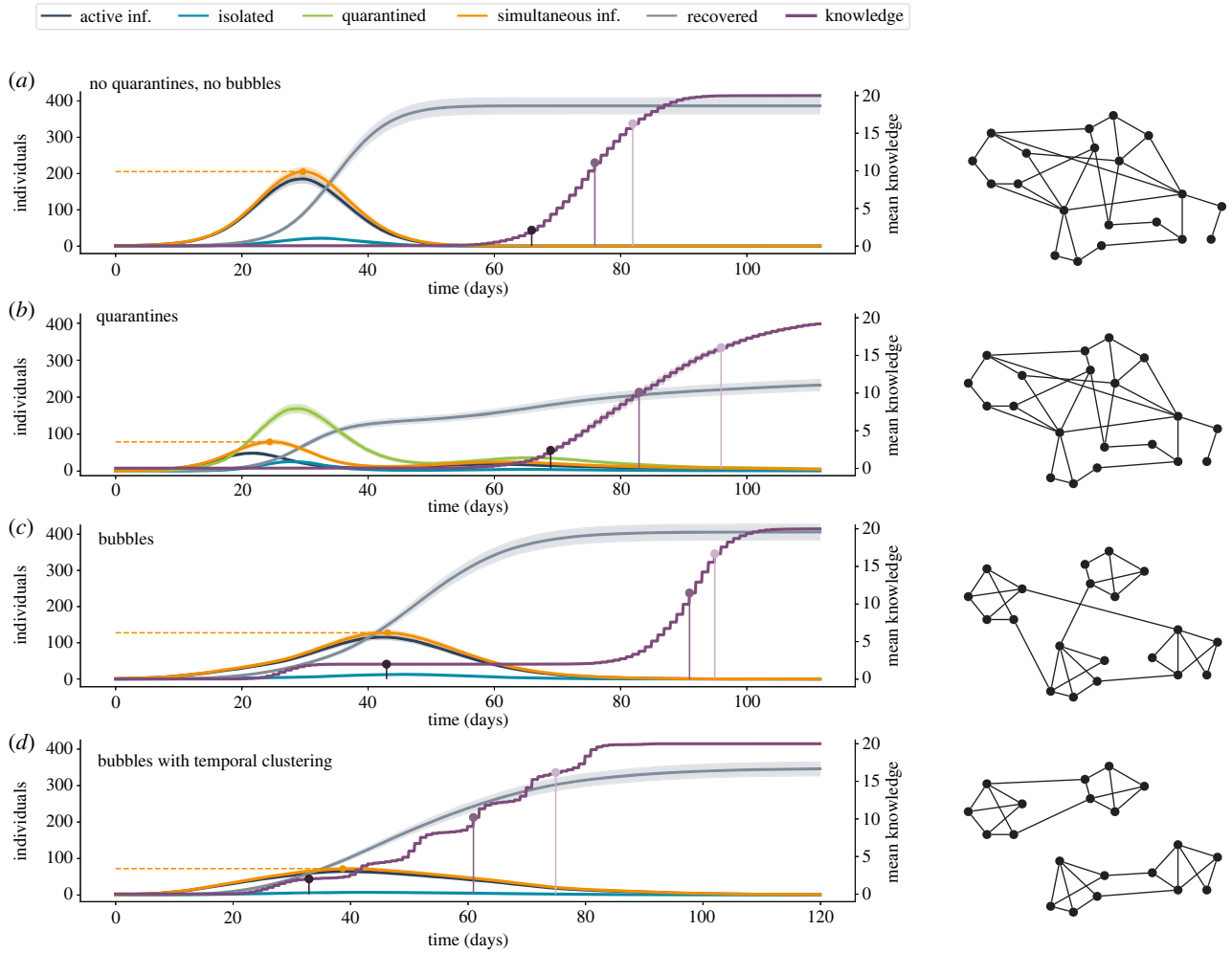

Figure 2. Time evolution of disease and knowledge spreading. Results of 200 simulations on temporal networks organized in 10 bubbles with 68 nodes. (*a*) $p = 5$ (i.e. connections inside and outside bubbles are of the same order; in practice, bubbles do not exist), $\varepsilon_T = 0$ (i.e. no quarantines). (*b*) $p = 5$, $\varepsilon_T = 0.1$ (i.e. quarantines without bubbles). (*c*) $p = 199$, $\varepsilon_T = 0$ (i.e. bubbles without quarantines). (*d*) $p = 199$, $\varepsilon_T = 0$, plus temporal clustering of 10 days. The parameter $\varepsilon_I = 0.1$ for all these cases.

vice versa. Thus, we rely on a parameter $p = p_{\text{intra}}/p_{\text{inter}}$ which represents the network modularity: namely, tuning the value of p allows us to generate more self-contained bubbles (higher p) or more interconnected groups less connected inside (lower p). By doing this, we can explore the effect of stronger or weaker bubbles as possible ways to reorganize a network without cutting or adding any link (see Methods for a detailed description of network generation).

In figure 2, we depict the evolution in time of both disease and knowledge spreading quantities: the number of active infected (dark blue) individuals, isolated (light blue) individuals and quarantined (green) individuals, plus the number of simultaneous infected (orange) individuals corresponding to the sum of active infected, isolated and true positive quarantined individuals (i.e. those that are actually infected). The horizontal dashed orange line highlights the maximum number of simultaneous active infected individuals reached during the temporal evolution. Additionally, we plot in purple the evolution of the mean number of different knowledge obtained by the nodes (averaged on all the nodes). We report mean knowledge in the y-axis on the right, which spans from 0 to 20 since we are considering 20 pieces of knowledge spreading in the network (at time 0 each piece of knowledge is only possessed by one node so the mean is $1/N = 0.0015$). We highlight three significant steps of knowledge spreading: the time when the average reaches the number of pieces of knowledge initially contained in one cluster, $20/N_c$ with N_c being the number of clusters ($N_c = 10$ in the example of figure 2); the time when it reaches 50% of total pieces of knowledge (10 in this case); and the time when it reaches 80% of total knowledge (16 in this case). All the reported results are obtained as averages over 200 stochastic simulations.

The first scenario that is depicted (figure 2*a*) corresponds to the case without quarantines ($\varepsilon_T = 0$) and where the social bubble strategy is not at play ($p = 5$, meaning that intra-bubble and inter-bubble connections are of the same order of magnitude). The only NPI is represented by the isolation of individuals who are identified as infected. This is the worst case: we have a peaked curve of infected individuals and at the maximum peak almost one third of the population is simultaneously infected. The average knowledge stays around 0 for the entire time span where people are infected and starts to grow only after the epidemic has been controlled.

Then, we introduce the scenario that serves as a benchmark to compare the social bubbles' strategy: it is the one with quarantines ($\varepsilon_T = 0.1$), still without the bubbles' organization of the network of interactions ($p = 5$). Quarantines clearly have an effect on infection number, managing to flatten the curves so as to reduce the number of simultaneously

active infected individuals. However, this reduction comes at a great social cost, confining a significative fraction of healthy individuals. In the example reported in figure 2*b*, while the maximum average number of simultaneous infected is only the 12% of the population, the average percentage of population confined at least once without being infected (collateral confinement) is as high as 34% (see electronic supplementary material, section S7, for collateral confinement in simulations with different parameters). Moreover, we observe that flattening the curve also implies extending the time span of the epidemics and this, in turn, badly affects the possibility of people having face-to-face interactions, thus slowing down knowledge diffusion. In fact, we observe that the purple curve starts to grow only when the other curves are very low. The time needed to acquire knowledge is hence longer than in the previous case. In other words, also in this case knowledge spreading via physical interactions can hardly coexist with an ongoing epidemic.

In figure 2*c*, we finally introduce the social bubbles strategy. In this case, no quarantine strategy is put in place (i.e. $\varepsilon_T = 0$) but the network is generated with $p = 199$, meaning that the intra-bubble interactions are, on average, 199 times more frequent than inter-bubble interactions. The number of infected individuals is higher with respect to the previous case: a simple organization of the network of interactions in bubbles is not able to reduce infections as quarantines, even if they are reduced with respect to the case without a bubble structure (figure 2*a*). However, we note that in this case knowledge starts to spread inside bubbles already during the unfolding of the epidemic. So, even if complete knowledge spreading remains quite slow, pieces of information that circulate inside bubbles guarantee that part of the knowledge is acquired from the beginning (in the reported case, already at day 43). This important achievement still has to pay the price of (i) a high number of simultaneously active infected individuals and (ii) a still long time before all the pieces of knowledge are able to reach all the nodes.

Hence, we consider an additional containment strategy: we add to the node clustering, represented by the social bubbles, a temporal clustering, thus obtaining temporal social bubbles. We leverage again the value of p setting the ratio of intra-bubble and inter-bubble connections but, in this strategy, the inter-bubble connections, instead of involving nodes of random different bubbles, are now concentrated only between specific couples of bubbles (as in the toy example depicted in figure 2*d*). So, each bubble (e.g. a team or department within an organization) only interacts with another bubble (e.g. a different team or department) at a specific time. Then, with a time periodicity d the couples change in such a way that, at the end of the simulation, each bubble has interacted with each other bubble. The result is that knowledge starts to grow from the early stages of the simulations inside each couple of bubbles and each node of one bubble easily acquires the ideas of its matched bubble, doubling the nodes' average knowledge. In a similar manner, once matches are updated, nodes are able to acquire knowledge from another bubble, and gradually augment their knowledge following a staggered growth. In figure 2*d*, this dynamic is clearly visible: the purple line shows the process of knowledge acquisition, which grows more rapidly in correspondence with the update of bubble matches (every $d = 10$ days in the reported simulations), and slows down once most of the matched bubble nodes have acquired the new piece of knowledge. For what concerns the disease spreading, leveraging the peculiar structure of node interactions, it naturally remains confined between a limited number of bubbles and, as a consequence, the number of infected individuals grows more slowly. Moreover, the number of simultaneously active infected does not reach a high value (only 11% of the population, similar to that of the quarantine strategy) since a fraction of the infected nodes can recover within the time that matches between bubbles are changed. By observing the curves in figure 2*d* we note that the disease curve and the knowledge one are partially overlapping, indicating that, in this framework, the epidemic can coexist with the diffusion of knowledge by social face-to-face interactions. Note also that in the simulations the epidemic variables are updated every 2 h while the knowledge variables only every 24 h. For this reason the epidemic curves are smooth while the knowledge ones reflect the daily updates.

In figure 3 we report, for different strategies and different parameters' settings, four significant quantities characterizing disease and knowledge spreading: (i) the maximum number of simultaneously active infected individuals (orange horizontal dashed lines in figure 2), (ii) the time at which the average knowledge acquired by nodes becomes equivalent to the number of different pieces of knowledge initially present in one cluster, (iii) the time at which it reaches 50% of the total knowledge and (iv) the time at which it reaches 80% of the total knowledge (the three purple vertical lines in figure 2). For the quarantine strategy, these indicators are reported versus ε_T and while, clearly, infected individuals decrease with ε_T, the knowledge times it is only marginally affected by it. For the bubbles' strategy, the indicators are instead depicted as functions of p, where increasing p means making the bubbles more self-contained. We consider the case with temporal clustering for three different values of d (i.e. 5, 10 and 20 days), and without temporal clustering. In all cases, we note that the number of infected individuals decreases with p, not drastically as for ε_T, but significantly. This is due to the fact that closer bubbles tend to maintain the disease confined to a few bubbles, while the other nodes remain safe. The lowest numbers of infected individuals are obtained with temporal clustering of 20 days; in fact with longer temporal clustering there is more chance that infected individuals inside a bubble recover or are quarantined before they have the possibility to meet new susceptible nodes. For what concerns knowledge times, instead, we note a peculiar behaviour with respect to the modularity, showing in all cases a minimum value in p, corresponding to an optimal value which varies according to different conditions. We also note that in order to shorten times to reach total knowledge, shorter temporal clustering is to be preferred (see light purple curves), while the partial knowledge times are not visibly affected by the length of the tournament d, at least for what concerns the three explored values.

3. Discussion

In this work, we have investigated the social bubbles framework as a potential strategy for controlling disease spreading and simultaneously allowing higher levels of face-to-face interactions, which in turn facilitate the process of knowledge diffusion. More specifically, we have considered different settings of the social interactions' network organization in bubbles with the aim of finding an alternative strategy to preventive quarantines.

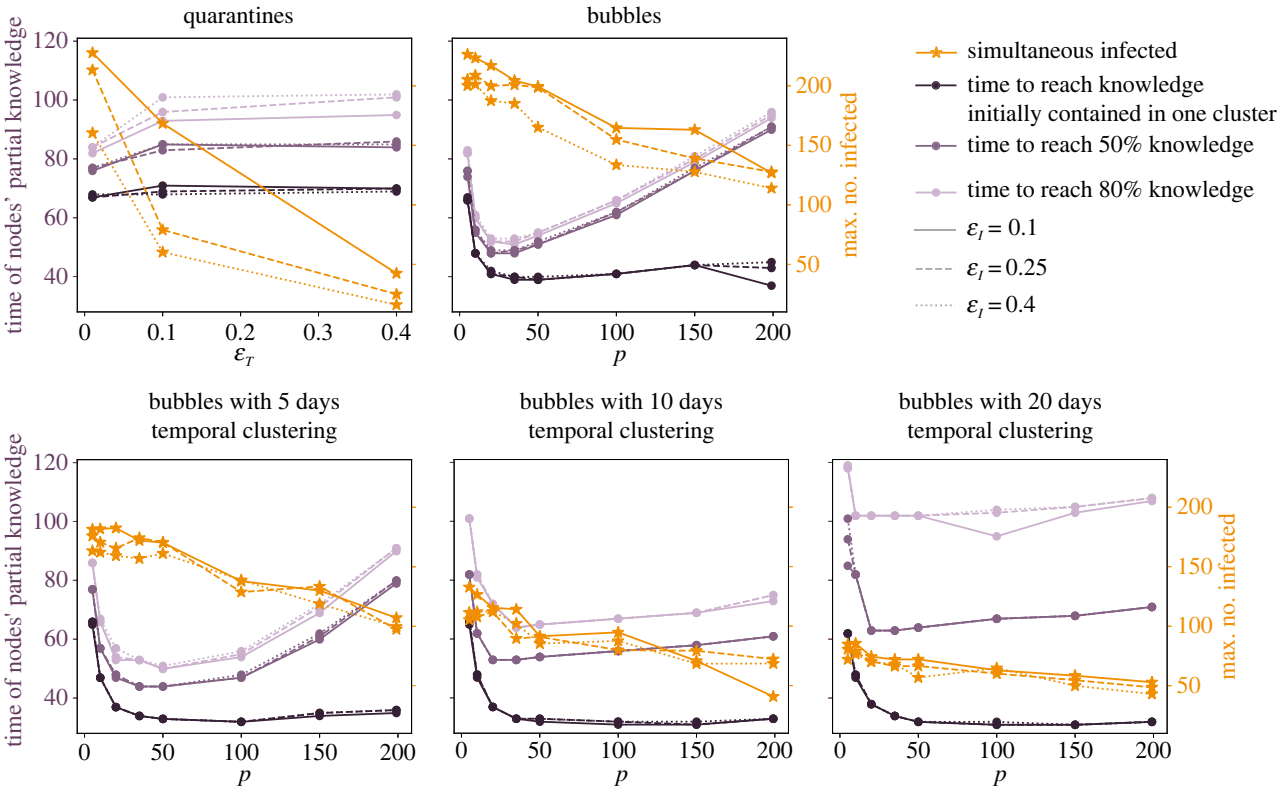

Figure 3. Scenario with 10 bubbles (i.e. 10 different teams or departments in an organization). Orange stars represent the maximum number of infected individuals, the other symbols represent the time at which the nodes reach on average 10%, 50% and 80% of knowledge. The first percentage corresponds to acquiring the entire knowledge initially contained in one bubble from all the nodes. We report results for $\varepsilon_I = 0.1$ (continuous lines), $\varepsilon_I = 0.25$ (dashed lines), $\varepsilon_I = 0.4$ (dotted lines). The quarantine results are reported versus ε_T, the bubble results versus p and for different levels of temporal clustering.

From the point of view of reducing the number of infected individuals, the quarantine strategy is the most effective one. However, arguably, quarantines impede social contacts and, as a consequence, all the diffusion processes that do not rely on simple contagion models but require multiple contacts to spread, as is the case for the considered knowledge spreading process. The bubble strategy instead permits face-to-face interactions between people, albeit limited to smaller groups than to an entire organization, school, university campus, workplace, etc. These free physical interactions allow complex dynamics (like complex contagions) to emerge, as is demonstrated by the progress of knowledge spreading, here represented by the threshold process. We note in fact that, despite the acquirement of the entire knowledge set still being reached after a long time, the singular ideas are free to circulate since the beginning inside the bubbles and are then easily acquired by the nodes (figure 2*c*).

Other strategies to flatten the infection curve by making use of social bubbles consist in cutting inter-bubble links instead of rewiring them [29]. However, the reduction of connections incontestably induces a slowdown of the epidemic spreading, which is not entirely due to the organization in bubbles. In our work, we show instead that the bubbles strategy is effective even if the number of links is maintained constant. In particular, we use networks with different levels of modularity, quantified by different values of p, but we generate the networks always with the same average number of links.

One of the most salient results of this work is that the existence of communities in a network facilitates knowledge diffusion (i.e. complex contagion), limiting collateral confinement, while it mitigates disease diffusion (i.e. simple contagion). An analogous conclusion about knowledge diffusion, implemented as a linear threshold model, has been obtained by Nematzadeh *et al.* [37] (and by Peng *et al.* [38] in a follow-up of the first work). They observe that strong communities enhance local spreading, while weak communities enhance global spreading, and they find an optimal range of intermediate values for community strength (analogous to our p) that maximize diffusion and the speed of cascades. This result agrees with our results on knowledge spreading; in fact, with our analyses we also observe the existence of an optimal value of p corresponding to the shortest times of knowledge diffusion (figure 3). The minimum in p tells us that the best performances of the bubbles strategy are not monotonic with p; on the contrary, very high or very low values of p slow down the knowledge diffusion. In fact, trivially, if p is too small we lose the effect of population partition and what we observe is a mixed population without bubbles and without quarantines, so a very inefficient network structure. If, instead, we increase p, i.e. we make the bubbles progressively more self-contained, we interestingly observe progressively longer knowledge spreading times, suggesting that a certain level of promiscuity between bubbles is instead advisable. It is important to note that the existence of quarantines and isolations does not affect the ratio p between the amount of intra- and inter-bubble connections. In electronic supplementary material, figure S9, we show indeed that, while the number of overall connections in the network is significantly reduced during the central phase of the disease epidemic, both within and between clusters, the ratio $p_{\text{intra}}/p_{\text{inter}}$ remains stable around the value p set to generate the

networks. This confirms the validity of the parameter p to discern the different networks and the resulting processes.

The greatest advantage is, however, obtained with the insertion of bubbles' temporal clustering. In fact in that case we have a combined effect: the existence of bubbles allows to keep the epidemics under control, and the temporal clustering enforces the nodes to be exposed to periodically different ideas. The result is a faster acquisition of knowledge while the disease is kept under control. This strategy reveals impressively effective and, to our knowledge, it is a completely novel idea.

All these results are confirmed by considering different sizes of bubbles (with the same number of nodes and the same average number of links), in particular 5 bubbles of 136 individuals and 20 bubbles of 34 individuals, as reported in section S2 of the electronic supplementary material.

Our work comes with some limitations that could be addressed by further exploring this research direction in the future. First of all, we are only considering synthetic networks, which are generated with random interactions without temporal correlations, clustering or other structural information that would make them more similar to real networks of interactions. The reason for this choice stemmed from the need to investigate the effect of social bubbles, disentangled from other possible structural constraints that could affect the dynamics. The choice of random networks ensures that the only structure existing in the considered graphs is the modularity, which we introduce and control by setting the parameter p when generating the networks. This allows us to directly scrutinize the phenomenon and draw untwisted conclusions. Envisaging possible applications to real networks will be a matter of future investigations. Moreover, the effect of network density is not explored, but we test different network sizes (increasing or decreasing the number of nodes and changing consequently the number of links so as to keep density fixed) and different bubble sizes and we observe that the results essentially do not change (see electronic supplementary material, section S3). Finally, only some of the simple and complex spreading parameters are explored, which implies that the results about optimal p are not general, they only apply in this specific context. However, the interesting result is that a minimum exists, even if its exact value will probably change by changing the parameters. In conclusion, we realize that the temporal social bubbles strategy represents a valid alternative to other NPIs like preventive quarantines when looking for a solution permitting to coexist with a spreading disease. This strategy, while still affecting the social structure of interactions, allows the pursuit of a series of otherwise hardly attainable collective goals that prove fundamental to the growth and social enrichment of society and individuals, from collaborations to social relationships, from knowledge transfer to opinion exchange. These results could spur innovative approaches to epidemic control strategies, for example, based on different interaction-mixing prescriptions in different settings, such as home, work and leisure places. While a systematic assessment of mixed strategy approaches is needed to better inform policymakers, our results provide a first insight about social bubbles strategies, exploring their efficacy under the prescription of temporal clustering. This study could help find viable solutions that are alternatives to standard ones, given that strategies solely focused on flattening the curve can dramatically affect knowledge diffusion and social cohesion.

4. Methods

4.1. Disease spreading

The disease-spreading model is non-Markovian and it is inspired by previous literature on COVID-19 models [53–55]. Each numerical simulation starts with one random infected individual, who has been infected for a number of days, τ, randomly sampled between 0 and 10. The variable τ is important in the spreading process since we assume that infectiousness, i.e. the probability of transmitting the disease, of an infected individual depends on the time since their own contagion, with a function $\omega(\tau)$ which has a maximum peak at around 5 days (see electronic supplementary material, figure S1). Such probability governs which individuals, among those that the infected seed meets according to the temporal network, will contract the disease. These can in turn infect their contacts. We assume that every infected individual becomes recovered, hence immune,[2] after 10 days [56]. We assume that 80% of infected individuals become symptomatic after being infected, with a symptom onset probability which increases in time according to a function $s(\tau)$ [55] (see electronic supplementary material, figure S1). As soon as they show symptoms they have a probability ε_I of being isolated. If this does not happen they go on spreading, otherwise their contacts are cut for the next 10 days, at the end of which they will become recovered. When an individual is isolated, their past contacts (last 7 days) are traced and preventively quarantined with a probability ε_T. The quarantined individuals can be infected (true positive) or susceptible (false positive). In the first case, if they show symptoms during quarantine, they will become isolated (the only difference with quarantine is that their past contacts are traced and, in the end, they will be recovered), otherwise, they finish the quarantine after 10 days and they are released. The reproduction number [57] with this parameter setting is $R_0 = 1.64$.

4.2. Network generation

We generate a synthetic temporal network with N nodes organized in N_c clusters, characterized by a strength of network modularity $p = p_{\text{intra}}/p_{\text{inter}}$, and a total number of links L. We hence generate a series of static networks which are going to constitute the layers of the temporal network. In each layer, nodes are classified into clusters (the first n_1 nodes in the first cluster, the second n_2 nodes in the second cluster, and so on, with n_i the *a priori* chosen number of nodes in cluster i, the same for each static network). All the temporal layers are characterized by the same values of intra-cluster connection probability p_{intra} and of inter-cluster connection probability p_{inter}.

The static networks are generated using the stochastic block model [52], once all the parameters (N_c, n_i, p_{intra}, p_{inter}) have been fixed.

Since p_{intra} and p_{inter} are not independent of each other but are constrained by the value of L, we need first to find the function that associates these three variables. We consider the general case where clusters do not contain the same number of nodes, but each cluster i contains a number n_i of nodes. This means that, statistically, in cluster i we can find a number of internal links given by

$$l^i_{\text{intra}} = \frac{n_i(n_i-1)}{2} p_{\text{intra}}, \tag{4.1}$$

and summing over all clusters we obtain

$$l_{\text{intra}} = \sum_{i=1}^{N_c} l^i_{\text{intra}} = \sum_{i=1}^{N_c} \frac{n_i(n_i-1)}{2} p_{\text{intra}}. \tag{4.2}$$

For what concerns the inter-cluster links, instead, we should consider that each node in cluster i can be connected to a number of external nodes $(N - n_i)p_{\text{inter}}$. This is true for every

node in the cluster, so it should be multiplied by n_i to find the number of connections of cluster i that point to other clusters:

$$l^i_{\text{inter}} = n_i(N - n_i)p_{\text{inter}}. \tag{4.3}$$

To obtain the total number of inter-cluster links we just have to sum over all the clusters and divide by 2, to avoid double counting:

$$l_{\text{inter}} = \sum_{i=1}^{N_c} \frac{l^i_{\text{inter}}}{2} = \sum_{i=1}^{N_c} \frac{n_i(N - n_i)}{2} p_{\text{inter}}. \tag{4.4}$$

The total number of links can, therefore, be written as

$$L = l_{\text{intra}} + l_{\text{inter}} = \sum_{i=1}^{N_c} \left[\frac{n_i(n_i - 1)}{2} p_{\text{intra}} + \frac{n_i(N - n_i)}{2} p_{\text{inter}} \right]. \tag{4.5}$$

This equation represents the constraint between L, p_{intra}, and p_{inter}. In our case, we consider all the clusters with the same number of nodes: $n_i = N/N_c \equiv n_c \ \forall i$, hence it reduces to

$$L = N_c \left[\frac{n_c(n_c - 1)}{2} p_{\text{intra}} + \frac{n_c(N - n_c)}{2} p_{\text{inter}} \right]. \tag{4.6}$$

The networks that we generate for this paper have a fixed number of links, $L = 400$. By varying the chosen value for p_{intra} we can obtain different choices of p_{inter} (and hence of p), always maintaining constant the number of links L, by inverting equation (4.6):

$$p_{\text{inter}} = \frac{2L}{N_c n_c (N - n_c)} - \frac{n_c - 1}{N - n_c} p_{\text{intra}}. \tag{4.7}$$

Data accessibility. This work did not include the use of real data. The networks used for the numerical simulations are synthetic, generated with the procedure described in the Methods section. The code used for the generation of temporal network, simulations and analysis is available from the GitHub repository: https://github.com/giuliacencetti/Social'bubbles [58].

Supplementary material is available online [59].

Declaration of AI use. We have not used AI-assisted technologies in creating this article.

Authors' contributions. G.C.: conceptualization, formal analysis, investigation, methodology, resources, software, validation, visualization, writing—original draft, writing—review and editing; L.L.: conceptualization, formal analysis, methodology, writing—review and editing; G.S.: conceptualization, formal analysis; F.B.: supervision; E.M.: supervision; A.P.: supervision; B.L.: conceptualization, project administration, supervision, writing—review and editing.

All authors gave final approval for publication and agreed to be held accountable for the work performed therein.

Conflict of interest declaration. We declare we have no competing interests.

Funding. G.C., G.S. and B.L. acknowledge the support of the PNRR ICSC National Research Centre for High Performance Computing, Big Data and Quantum Computing (grant no. CN00000013), under the NRRP MUR programme funded by the NextGenerationEU. B.L. also acknowledges the support of the PNRR project FAIR-Future AI Research (grant no. PE00000013), under the NRRP MUR programme funded by the NextGenerationEU. G.C. also acknowledges the support of the European Union's Horizon 2020 research and innovation programme under Marie Skłodowska-Curie grant agreement no. 101103026. L.L. has been supported by the ERC project 'IMMUNE' (grant agreement ID: 101003183). F.B. acknowledges support from the Air Force Office of Scientific Research under award no. FA8655-22-1-7025.

Endnotes

[1]World Health Organization: https://covid19.who.int.

[2]If we extend the time span of the simulations we should consider the fact that immunity only exists for a finite period of time; however, since we are considering only around four months, considering that recovered people remain immune until the end of the simulations is a good approximation of reality.

References

1. Perra N. 2021 Non-pharmaceutical interventions during the COVID-19 pandemic: a review. *Phys. Rep.* **913**, 1–52. (doi:10.1016/j.physrep.2021.02.001)
2. Hsiang S *et al.* 2020 The effect of large-scale anti-contagion policies on the COVID-19 pandemic. *Nature* **584**, 262–267. (doi:10.1038/s41586-020-2404-8)
3. Dehning J, Zierenberg J, Spitzner FP, Wibral M, Neto JP, Wilczek M, Priesemann V. 2020 Inferring change points in the spread of COVID-19 reveals the effectiveness of interventions. *Science* **369**, eabb9789. (doi:10.1126/science.abb9789)
4. Brauner JM *et al.* 2021 Inferring the effectiveness of government interventions against COVID-19. *Science* **371**, eabd9338. (doi:10.1126/science.abd9338)
5. Soltesz K *et al.* 2020 The effect of interventions on COVID-19. *Nature* **588**, E26–E28. (doi:10.1038/s41586-020-3025-y)
6. Bank W. 2022 *Helping countries adapt to a changing world*, vol. 1. Washington, DC: World Bank Group.
7. Yabe T, Bueno BGB, Dong X, Pentland A, Moro E. 2023 Behavioral changes during the COVID-19 pandemic decreased income diversity of urban encounters. *Nat. Commun.* **14**, 2310. (doi:10.1038/s41467-023-37913-y)
8. Hunter RF, Garcia L, de Sa TH, Zapata-Diomedi B, Millett C, Woodcock J, Pentland A, Moro E. 2021 Effect of COVID-19 response policies on walking behavior in US cities. *Nat. Commun.* **12**, 3652. (doi:10.1038/s41467-021-23937-9)
9. Lucchini L, Langle-Chimal O, Candeago L, Melito L, Chunet A, Montfort A, Lepri B, Lozano-Gracia N, Fraiberger SP. 2023 Socioeconomic disparities in mobility behavior during the COVID-19 pandemic in developing countries. (http://arxiv.org/abs/2305.06888)
10. Lucchini L, Centellegher S, Pappalardo L, Gallotti R, Privitera F, Lepri B, De Nadai M. 2021 Living in a pandemic: changes in mobility routines, social activity and adherence to COVID-19 protective measures. *Sci. Rep.* **11**, 24452. (doi:10.1038/s41598-021-04139-1)
11. Seitz BM *et al.* 2020 The pandemic exposes human nature: 10 evolutionary insights. *Proc. Natl Acad. Sci. USA* **117**, 27 767–27 776. (doi:10.1073/pnas.2009787117)
12. Van de Groep S, Zanolie K, Green KH, Sweijen SW, Crone EA. 2020 A daily diary study on adolescents' mood, empathy, and prosocial behavior during the COVID-19 pandemic. *PLoS ONE* **15**, e0240349. (doi:10.1371/journal.pone.0240349)
13. Myers KR *et al.* 2020 Unequal effects of the COVID-19 pandemic on scientists. *Nat. Hum. Behav.* **4**, 880–883. (doi:10.1038/s41562-020-0921-y)
14. Probst T, Stippl P, Pieh C. 2020 Changes in provision of psychotherapy in the early weeks of the COVID-19 lockdown in Austria. *Int. J. Environ. Res. Public Health* **17**, 3815. (doi:10.3390/ijerph17113815)
15. Jeste S, Hyde C, Distefano C, Halladay A, Ray S, Porath M, Wilson R, Thurm A. 2020 Changes in access to educational and healthcare services for individuals with intellectual and developmental disabilities during COVID-19 restrictions. *J. Intellect. Disabil. Res.* **64**, 825–833. (doi:10.1111/jir.12776)
16. Khan MA, Smith JEM. 2020 'Covibesity,' a new pandemic. *Obes. Med.* **19**, 100282. (doi:10.1016/j.obmed.2020.100282)
17. Sutin AR, Luchetti M, Aschwanden D, Lee JH, Sesker AA, Strickhouser JE, Stephan Y, Terracciano A. 2020 Change in five-factor model personality traits during the acute phase of the coronavirus pandemic. *PLoS ONE* **15**, e0237056. (doi:10.1371/journal.pone.0237056)
18. Whittaker S, Frohlich D, Daly-Jones O. 1994 Informal workplace communication: what is it like and how might we support it? In *Proc. SIGCHI Conf. on Human Factors in Computing Systems, Boston, MA, USA, 24–28 April 1994*, pp. 131–137. New York, NY: ACM. (doi:10.1145/191666.191726)

19. Teasley S, Covi L, Krishnan MS, Olson JS. 2000 How does radical collocation help a team succeed? In *Proc. 2000 ACM Conf. on Computer Supported Cooperative Work, Philadelphia, PA, USA, 2–6 December 2000*, pp. 339–346. New York, NY: ACM. (doi:10.1145/358916.359005)
20. Kirkman BL, Rosen B, Tesluk PE, Gibson CB. 2004 The impact of team empowerment on virtual team performance: the moderating role of face-to-face interaction. *Acad. Manag. J.* **47**, 175–192. (doi:10.2307/20159571)
21. Salis S, Williams AM. 2010 Knowledge sharing through face-to-face communication and labour productivity: evidence from British workplaces. *Brit. J. Ind. Relat.* **48**, 436–459. (doi:10.1111/j.1467-8543.2009.00762.x)
22. Wu L, Waber BN, Aral S, Brynjolfsson E, Pentland A. 2008 Mining face-to-face interaction networks using sociometric badges: predicting productivity in an IT configuration task. In *Proc. of Int. Conf. on Information Systems, Miami, FL, USA, 7–10 December 2008*. New York, NY: ACM.
23. Pentland A. 2012 The new science of building great teams. *Harv. Bus. Rev.* **1**, 60–69.
24. Wilson HJ. 2013 Wearables in the workplace. *Harv. Bus. Rev.* **91**, 23–25. (https://hbr.org/2013/09/wearables-in-the-workplace)
25. Shirley MD, Rushton SP. 2005 The impacts of network topology on disease spread. *Ecol. Complex.* **2**, 287–299. (doi:10.1016/j.ecocom.2005.04.005)
26. Ikeda Y, Hasegawa T, Nemoto K. 2010 Cascade dynamics on clustered network. *J. Phys.: Conf. Ser.* **221**, 012005. (doi:10.1088/1742-6596/221/1/012005)
27. Hackett A, Melnik S, Gleeson JP. 2011 Cascades on a class of clustered random networks. *Phys. Rev. E* **83**, 056107. (doi:10.1103/PhysRevE.83.056107)
28. Ansari S, Anvari M, Pfeffer O, Molkenthin N, Moosavi MR, Hellmann F, Heitzig J, Kurths J. 2021 Moving the epidemic tipping point through topologically targeted social distancing. *Eur. Phys. J. Spec. Top.* **230**, 3273–3280. (doi:10.1140/epjs/s11734-021-00138-5)
29. Block P, Hoffman M, Raabe IJ, Dowd JB, Rahal C, Kashyap R, Mills MC. 2020 Social network-based distancing strategies to flatten the COVID-19 curve in a post-lockdown world. *Nat. Hum. Behav.* **4**, 588–596. (doi:10.1038/s41562-020-0898-6)
30. Willem L *et al.* 2021 The impact of contact tracing and household bubbles on deconfinement strategies for COVID-19. *Nat. Commun.* **12**, 1524. (doi:10.1038/s41467-021-21747-7)
31. Gemmetto V, Barrat A, Cattuto C. 2014 Mitigation of infectious disease at school: targeted class closure vs school closure. *BMC Infect. Dis.* **14**, 695. (doi:10.1186/s12879-014-0695-9)
32. Gauvin L, Panisson A, Barrat A, Cattuto C. 2015 Revealing latent factors of temporal networks for mesoscale intervention in epidemic spread. (http://arxiv.org/abs/1501.02758)
33. McGee RS, Homburger JR, Williams HE, Bergstrom CT, Zhou AY. 2021 Model-driven mitigation measures for reopening schools during the COVID-19 pandemic. *Proc. Natl Acad. Sci. USA* **118**, e2108909118. (doi:10.1073/pnas.2108909118)
34. Leoni E, Cencetti G, Santin G, Istomin T, Molteni D, Picco GP, Farella E, Lepri B, Murphy AL. 2022 Measuring close proximity interactions in summer camps during the COVID-19 pandemic. *EPJ Data Sci.* **11**, 5. (doi:10.1140/epjds/s13688-022-00316-y)
35. Valdano E, Poletto C, Boëlle P-Y, Colizza V. 2021 Reorganization of nurse scheduling reduces the risk of healthcare associated infections. *Sci. Rep.* **11**, 7393. (doi:10.1038/s41598-021-86637-w)
36. Gleeson JP. 2008 Cascades on correlated and modular random networks. *Phys. Rev. E* **77**, 046117. (doi:10.1103/PhysRevE.77.046117)
37. Nematzadeh A, Ferrara E, Flammini A, Ahn Y-Y. 2014 Optimal network modularity for information diffusion. *Phys. Rev. Lett.* **113**, 088701. (doi:10.1103/PhysRevLett.113.088701)
38. Peng H, Nematzadeh A, Romero DM, Ferrara E. 2020 Network modularity controls the speed of information diffusion. *Phys. Rev. E* **102**, 052316. (doi:10.1103/PhysRevE.102.052316)
39. Sapiezynski P, Stopczynski A, Lassen DD, Lehmann S. 2019 Interaction data from the Copenhagen Networks Study. *Sci. Data* **6**, 315. (doi:10.1038/s41597-019-0325-x)
40. Antelmi A, Cordasco G, Scarano V, Spagnuolo C. 2021 Modeling and evaluating epidemic control strategies with high-order temporal networks. *IEEE Access* **9**, 140 938–140 964. (doi:10.1109/ACCESS.2021.3119459)
41. Meidan D, Schulmann N, Cohen R, Haber S, Yaniv E, Sarid R, Barzel B. 2021 Alternating quarantine for sustainable epidemic mitigation. *Nat. Commun.* **12**, 220. (doi:10.1038/s41467-020-20324-8)
42. Cowan R, Jonard N. 2004 Network structure and the diffusion of knowledge. *J. Econ. Dyn. Control* **28**, 1557–1575. (doi:10.1016/j.jedc.2003.04.002)
43. Ok J, Jin Y, Shin J, Yi Y. 2014 On maximizing diffusion speed in social networks: impact of random seeding and clustering. *ACM SIGMETRICS Perform. Eval. Rev.* **42**, 301–313. (doi:10.1145/2637364.2591991)
44. Lambiotte R, Panzarasa P. 2009 Communities, knowledge creation, and information diffusion. *J. Informetrics* **3**, 180–190. (doi:10.1016/j.joi.2009.03.007)
45. Chung K, Baek Y, Kim D, Ha M, Jeong H. 2014 Generalized epidemic process on modular networks. *Phys. Rev. E* **89**, 052811. (doi:10.1103/PhysRevE.89.052811)
46. She B, Liu J, Sundaram S, Paré PE. 2022 On a networked SIS epidemic model with cooperative and antagonistic opinion dynamics. *IEEE Trans. Control Netw. Syst.* **9**, 1154–1165. (doi:10.1109/TCNS.2022.3145748)
47. Peng K, Lu Z, Lin V, Lindstrom MR, Parkinson C, Wang C, Bertozzi AL, Porter MA. 2021 A multilayer network model of the coevolution of the spread of a disease and competing opinions. *Math. Models Methods Appl. Sci.* **31**, 2455–2494. (doi:10.1142/S0218202521500536)
48. Anderson RM, May RM. 1991 *Infectious diseases of humans: dynamics and control*. New York, NY: Oxford University Press.
49. Pastor-Satorras R, Castellano C, Van Mieghem P, Vespignani A. 2015 Epidemic processes in complex networks. *Rev. Mod. Phys.* **87**, 925. (doi:10.1103/RevModPhys.87.925)
50. Dodds PS, Watts DJ. 2004 Universal behavior in a generalized model of contagion. *Phys. Rev. Lett.* **92**, 218701. (doi:10.1103/PhysRevLett.92.218701)
51. Lucas M, Iacopini I, Robiglio T, Barrat A, Petri G. 2023 Simplicially driven simple contagion. *Phys. Rev. Res.* **5**, 013201. (doi:10.1103/PhysRevResearch.5.013201)
52. Holland PW, Laskey KB, Leinhardt S. 1983 Stochastic blockmodels: first steps. *Soc. Netw.* **5**, 109–137. (doi:10.1016/0378-8733(83)90021-7)
53. Ferretti L, Wymant C, Kendall M, Zhao L, Nurtay A, Abeler-Dörner L, Parker M, Bonsall D, Fraser C. 2020 Quantifying SARS-CoV-2 transmission suggests epidemic control with digital contact tracing. *Science* **368**, eabb6936. (doi:10.1126/science.abb6936)
54. He X *et al.* 2020 Temporal dynamics in viral shedding and transmissibility of COVID-19. *Nat. Med.* **26**, 1491–1493. (doi:10.1038/s41591-020-1016-z)
55. Cencetti G, Santin G, Longa A, Pigani E, Barrat A, Cattuto C, Lehmann S, Salathe M, Lepri B. 2021 Digital proximity tracing on empirical contact networks for pandemic control. *Nat. Commun.* **12**, 1655. (doi:10.1038/s41467-021-21809-w)
56. Barman MP, Rahman T, Bora K, Borgohain C. 2020 COVID-19 pandemic and its recovery time of patients in India: a pilot study. *Diabetes Metab. Syndr.: Clin. Res. Rev.* **14**, 1205–1211. (doi:10.1016/j.dsx.2020.07.004)
57. Van den Driessche P. 2017 Reproduction numbers of infectious disease models. *Infect. Dis. Modell.* **2**, 288–303. (doi:10.1016/j.idm.2017.06.002)
58. Cencetti G, Lucchini L, Santin G, Battiston F, Moro E, Pentland A, Lepri B. 2023 Data from: Temporal clustering of social interactions trades-off disease spreading and knowledge diffusion. GitHub repository. (https://github.com/giuliacencetti/Social˙bubbles)
59. Cencetti G, Lucchini L, Santin G, Battiston F, Moro E, Pentland A, Lepri B. 2023 Temporal clustering of social interactions trades-off disease spreading and knowledge diffusion. Figshare. (doi:10.6084/m9.figshare.c.6961112)