Strong Dependence of Infection Profiles on Grouping Dynamics during Epidemiological Spreading

Zhenyuan Zhao[1], Guannan Zhao[1], Chen Xu[2],
Pak Ming Hui[3], and Neil F. Johnson[1,*]

[1] Physics Department, University of Miami, Coral Gables, FL 33146, USA
njohnson@physics.miami.edu
[2] Department of Physics, Suzhou University,
Suzhou 215006, People's Republic of China
[3] Department of Physics, Chinese University of Hong Kong,
Shatin, Hong Kong, China

Abstract. The spreading of an epidemic depends on the connectivity of the underlying host population. Because of the inherent difficulties in addressing such a problem, research to date on epidemics in networks has focused either on static networks, or networks with relatively few rewirings per timestep. Here we employ a simple, yet highly non-trivial, model of dynamical grouping to investigate the extent to which the underlying dynamics of tightly-knit communities can affect the resulting infection profile. Individual realizations of the spreading tend to be dominated by large peaks corresponding to infection resurgence, and a generally slow decay of the outbreak. In addition to our simulation results, we provide an analytical analysis of the run-averaged behaviour in the regime of fast grouping dynamics. We show that the true run-averaged infection profile can be closely mimicked by employing a suitably weighted static network, thereby dramatically simplifying the level of difficulty.

Keywords: complex systems, networks, epidemics, group dynamics.

1 Introduction

The way in which a given virus, idea, rumour or activity spreads throughout a population, will depend on the underlying connectivity structure of the network describing the population. If the population is sufficiently well connected that it can be considered well-mixed, then the standard approximations of mass-action epidemiology should apply [1]. Much research in the complex systems literature has focused on understanding spreading phenomena in static networks with limited connectivity [1,2,3]. More recently, researchers have begun to relax the assumptions of a static network, by allowing rewirings from one timestep to the next [4,5,6,7,8].

The implicit complexity of competing time and length scales, means that there is still an outstanding challenge to describe spreading phenomena in networks in which there may be an arbitrary number of rewirings per timestep, and where communities

* Corresponding author.

J. Zhou (Ed.): Complex 2009, Part I, LNICST 4, pp. 960–970, 2009.

can abruptly change their identity and number in time. This is particularly important since dynamical grouping processes, in which groups coalesce and fragment over time, are thought to occur widely in biological and social systems [9,10,11,12,13].

In this paper, we focus on this problem by considering a variety of spreading processes in a population which is characterized by a particularly simple, but highly non-trivial, dynamical grouping mechanism [11]. Our focus is to understand how this underlying group dynamics (GD) affects the spreading process for a number of standard epidemiological models. The shift from grouping models to dynamical networks, can be achieved simply by allowing the network nodes to represent individuals within the population – and then specifying that any two individuals have a strong link whenever they are within the same group, but have a negligibly weak link when they are in different groups. Since the groupings change in time, so too will the links between them. Most importantly, as groups break up (fragment) or join together (coalesce), a variable number of links will be broken or formed. At the two extremes, a very large group of size $\sim N$ breaking up will represent the net loss of $\sim N(N-1)/2$ links in a given timestep, while two groups of size $\sim N/2$ joining together will represent an increase in the number of links from $\sim 2.\frac{N}{2}(\frac{N}{2}-1)/2$ to $\sim N(N-1)/2$. This shift in descriptive emphasis from groups to networks also means that the group size distribution $n(k)$, which is proportional to the probability that a randomly chosen node is in a group of size k, is simply related to the network connectivity distribution $P(k)$, which is the probability that a randomly chosen node has k contacts.

Our interest in this paper is confined to problems in which there is no obvious spatial component. In other words, instead of dealing with the well-studied patch setups typified by the spread of a real virus between geographically distinct regions [1], we are interested in populations within which the connections have a more general and variable topology – such as communities on the Internet, or in a financial market – and hence are arguably better described by general coalescence and fragmentation probabilities with no direct spatial interpretation. With these applications in mind, we employ a generalization of a simple but highly non-trivial model of grouping dynamics that exhibits a power-law distribution of group sizes with an exponential cutoff [11,13]. In addition to being qualitatively consistent with a variety of real-world social phenomena, including financial market behaviour [11,13], the model has the beauty of just featuring two parameters: the probability that a chosen group fragments in a given timestep ν_{frag}, and the probability that it instead coalesces with another group ν_{coal} [11,13]. We consider the dynamics of several different types of epidemic model: SIR, SIS, SIRS. Using standard terminology [1], individual nodes are either **S**usceptible, **I**nfected or **R**ecovered. There are two particular forms of infection process $S \to I$ which one might reasonably consider using. The linear form states that the susceptible node gets infected with rate $i.p$ if it is connected to i infected nodes; the more correct nonlinear form says that if it is connected to i infected nodes, the probability will be $1-(1-p)^i$ [3]. We use this more correct latter version in our simulations, noting that for $p \ll 1$ they tend to converge as can be seen from a simple binomial expansion of the nonlinear form.

In the limit of a well-mixed population, the mass-action differential equations for all these standard epidemic processes can be written in the following compact form:

$$\begin{aligned}\frac{dS}{dt} &= \mu - (1-(1-p)^I)S - \mu S + \gamma I + \omega R \\ \frac{dI}{dt} &= (1-(1-p)^I)S - qI - \mu I - \gamma I \\ \frac{dR}{dt} &= qI - \mu R - \omega R\,.\end{aligned} \tag{1}$$

where p is the infection rate, q is the recovery rate, μ is birth and death rate, γ is the rate of transmission from I to S, and ω is the rate of transmission from R to S.

We now discuss the initial setup for the simulations. Prior to the initial infection, we suppose that the grouping dynamics have been underway for a sufficiently long time that the system has reached its steady state. At some arbitrary time which we call $t = 0$, the largest group is selected and an arbitrary individual in this group becomes infected. This mimics a situation of real-world relevance, such as a financial market [11,13], in which there is a power-law steady-state distribution of group sizes or communities, and then one member of the population gets infected with a virus. Comparing the time scales for epidemic spreading and grouping dynamics, our investigations show that there are three broad regimes of subsequent spreading behaviour:

(1) The grouping dynamics are much slower than the epidemic spreading: The virus then tends to remain within the initial group. The spreading dynamics therefore become equivalent to mass-action spreading in a population whose size is equal to the initial group size. The infection is oblivious to the larger population of susceptibles. This large population of susceptibles is granted an accidental but effective immunity, by happening to be members of different groups at the right time.
(2) The grouping dynamics are of comparable timescale to the epidemic spreading: The grouping dynamics can now play a significant role in suppressing or amplifying the spreading, because they will determine the contact condition for the infected nodes at any given timestep. Due to the existence of relatively large groups which can change their membership on a timescale similar to the spreading, a rather generic spiky infection profile $I(t)$ tends to emerge.
(3) The grouping dynamics are much faster than the epidemic spreading: The virus now sees a time-average of the network's connectivity, where this average is taken over some suitably large time-window. In the limit of very fast grouping dynamics, this static, time-averaged network will correspond to a fully-connected weighted network, in stark contrast to the sparse, disconnected network which is characteristic of individual timesteps. Section 3 explains how we have managed to develop an analytical analysis in this regime, based on a suitably averaged measure of the grouping dynamics.

2 Epidemics within a Population Undergoing Dynamical Grouping

Since the dynamics describing the spreading of an infectious disease will depend on the contact structure which underlies the population, our simulations and analysis are chosen to highlight this interplay between epidemic spreading and the grouping dynamics – in particular, the virus is able to spread within groups but not between groups. As

a result of the grouping dynamics, the infectious nodes can themselves move between groups as a result of the fragmentation-coalescence process, and hence propagate the virus. As mentioned earlier, we wish to mimic aspects of modern-day human interaction, e.g. long distance travel and communication, as well as rapidly evolving social networks in cyberspace, hence the grouping dynamics we use is independent of any spatial length scale. In addition, the model allows a group's size and membership to change over time – moreover, groups may disappear or form in time and hence the total number of groups may change in time. We consider a fixed population N of individual objects (e.g. people). At each timestep, a group or cluster is chosen by random-picking a node among the N possible nodes, and then looking to see to which group this node belongs at that timestep. In other words, groups are picked proportional to their group size, reflecting the fact that any one individual is equally likely to initiate a group-formation or group-breakup event. With probability ν_{frag}, the group fragments, while with probability ν_{coal} a second group is picked in a similar way, to join with the first one [11,13]. Naturally, $\nu_{frag} + \nu_{coal} \leq 1$.

2.1 SIR Process within the Dynamical Grouping Model

Here we consider the results for the basic SIR process passing through the dynamical grouping model. In the language of the well-mixed mass-action equations of Eq. 1, this process is described by setting μ, γ, and ω equal to zero (see Fig. 1(a)).

Figures 1(b) and (c) show the fraction of the population N who are infected at time t, obtained in a single run of the simulation. It is remarkable that even though the profile looks reasonably standard in Fig. 1 (b), the profile in Fig. 1(c) is completely different in character. Most noticeably, Fig. 1(c) features multiple peaks and a far slower decay than that of Fig. 1(b). Indeed, the overall decay of infection in Fig. 1(b) is of order 20 timesteps while that in Fig. 1(c) is 2000, which is two orders of magnitude larger. Despite the fact that there are only four parameters – the p and q values of the SIR process together with ν_{frag} and ν_{coal} for the dynamical grouping – our exploration of the parameter space shows that it is easy to produce a broad range of qualitatively different behaviors [14]. Common to many of the runs, are the features of multiple, large peaks and a very slow, possibly fat-tailed decay. These $I(t)$ profiles cannot be reproduced in general by the well-mixed SIR model, either in its deterministic form (Eq. 1) or in stochastic simulation. Instead, the changing contact structure acts to drive these large fluctuations and slow decay. Even if we employ a weighted network (Fig. 1(d)) as described in Section 3, the individual run profiles $I(t)$ *and* the run-averaged behaviour $I(t)_{\mathrm{RunAvge}}$ differ significantly from it.

2.2 SIS Process within the Dynamical Grouping Model

We now turn to the standard SIS process (Fig. 2(a)) in which any infected individual recovers to the susceptible state – in other words, there is no immunity conferred by having had the disease [1]. To analyze this case in the well-mixed limit corresponding to the mass-action equations of Eq. 1, we set q, μ and ω equal to zero.

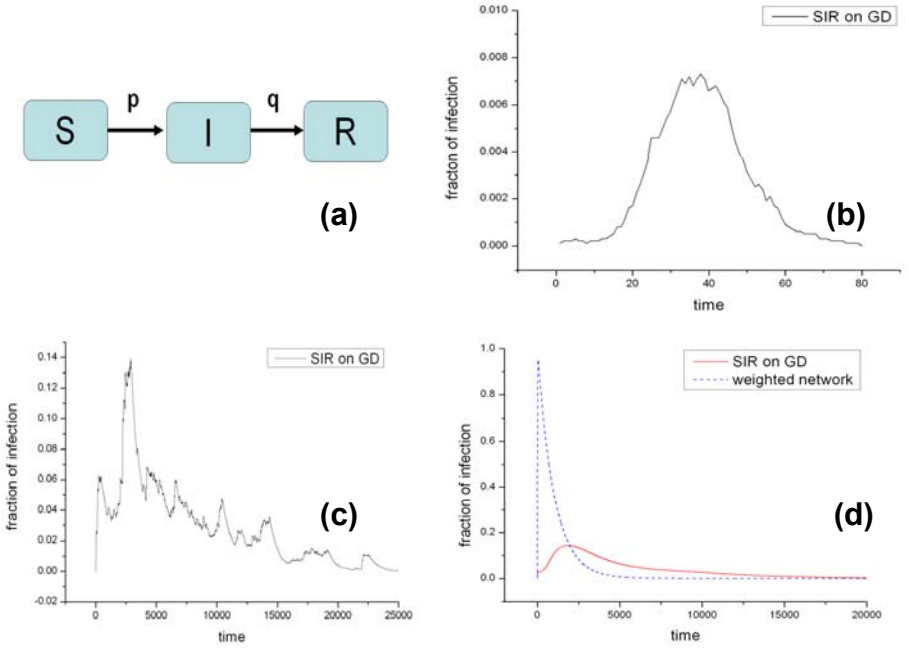

Fig. 1. SIR process within the dynamical grouping model. (a) Schematic of the SIR process. (b) An individual run showing the infection profile $I(t)$ under GD, which corresponds to the fraction of the entire population N who are infected at time t. Here $N = 10^4$, $\nu_{frag} = 0.001$, $\nu_{coal} = 0.99$, $p = 0.001$ and $q = 0.1$. (c) An individual run showing $I(t)$ for $N = 10^4$, $\nu_{frag} = 0.01$, $\nu_{coal} = 0.9$, $p = 0.001$ and $q = 0.001$. (d) Solid red curve shows the run-averaged profile $I_{\mathrm{RunAvge}}(t)$, obtained by averaging $I(t)$ over many runs. Dashed blue curve shows the run-averaged $I(t)$ for a *static* weighted network, where nodes i and j are connected by a time-independent link of strength $P_{i,j}$. The same parameters are used as in panel (c).

The finite value of $I(t)$ in Fig. 2(b) corresponds to an endemic equilibrium, and is reached after a slow growth. Both these features arise as a direct result of the suppressing effect of the grouping dynamics. Specifically, the grouping dynamics only allows a small fraction of the entire population N to be in contact with an infected individual at a given timestep. The remainder of the population are isolated from the infection since they are located in groups comprising entirely susceptibles. As a result, before any infected node can contact the rest of the population and infect them, it gets turned back to susceptible.

Comparing Figs. 2(b) and (c) shows how changing the probabilities γ and ν_{frag}, affects the stability of the $I(t)$ curve in the steady state. Increasing γ aids in producing large jumps because it replenishes the pool of susceptibles. Likewise reducing the fragmentation probability ν_{frag} increases the effective timescale over which groups break up – this in turn allows larger spikes to appear by allowing time for significant decay between jumps. Once again, the weighted network cannot reproduce the run-averaged behavior, as evidenced in Fig. 2(d).

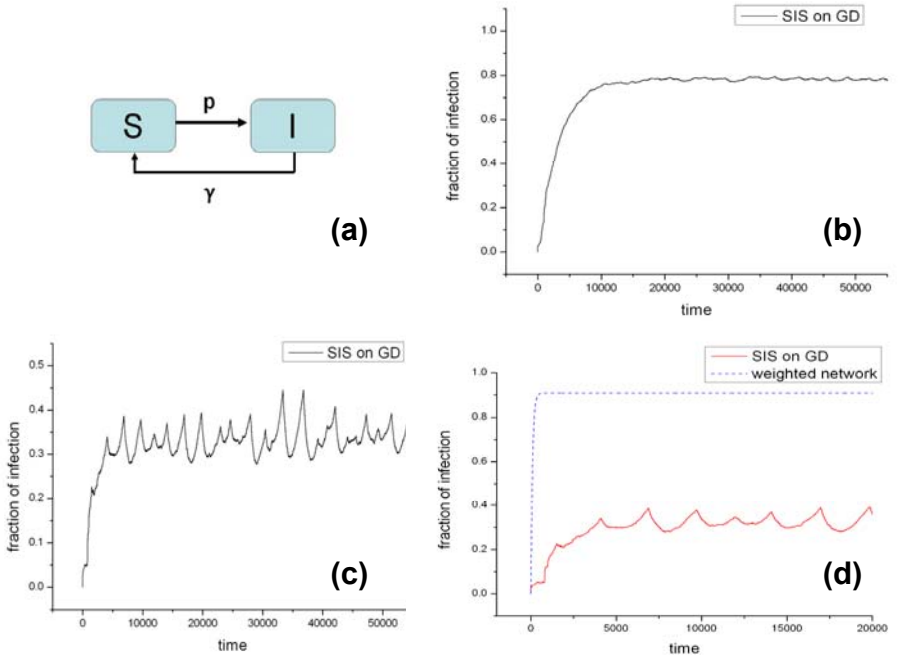

Fig. 2. SIS process within the dynamical grouping model. (a) Schematic of the SIS process. (b) An individual run showing the infection profile $I(t)$ with $N = 10^4$, $\nu_{frag} = 0.01$, $\nu_{coal} = 0.99$, $p = 0.01$ and $\gamma = 0.0001$. (c) An individual run showing $I(t)$ with $N = 10^4$, $\nu_{frag} = 0.001$, $\nu_{coal} = 0.99$, $p = 0.01$ and $\gamma = 0.001$. (d) Solid red curve shows the run-averaged profile $I_{\mathrm{RunAvge}}(t)$, obtained by averaging $I(t)$ over many runs. Dashed blue curve shows the run-averaged $I(t)$ for a *static* weighted network, where nodes i and j are connected by a time-independent link of strength $P_{i,j}$. The same parameters are used as in panel (c).

2.3 SIRS Process within the Dynamical Grouping Model

The SIRS process is summarized in Fig. 3(a), and corresponds to setting μ and γ equal to zero in Eq. 1. As can be seen in Figs. 3(b) and (c) it is qualitatively similar to SIS. However, the presence of a temporary, intermediate R-state (i.e. waning immunity) allows the dynamics to build up even stronger decays – hence the observable peaks appear even stronger as compared to the SIS case.

2.4 Demographic SIR Process within the Dynamical Grouping Model

Finally we discuss the introduction of demography into the SIR model, which we call SIRD. This epidemiological process corresponds to setting γ and ω equal to zero in Eq. 1 – in addition, we assume that the natural mortality is μ. In other words, each individual has a lifespan given by $1/\mu$ [1]. In order to keep the total population constant $S + I + R = N$, μ is also taken to represent the birth rate within the population (Fig. 4(a)). Figure 4(b) shows that when the birth-death processes are much slower than the infection and recovery processes, it is effectively the same as the SIR model

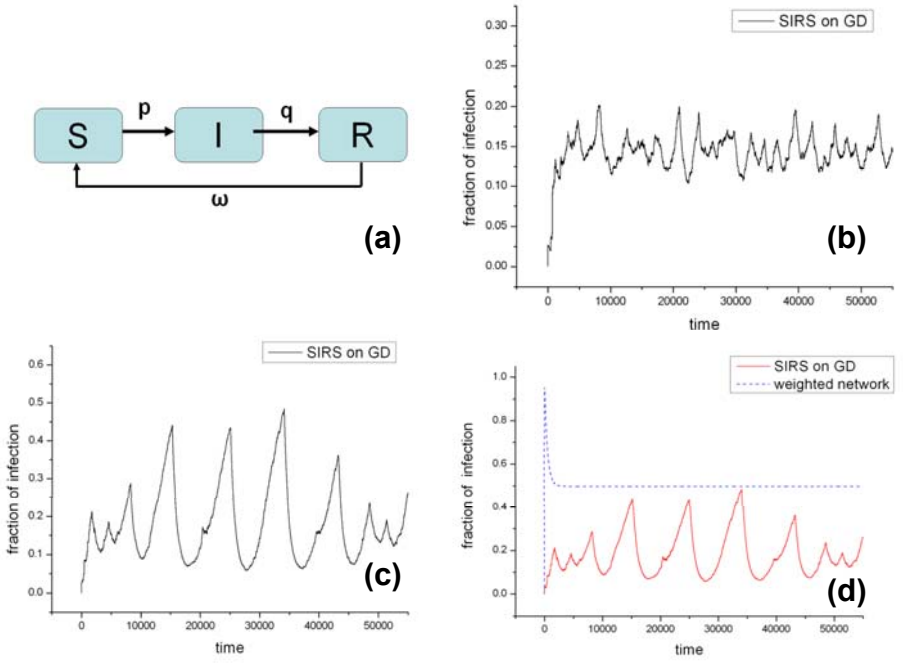

Fig. 3. SIRS process within the dynamical grouping model. (a) Schematic of the SIRS process. (b) An individual run showing the infection profile $I(t)$ with $N = 10^4$, $\nu_{frag} = 0.01$, $\nu_{coal} = 0.99$, $p = 0.01$, $q = 0.001$ and $\omega = 0.001$. (c) An individual run showing $I(t)$ with $N = 10^4$, $\nu_{frag} = 0.001$, $\nu_{coal} = 0.99$, $p = 0.01$, $q = 0.001$ and $\omega = 0.001$. (d) Solid red curve shows the run-averaged profile $I_{\mathrm{RunAvge}}(t)$, obtained by averaging $I(t)$ over many runs. Dashed blue curve shows the run-averaged $I(t)$ for a *static* weighted network, where nodes i and j are connected by a time-independent link of strength $P_{i,j}$. The same parameters are used as in panel (c).

within the dynamical grouping model, as discussed earlier. The associated endemic equilibrium is shown not to be stable in Fig. 4(b), however this stability returns by increasing μ, as shown by comparing Fig. 4(b) to Fig. 4(c). Again, the results for the run-averaged case cannot be reproduced by the static weighted network, as evidenced by Fig. 4(d).

3 Regime of Fast Grouping Dynamics

So far, we have established that new features occur in the $I(t)$ profile both for individual runs, and when run-averaged. In particular, the $I(t)$ curve typically has more peaks and has a slower general decay – in short, it appears more bursty or 'noisy', however we emphasize that this 'noise' is actually meaningful (and interpretable) dynamics within our model, resulting from the intrinsic group formation and breakup processes of coalescence and fragmentation. Even though other group fusion-fission mechanisms could be introduced and the same procedure followed to investigate the various epidemiological processes, we expect the features that we have found to remain qualitatively similar.

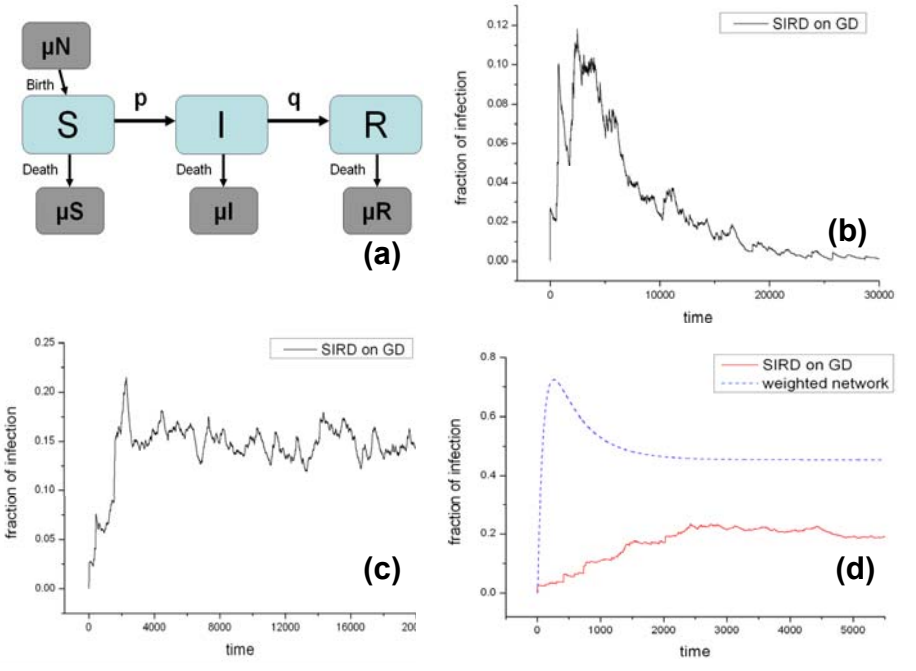

Fig. 4. SIR process with demography, within the dynamical grouping model. (a) Schematic of the SIRS process. (b) An individual run showing the infection profile $I(t)$ with $N = 10^4$, $\nu_{frag} = 0.01$, $\nu_{coal} = 0.99$, $p = 0.01$, $q = 0.001$ and $\mu = 0.000001$. (c) An individual run showing $I(t)$ with $N = 10^4$, $\nu_{frag} = 0.01$, $\nu_{coal} = 0.99$, $p = 0.01$, $q = 0.001$ and $\mu = 0.001$. (d) Solid red curve shows the run-averaged profile $I_{\mathrm{RunAvge}}(t)$, obtained by averaging $I(t)$ over many runs. Dashed blue curve shows the run-averaged $I(t)$ for a *static* weighted network, where nodes i and j are connected by a time-independent link of strength $P_{i,j}$. The same parameters are used as in panel (b), but with $\mu = 0.001$.

This is indeed exactly what we find when, for example, we consider another popular dynamical grouping mechanism such as that introduce by Levin [9].

Looking back at the first case that we discussed, of the basic SIR process within the dynamical grouping model, there are essentially four timescales: the group fragmentation timescale $\tau_{frag} \sim \nu_{frag}^{-1}$, the group coalescence timescale $\tau_{coal} \sim \nu_{coal}^{-1}$, the infection timescale $\tau_p \sim p^{-1}$, and an individual's recovery timescale $\tau_q \sim q^{-1}$. The first two timescales correspond to the timescale of the grouping dynamics, while the latter two are instead related to the epidemic process itself. Although there are only four parameters, and they are all physically meaningful, this is still a very large parameter space to analyze – hence for brevity, we will here focus on the regime of fast grouping dynamics, i.e. the first two timescales τ_{frag} and τ_{coal} are much smaller than $\tau_p \sim p^{-1}$ and $\tau_q \sim q^{-1}$.

With fast group breaking and merging processes, one can imagine that the grouping dynamics would manage to refresh or reshuffle the whole population within an infection/recovery period. In other words, the heterogenous population would become

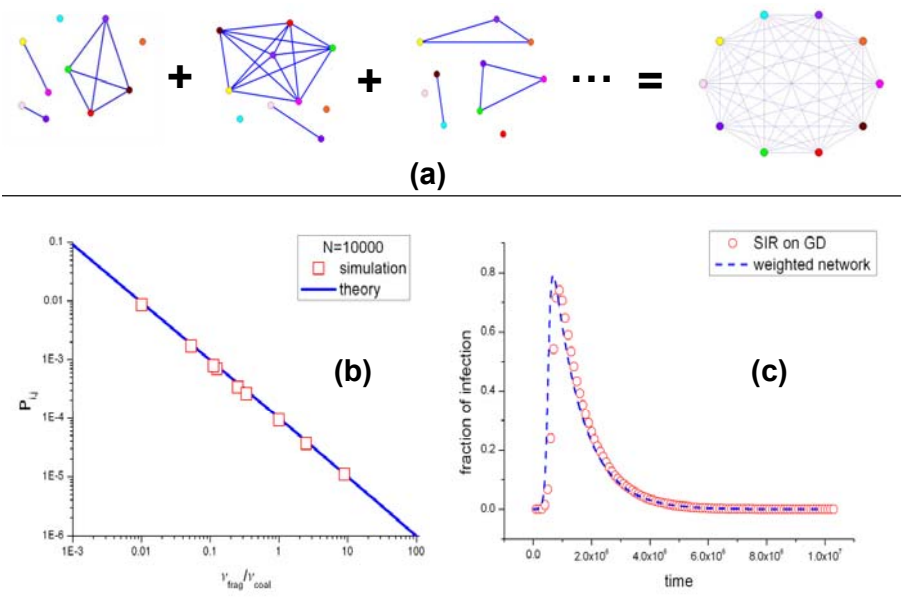

Fig. 5. Regime of fast grouping dynamics. (a) Schematic showing the aggregation of the instantaneous network links, to form a fully connected, weighted network. (b) Comparison between our analytic result for $P_{i,j}$ and numerical results obtained directly from the simulation. (c) Results for the run-averaged infection profile I_{RunAvge} obtained by simulation (red circles) and using the generalized mass-action partial differential equations of Eq.5. Here $\nu_{frag} = 0.05$, $\nu_{coal} = 0.95$, $p = 10^{-6}$, and $q = 10^{-6}$.

effectively homogenous as far as the spread of epidemics is concerned. In this limit, we find that we can capture the features of the run-averaged dynamics using a suitably averaged quantity to capture the average network properties. With this purpose in mind, we focus on a probability $P_{i,j}$ corresponding to the averaged probability that two random nodes i and j are connected. Figure 5(a) shows the effect of aggregating the network dynamics over some sufficiently large time-window that it eventually becomes a fully connected, weighted network. $P_{i,j}$ is the resulting weight of the links, which for our grouping dynamics then satisfies the following master equation:

$$\frac{dP_{i,j}}{dt} = -P_{i,j}^3 \frac{1}{N} \nu_{frag} + (1 - P_{i,j}) \frac{1}{N^2} P_{i,j}^2 \nu_{coal} \ , \tag{2}$$

where we have assumed that all $P_{i,j}$'s are equal. The first term on the right-hand side of the master equation corresponds to the fragmentation of a group, while the second term corresponds to the coalescence of two groups into one larger group. Solving Eq. 2 by setting the left-hand side equal to zero, yields:

$$P_{i,j} = \frac{1}{1 + N \frac{\nu_{frag}}{\nu_{coal}}} \ . \tag{3}$$

This expression for $P_{i,j}$ fits remarkably well with simulation results, as evidenced by the excellent agreement in Fig. 5(b). In the limit $N\nu_{frag} \gg \nu_{coal}$, Eq. 3 reduces to the more approximate form

$$P_{i,j} = \frac{\nu_{coal}}{N\nu_{frag}} \quad . \tag{4}$$

Using this result, the effective infection rate becomes $p \cdot P_{i,j}$ instead of p. In this particular regime, the mass-action partial differential equations from Eq. (1) can be modified to the following approximate form for the SIR process in the presence of a dynamically evolving network:

$$\begin{aligned} \frac{dS}{dt} &= -pP_{i,j}SI \\ \frac{dI}{dt} &= pP_{i,j}SI - qI \\ \frac{dR}{dt} &= qI \quad . \end{aligned} \tag{5}$$

Because of the small values of the epidemic parameters (i.e. small p, q, μ, γ, and ω) we can use the linear form for the infection process. The situation therefore becomes a virus spreading on top of a static weighted network, where all the nodes are fully connected with each other and where the links all have a uniform strength, $P_{i,j}$. Figure 5(c) shows that a good fit can be achieved for the SIR process. As a result, the basic reproductive ratio R_0 should instead be written as R_0^*, with $p \cdot P_{i,j}$ replacing p. For a large population $N \to \infty$, we therefore obtain $R_0^* = \frac{p\nu_{coal}}{qN\nu_{frag}}$.

4 Discussion

We have addressed the fascinating theoretical, yet practically relevant, question of how a population's underlying grouping dynamics might affect epidemiological spreading processes. We have considered the specific case of a coalescence-fragmentation grouping model subject to various standard epidemiological processes. This led to the appearance of multiple peaks and a slower overall decay, as compared to the usual mass-action limit for well-mixed populations. Such features are not uncommon in empirical infection data both in real viral infections, and the profile of activity in online communities [15]. For the specific regime of fast group dynamics, we were able to provide a modified mathematical model to explain the run-averaged infection profile.

We hope that our work provides useful insight and additional motivation to other researchers in the field of complex networks and complex systems, in the common quest to develop a general theory of epidemic spreading in the presence of arbitrarily complex population dynamics – in particular, in the presence of internal group formation and breakup processes.

P.M.H. acknowledges the support of a grant CUHK-401005 from the Research Grants Council of the Hong Kong SAR Government.

References

1. Keeling, M.J., Rohani, P.: Modeling Infectious Diseases in Humans and Animals. Princeton University Press, New York (2007)
2. Pastor-Satorras, R., Vespignani, A.: Epidemic dynamics in finite size scale-free networks. Phys. Rev. E 65, 035108–035112 (2002)

3. Petermann, T., Rios, P.D.L.: The role of clustering and gridlike ordering in epidemic spreading. Phys. Rev. E 69, 066116 (2004)
4. Watts, D.J., Muhamad, R., Medina, D.C., Dodds, P.S.: Multiscale, resurgent epidemics in a hierarchical metapopulation model. Proc. Natl. Acad. of Sci. 102, 11157–11162 (2005)
5. Gross, T., Dommar, C., Blasius, B.: Epidemic dynamics on an adaptive network. Phys. Rev. Lett. 96, 20–23 (2006)
6. Gross, T., Blasius, B.: Adaptive Coevolutionary Networks: A Review. J. R. Soc. Interface 5, 259–271 (2008)
7. Colizza, V., Vespignani, A.: Invasion threshold in heterogenous metapopulation networks. Phys. Rev. Lett. 99, 148701–148705 (2007)
8. Shaw, L.B., Schwartz, I.B.: Noise induced dynamics in adaptive networks with applications to epidemiology. E-print arXiv:0807.3455 on xxx.lanl.gov
9. Gueron, S., Levin, S.A.: The dynamics of group formation. Mathematical Biosciences 128, 243–246 (1995)
10. Gonzalez, M.C., Hidalgo, C.A., Barabási, A.-L.: Understanding individual human mobility pattens. Nature 453, 779–782 (2008)
11. Eguíluz, V.M., Zimmermann, M.G.: Transmission of information and herd behaviour: an application to financial markets. Phys. Rev. Lett. 85, 5659–5662 (2000)
12. McDonald, M., Suleman, O., Williams, S., Howison, S., Johnson, N.F.: Impact of unexpected events, shocking news, and rumors on foregin exchange market dynamics. Phys. Rev. E 77, 046110–046122 (2008)
13. Johnson, N.F., Jefferies, P., Hui, P.M.: Financial Market Complexity. Oxford University Press, Oxford (2003)
14. Zhao, Z., Calderon, J.P., Xu, C., Hui, P.M., Johnson, N.F.: (in preparation)
15. Sornette, D., Deschâtres, F., Gilbert, T., Ageon, Y.: Endogenous Versus Exogenous Shocks in Complex Networks: An Empirical Test Using Book Sale Rankings. Phys. Rev. Lett. 93, 228701 (2004)