Social fluidity mobilizes contagion in human and animal populations

Ewan Colman1,2, Vittoria Colizza3, Ephraim M. Hanks4, David P. Hughes5, and Shweta Bansal1

1Department of Biology, Georgetown University, Washington DC, United States
2Roslin Institute, University of Edinburgh, Easter Bush, Midlothian, United Kingdom
3INSERM, Sorbonne Université, Institut Pierre Louis d’Épidémiologie et de Santé Publique (IPLESP UMR5136), F75012, Paris, France.
4Department of Statistics, Eberly College of Science, Penn State University, State College, United States
5Department of Entomology, College of Agricultural Sciences, Penn State University, State College, United States

Humans and other group-living animals tend to distribute their social effort disproportionately. Individuals predominantly interact with a small number of close companions while maintaining weaker social bonds with less familiar group members. By incorporating this behaviour into a mathematical model we find that a single parameter, which we refer to as social fluidity, controls the rate of social mixing within the group. We compare the social fluidity of 13 species by applying the model to empirical human and animal social interaction data. To investigate how social behavior influences the likelihood of an epidemic outbreak we derive an analytical expression of the relationship between social fluidity and the basic reproductive number of an infectious disease. For highly fluid social behaviour disease transmission is revealed to be density-dependent. For species that form more stable social bonds, the model describes frequency-dependent transmission that is sensitive to changes in social fluidity.

Social behavior is fundamental to the survival of many species. It allows the formation of social groups providing fitness advantages from greater access to resources and better protection from predators [1]. Structure within these groups can be found in the way individuals communicate across space, cooperate in sexual or parental behavior, or clash in territorial or mating conflicts [2]. While animal societies are usually studied independently of each other, studying their differences has potential to reveal new insights into the nature of social living [3, 4].

When social interaction requires shared physical space it can also be a conduit for the transmission of infectious disease [5]. In a typical infectious disease model, if the disease spreads through the environment then the transmission rate is assumed to scale proportionally to the local population density [6, 7]. Alternatively, if transmission requires close proximity encounters that only occur between bonded individuals, then we expect social connectivity to determine the outcome. These two paradigms are known in the literature as density-dependence and frequency-dependence [8].

The problem, however, is that real diseases are not so easy to categorize [9]. For example, as social groups grow in size, new bonds must be created to maintain cohesiveness [10]. To manage the time and cognitive effort required to create these bonds, individuals tend to interact mostly with a small number of close companions while maintaining cohesion with the wider group through less frequent contact [11–13]. For an infectious disease, this creates fewer transmission opportunities than we would expect to see in a group with highly fluid social dynamics. The extent to which group size amplifies the transmission rate therefore depends on how individuals choose to distribute their social effort between strong and weak ties [14].

The purpose of this study is to address two questions. Firstly, can we quantify the variability in how individuals choose to distribute their social effort within a group, and secondly, what will this tell us about the effect that population density has on disease transmission?

There is growing evidence for the disproportionate distribution of social effort in human telecommunication [15–18]. Quantifying this aspect of sociality in animal systems, however, has been held back by the limitations of the data. One challenge, which we address here, is the bias introduced by variation in activity levels across the social group [19]. Additionally, while heterogeneous interaction frequencies and temporal dynamics have become common in epidemiological models [20, 21], little has been done to incorporate the variability in how the individual chooses to expend their social effort.

In the first part of this paper we introduce a mathematical model founded on the concept of social fluidity which we define as variability in the amount of social effort the individual invests in each member of their social group. Using openly available data, we estimate the social fluidity of 57 human and animal social systems. In the second part we derive an expression for the basic reproductive number of an infectious disease in the social fluidity model and demonstrate its accurately in predicting simulated outcomes. Furthermore, social fluidity emerges as a coherent mathematical framework providing the smooth connection between density-dependent and frequency-dependent disease systems.

Characterizing social behaviour

Our first objective is to measure social behaviour in a range of human and animal populations. We start by introducing a model that captures a hidden element of social dynamics: how individual group members distribute their social effort. We mathematically describe the relationships between social variables that are routinely found in studies of
animal behavior, the number of social ties and the number of interactions observed, and apply the model to empirical data to reveal behavioural differences between several species.

Social behavior model Consider a closed system of $N$ individuals and a set of interactions between pairs of individuals that were recorded during some observation period. These observations can be represented as a network: each individual, $i$, is a node; an edge exists between two nodes $i$ and $j$ if at least one interaction was observed between them; the edge weight, $w_{i,j}$, denotes the number of times this interaction was observed. The total number of interactions of $i$ is denoted strength, $s_i = \sum_{j} w_{i,j}$, and the number of nodes with whom $i$ is observed interacting is its degree, $k_i$ [22].

We define $x_{j|i}$ to be the probability that an interaction involving $i$ will also involve node $j$. Therefore the probability that at least one of these interactions is with $j$ is $1 - (1 - x_{j|i})^s_i$. The main assumption of the model is that the values of $x_{j|i}$ over all $i, j$ pairs are distributed according to a probability distribution, $\rho(x)$.[1] Thus, if a node interacts $s$ times, the marginal probability that an edge exists between that node and any other given node in the network is

$$\Psi(s) = 1 - \int \rho(x)(1 - x)^s dx.$$ (1)

Our goal is to find a form of $\rho$ that accurately reproduces network structure observed in real social systems. Motivated by our exploration of empirical interaction patterns from a variety of species (Fig. S1), we propose that $\rho$ has a power-law form:

$$\rho(x) = \frac{\phi x^\phi}{1 - x^\phi} x^{-(1+\phi)}$$ for $\epsilon < x < 1,$ (2)

where $\phi (> 0)$ controls the variability in the values of $x$, and $\epsilon$ simply truncates the distribution to avoid divergence. Combining (1) and (2) we find

$$\Psi(s, \phi, \epsilon) = 1 - \frac{\phi s^{\phi}(1 - \epsilon)^{s+1}}{(1 - \epsilon^\phi)(s + 1)} 2F_1(s + 1, 1 + \phi, s + 2, 1 - \epsilon)$$ (3)

where the notation $2F_1$ refers to the Gauss hypergeometric function [23]. It follows from $\sum_j x_{j|i} = 1$ that

$$N = 1 + \frac{(1 - \phi)(1 - x^\phi)}{\phi x^\phi(1 - \epsilon^{1-\phi})},$$ (4)

which can be solved numerically to find $\epsilon$ for given values of $N$ and $\phi$. The expectation of the degree is $\Psi(s, \phi, N) = (N - 1)\Psi(s, \phi, \epsilon)$.

Fig. 1 illustrates how the value of $\phi$ can produce different types of social behavior. As $\phi$ is the main determinant of social behaviour in our model, we use the term social fluidity to refer to this quantity. Low social fluidity ($\phi < 1$) produces what we might describe as “allegiant” behavior: interactions with the same partner are frequently repeated at the expense of interactions with unfamiliar individuals. As $\phi$ increases, the model produces more “gregarious” behavior: interactions are repeated less frequently and the number of partners grows faster. While names like “social strategy” and “loyalty” have been applied to similar concepts [24, 25], fluidity, as a property of matter, is a useful metaphor for communicating the main idea behind this model.

**Estimating social fluidity in empirical networks:** To understand the results of the model in the context of real systems we estimate $\phi$ in 57 networks from 20 studies of human

---

1. $x_{j|i}$ are subject to network interdependencies. Specifically, $AX = X^T A$ and $X_1 = 0$, where $X$ is a matrix whose $i, j$ entry is $-1$ if $i = j$ and $x_{j|i}$ otherwise, $A$ is any diagonal matrix with positive entries, and 0 and 1 are column vectors of length $N$ containing only 0 and 1, respectively. Thus, $\rho(x)$ is the distribution of marginal $x_{j|i}$ values of the joint distribution $P(X).$
Social fluidity, a measure of how readily individuals change their social connections, is useful for comparing social behavior across different systems. It is not their absolute values, but rather the relationships between strength and degree that matter. We estimate social fluidity, squared squared error between the observed degrees and the expected ones given by the model. By applying this method to empirical networks, we can predict disease transmission, as the system is closed, and the total network size is equal to the number of individuals observed in at least one interaction. To estimate social fluidity, we find the squared error between the observed degrees and the expectation given by the model. Being estimated from the relationship between strength and degree, and not their absolute values, social fluidity is a good candidate for comparing social behavior across different systems as it is independent of the distributions of $s_i$ or $k_i$, and of the timescale of interactions.

Fig. 2 shows the estimated values of $\phi$ for all networks in our study. We organize the measurements of social fluidity by interaction type. Aggressive interactions have the highest fluidity (which implies that most interactions are rarely repeated between the same individuals), while grooming and other forms of social bonding have the lowest (which implies frequent repeated interactions between the same individuals). Social fluidity also appears to be related to species: ant systems cluster around $\phi = 1$, monkeys around $\phi = 0.5$, humans take a range of values that depend on the social environment. Sociality type does not appear to affect $\phi$; sheep, bison, and cattle have different social fluidity compared to kangaroos and bats, though they are all categorized as fission-fusion species [3].

There is no significant correlation between the mean number of interactions per individual ($\bar{s}$) and social fluidity ($\rho^2 = 0.02$, $p = 0.26$), which implies that sampling bias does not affect the estimation of social fluidity. Similarly, network size does not correlate with $\phi$ ($\rho^2 = 0.02$, $p = 0.33$). Larger values of $\phi$ correspond to higher mean degrees ($\rho^2 = 0.27$, $p < 0.001$) and lower variability in the distribution of edge weights (measured as the index of dispersion of $w_{ij}$; $\rho^2 = 0.26$, $p < 0.001$). Weight variability and mean degree are uncorrelated in these data ($\rho^2 = 0.01$, $p = 0.59$) implying that $\phi$ combines these two entirely distinct features of social behavior.

Finally, the modularity of the network (computed by the Louvain method on the unweighted network [47]) is negatively correlated with $\phi$ ($\rho^2 = 0.57$, $p < 0.001$). This is expected as individuals tend to be loyal to those within the same module while maintaining weaker connections with the remaining network - in all but one network the mean weight of edges within modules is higher than the mean weight of edges between modules (supplementary document).

**Characterizing disease spread with social fluidity**

Our objective is to characterize how social behavior influences the susceptibility of the group to infectious disease in a range of human and animal social systems. Intuitively, we expect an infected individual in a group with low social fluidity to expose fewer susceptible group members to the pathogen than they would in a group with highly fluid social dynamics. We explore this idea by introducing an analytical transmission model that incorporates social fluidity. Using this model, we mathematically characterize the impact of social fluidity on density dependence, and apply the model to empirical networks to predict disease spread.
Disease transmission model: We consider the transmission of an infectious disease on the social behavior model introduced in the previous section. An infectious node $i$ interacting with a susceptible node $j$ will transmit the infection with probability $\beta$. The node will recover from infection with rate $\gamma$, assuming an exponential distribution of the length of the infectious period. The probability that the infection is transmitted from $i$ to any given $j$ is

$$T_{i \rightarrow j}(\beta, \gamma, s, \tau, x_{ji}) = 1 - \exp(-s_i x_{ji} \beta / \gamma \tau), \quad (5)$$

assuming that the interactions $s_i$ of $i$ are distributed randomly across an observation period of duration $\tau$.

By integrating Eq. (5) over all possible values $x_{ji}$ and infection period durations and multiplying by the number of susceptible individuals $(N - 1)$ we obtain the expected number of infections caused by individual $i$,

$$r(s_i) = \frac{1 - \phi}{\phi(\epsilon^\phi - \epsilon)} \left[ 1 - e^\phi + e^\phi 2 F_1(-\phi, 1, 1 - \phi; -\beta s_i / \gamma \tau) - 2 F_1(-\phi, 1, 1 - \phi; -\epsilon \beta s_i / \gamma \tau) \right]. \quad (6)$$

The basic reproductive number (usually denoted $R_0$) is defined as the mean number of secondary infections caused by a typical infectious individual in an otherwise susceptible population [48]. We will use the notation $R_0^\phi$ to signify the social fluidity reproductive number, that is the analogue of $R_0$ derived from our social behavior model.

We assess the relation of the reproductive number with the population density by focusing on a special case where every node has the same strength, i.e. $s_i = s$ for all $i$, so that $R_0^\phi = r(s)$. Furthermore, we choose $\beta = \gamma \tau R_0^\infty / s$ where $R_0^\infty$ is $R_0^\phi$ as $\phi \rightarrow \infty$, i.e. a constant that represents what the basic reproductive number would be if every new interaction occurred between a pair of individuals who have not previously interacted with each other.

Fig. 3 shows the effect of social fluidity on the density dependence of the disease. At small population sizes, $R_0^\phi$ increases with $N$ and converges as $N$ goes to $\infty$ (Fig. 3A). The rate of this convergence increases with $\phi$, and the limit it converges to is higher, meaning that $\phi$ determines the extent to which density affects the spread of disease. As $N \rightarrow \infty$, we find that $R_0^\phi \rightarrow R_0^\infty$ for $\phi > 1$. When $\phi < 1$, $R_0^\phi \rightarrow [(1 - \phi) / \phi] [2 F_1(-\phi, 1, 1 - \phi; -R_0^\infty) - 1]$. At these values of $\phi$ the disease is constrained by individuals choosing to repeat interactions despite having the choice of infinitely many potential interaction partners (Fig 3B).

Estimating infection spread in empirical networks with heterogeneous connectivity: To apply this analogue of a reproductive number to an animal-disease system, we need to account for heterogeneous levels of social connectivity in the given population and thus the tendency for infected individuals to be those with a greater number of social partners [49]. For the basic reproductive number, this is often done using the mean excess degree, i.e. the degree of an individual selected with probability proportional to their degree [50]. Following a similar reasoning, we define $R_0^{\text{Est}}$, which incorporates the effect of social fluidity, as the expected number of infections $(r(s_i))$ caused by an individual that has been selected with probability proportional to their degree $(\kappa_i)$:

$$R_0^{\text{Est}}(\{s_i\}, \{\kappa_i\}, \tau, \beta, \gamma) = \frac{\sum_i \kappa_i r(s_i)}{\sum_i \kappa_i}. \quad (7)$$

Given the degree and strength of each individual in a network, the duration over which those interactions occurred, and the transmission and recovery rates of the disease, we are able to estimate $\phi$, compute Eq.(6) for each individual, and finally use Eq.(7) to derive a statistic that provides a measure of the risk of the host population to disease outbreak.

Numerical validation using empirical networks: We simulated the spread of disease through the interactions that occurred in the empirical data (materials and methods). We compute $R_0^{\text{Sim}}(g)$, defined as the ratio of the number of individuals infected at the $(g + 1)$-th generation to the number infected at the $g$-th generation over $10^3$ simulated outbreaks, for $g = 0, 1, 2$ ($g = 0$ refers to the initial seed of the outbreak).

Table 1 shows the Pearson correlation coefficient between $R_0^{\text{Sim}}(g)$ and its corresponding value $R_0^{\text{Est}}$ obtained Eq.(7). For comparison, the correlation is shown for other...
indicators and network statistics. The results correspond to one set of simulation conditions, and are robust across a wide range of parameter combinations (see supplementary tables). Note that a different value of $\beta$ was chosen for each network to control for the varying interaction rates between networks while keeping the upper bound ($R_0^\infty$) constant (materials and methods). Thus, the mean strength does not have a significant effect on $R_0^{\text{Sim}}(g)$.

Table 1: The Pearson correlation coefficient between quantities calculated on the network and the simulated disease outcomes (with $R_0^\infty = 3$). Results that are significant with $p < 0.01$ are labelled with *.

<table>
<thead>
<tr>
<th></th>
<th>Corr. with $R_0^{\text{Sim}}(g = 1)$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$R_0^{\text{Est}}$</td>
<td>0.91*</td>
</tr>
<tr>
<td>Social fluidity</td>
<td>0.73*</td>
</tr>
<tr>
<td>Excess degree</td>
<td>0.64*</td>
</tr>
<tr>
<td>Mean degree</td>
<td>0.53*</td>
</tr>
<tr>
<td>Network size</td>
<td>0.47*</td>
</tr>
<tr>
<td>Mean strength</td>
<td>$-0.07$</td>
</tr>
<tr>
<td>Mean clustering</td>
<td>$-0.15$</td>
</tr>
<tr>
<td>Mean edge weight</td>
<td>$-0.45*$</td>
</tr>
<tr>
<td>Edge weight variability</td>
<td>$-0.48*$</td>
</tr>
<tr>
<td>Modularity</td>
<td>$-0.60*$</td>
</tr>
</tbody>
</table>

These correlations support a known result regarding repeat contacts in network models of disease spread: that indicators of disease risk that are derived solely from the degree distribution are unreliable and the role of edge weights should not be neglected [51, 52]. After transmission has occurred from one individual to another, repeating the same interaction serves no advantage for disease (most directly-transmitted microparasites are not dose-dependent). Since a large edge weight implies a high frequency of repeated interactions, networks with a higher mean weight tend to have lower basic reproductive numbers. Furthermore, variability in the distribution of weights concentrates a yet larger proportion of interactions onto a small number of edges, further increasing the number of repeat interactions and reducing the reproductive number.

Correlation between modularity and $R_0^{\text{Sim}}(g)$ is partly due to the strong correlation between modular networks and those with high social fluidity. Consistent with other evidence [53], this suggests that transmission events occur mostly within the module of the seed node, with weaker social ties facilitating transmission to other modules. The effect of clustering (a measure of the number of connected triples in network [54]) correlates with smaller $R_0^{\text{Sim}}(2)$, consistent with other theoretical work [51, 55].

Finally, we find the model estimate of the social fluidity reproductive number $R_0^{\text{Est}}$ to be, on average, within 10% of the simulated value, $R_0^{\text{Sim}}(g)$ at $g = 1$. At $g = 2$ the amount of error is larger (to up to 29% for some parameter choices). Prediction accuracy at this generation is negatively correlated with the mean clustering coefficient. This is not surprising as $R_0^{\text{Est}}$ does not account for the accelerated depletion of susceptible neighbours that is known to occur in clustered networks [51, 55]. No other properties of the network affect the accuracy of $R_0^{\text{Est}}$ consistently across all parameter combinations (see supplementary tables).

### Discussion

We proposed a measure of fluidity in social behavior which quantifies how much mixing exists within the social relationships of a population. While social networks can be measured with a variety of metrics including size, connectivity, contact heterogeneity and frequency, our methodology reduces all such factors to a single quantity allowing comparisons across a range of human and animal social systems. Social fluidity correlates with both the density of social ties (mean degree) and the variability in the weight of those ties, though these quantities do not correlate with each other. Social fluidity is thus able to combine these two aspects seamlessly in one quantity.

By measuring social fluidity across a range of human and animal systems we are able to rank social behaviors. We identify aggressive interactions as the most socially fluid; this indicates a possible learning effect whereby each aggressive encounter is followed by a period during which individuals avoid further aggression with each other [56].

At the opposite end of the scale, we find interactions that strengthen bonds (and thus require repeated interactions) such as grooming in monkeys [57] and food-sharing in bats [33]. The fact that food-sharing ants are far more fluid than bats, despite performing the same kind of interaction, reflects their eusocial nature and the absence of any need to consistently reinforce bonds with their kin [58].

Most studies that aim to describe and quantify social structure are met with a number of challenges, including ours. First, the degree of an individual, for example, is known to scale with the length of the observation period [59]. By focusing not on the absolute value of degree, but instead on how degree scales with the number of observations, our analysis controls for this bias. Second, observed interactions have been assumed to persist over time [60]. In our model, only the distribution of edge weights remains constant through time, an assumption consistent with growing evidence [25, 61]. Third, duration of contacts is known to be important for disease spread [52]. We did not include explicitly the duration of each contact in our model, since this information was only available in a fraction of the datasets [62]. There is therefore potential to improve the applicability of this model as more high resolution data becomes openly available.

Our estimate of reproductive number derived from social fluidity provides a better predictor for the epidemic risk of a host population, going beyond predictors based on density or degree only. To illustrate this point, the social network of individuals at a conference ($R_0^{\text{Est}} = 1.60$; conference_0, supplementary document) is predicted to be at higher risk compared to the social network at a school ($R_0^{\text{Est}} = 1.39$; highschool_0), despite having a smaller size and lower connectivity ($N = 93$ vs. $N = 312$, and $k = 5.63$ vs. $k = 6.78$, respectively). The discrepancy in the risk prediction comes from the lower frequency of repeated contacts between individuals in the conference, compared to...
the school. Interactions between infectious individuals and those they have previously infected are redundant in terms of transmission. This dynamic is nicely captured by the social fluidity, with $\phi = 0.66$ for the conference and $\phi = 0.40$ for the high school.

Unlike previous work that explores the disease consequences of population mixing [63, 64], our analysis allows us to investigate this relation across a range of social systems. We see, for example, how the relationship between mixing and disease risk scales with population density. For social systems that have high values of social fluidity, $R_0^\phi$ is highly sensitive to changes in $N$, whereas this sensitivity is not present at low values of $\phi$. This corroborates past work on the scaling of transmission being associated to heterogeneity in contact [65,66]. Going beyond previous work, our model captures in a coherent theoretical framework both density-dependence and frequency-dependence, and social fluidity is the measure to tune from one to the other in a continuous way. Since many empirical studies support a transmission function that is somewhere between these two modeling paradigms [7,67–69], the modeling approaches applied in this paper can be carried forward to inform transmission relationships in future disease studies.

**Materials & Methods**

**A. Python libraries** Mean clustering coefficients were computed using the `networkx` Python library. To evaluate the hypergeometric function in (3) we used the `hyp2f1` function from the `scipy.special` Python library. Numerical solutions to Eq.(4) using the `fsolve` function from the `scipy.optimize` Python library. All scripts, data, and documentation used in this study are available through https://github.com/EwanColman/Social-Fluidity.

**B. Data handling** Only freely available downloadable sources of data have been used for this study. Details of the experimentation and data collection can be found through their respective publications. Here we note some additional processes we have applied for our study.

Each human contact dataset lists the identities of the people in contact, as well as the 20-second interval of detection [26–29,32]. Any sequence of consecutive time intervals for which contact is detected between two individuals is considered to be an interaction. To exclude contacts detected while participants momentarily walked past one another, only contacts detected in at least two consecutive intervals are considered interactions. Data were then separated into 24 hour subsets.

Bee trophallaxis provided experimental data for 5 unrelated colonies under continuous observation. We use the first hour of recorded data for each colony [46]. The ant trophallaxis study provided 6 networks: 3 unrelated colonies continuously observed under 2 different experimental conditions [30]. Ant antennation study provided 6 networks: 3 colonies, each observed in 2 sessions separated by a two week period. The bat study collected individual data at different times and under different experimental conditions [33]. For bats that were studied on more than one occasion we use only the first day they were observed.

Some data sets provided data for group membership collected through intermittent, rather than continuous, observa-

tion [34–38]. We construct networks from these data by recording an interaction when two individuals were seen to be in the same group during one round of observation. The shark data was divided into 6 datasets, each one constructed from 10 consecutive observations, and spread out through the full time period over which the data was collected.

For the grooming data [39,40], if one animal was grooming another during one round of observations then this would be recorded as a directed interaction. Similarly for aggressive interactions [41–45,56]. When an animal was determined to be the winner of a dominance encounter then this would be recorded as a directed interaction between the winner and the loser. We consider interaction in either direction to be a contact in the network.

We considered including two rodent datasets in which interaction is defined as being observed within the same territorial space [67,69]. We did not find this suitable for our analysis since the network we obtain, and the consequent results are sensitive to setting of arbitrary threshold values regarding what should, or should not, be considered sufficient contact for an interaction.

For data that did not contain the time of each interaction, contact time series were generated synthetically. For those datasets, the interactions between each pair were given synthetic timestamps in three different ways, Poisson: the time of each interaction is chosen uniformly at random from $\{0,1,\ldots,10^4\}$ seconds, Circadian: chosen uniformly at random from $\{0,1,\ldots,3333,6666,\ldots,10^4\}$, and Bursty: interaction times occur with power-law distributed inter-event times adjusted to give an expected total duration of $10^4$ seconds.

**C. Disease simulation** Simulations of disease spread were executed using the contacts provided by the datasets. The bat network was omitted from this part since these data were collected over a series of independent experiments carried out at different times and under different experimental treatments.

In one run of the simulation, one seed node is randomly chosen from the network and, at a randomly selected point in time during the duration of the data, transitions to the infectious state. The duration for which they remain infectious is a random variable drawn from an exponential distribution with mean $1/\gamma$. During this time any contact they have with other individuals who have not previously been infected will cause an infection with probability $\beta$.

The simulation runs until all individuals who were infected at the second generation of the disease, i.e. those infected by those infected by the seed, have recovered. The datasets are ‘looped’ to ensure that the timeframe of the data collection does not influence the outcome. In other words, immediately after the latest interaction, the interactions are repeated exactly as they were originally. This continues to happen until the termination criteria is met.

We set the parameters to normalise for the variation in contacts rates between networks. To achieve this we consider a hypothetical counterpart to each network in which the strength of every node is the same, but each interaction occurs between a pair of individuals who have not previously interacted. This is equivalent to $\phi \to \infty$. Under these conditions $x_{ij} = 1/(N-1)$ for all pairs $i,j$. It follows that Eq. (5) becomes $T_{i\to j} \approx s_i \beta / \gamma (N-1)$, then $r(s) \approx s_i \beta / \gamma \tau$, and, since $k_i = s_i$ for all nodes $i$, Eq. (7) gives

$$R_0^\infty = R_0^{\text{Ext}}(\{s_i\}, \{s_i\}, \tau, \beta, \gamma) = \frac{\beta \sum_i s_i^2}{\gamma \tau \sum_i s_i}.$$  

(8)
The value of $R_0^\infty$ can be chosen arbitrarily. Then, by setting $\gamma = 1/\tau$ and $\beta = R_0^\infty \sum_i s_i / \sum_i s_i^2$ we guarantee that Eq. (8) holds for every network. To test that our results hold over a range of disease scenarios we repeat our analysis with $R_0^\infty = 2$, 3, and 4.

Acknowledgments This work was supported by NSF grant number 141296. We are grateful for insightful feedback from Pratha Sah. We also thank all the researchers who have made their behavioral data openly accessible, making this study possible.

References


