Published in final edited form as:
J Theor Biol. 2012 January 21; 293: 87–100. doi:10.1016/j.jtbi.2011.10.010.

Invasion threshold in structured populations with recurrent mobility patterns

Duygu Balcan[a,b] and **Alessandro Vespignani**[c,d]

[a]Complex Networks and Systems Lagrange Laboratory, Institute for Scientific Interchange (ISI), Torino 10133, Italy

[b]Center for Complex Networks and Systems Research (CNetS), School of Informatics and Computing, Indiana University, Bloomington, IN 47408, USA

[c]College of Computer and Information Sciences and Department of Health Sciences, Northeastern University, Boston, MA 02115 USA

[d]Institute for Scientific Interchange (ISI), Torino 10133, Italy

Abstract

In this paper we develop a framework to analyze the behavior of contagion and spreading processes in complex subpopulation networks where individuals have memory of their subpopulation of origin. We introduce a metapopulation model in which subpopulations are connected through heterogeneous fluxes of individuals. The mobility process among communities takes into account the memory of residence of individuals and is incorporated with the classical susceptible-infectious-recovered epidemic model within each subpopulation. In order to gain analytical insight into the behavior of the system we use degree-block variables describing the heterogeneity of the subpopulation network and a time-scale separation technique for the dynamics of individuals. By considering the stochastic nature of the epidemic process we obtain the explicit expression of the global epidemic invasion threshold, below which the disease dies out before reaching a macroscopic fraction of the subpopulations. This threshold is not present in continuous deterministic diffusion models and explicitly depends on the disease parameters, the mobility rates, and the properties of the coupling matrices describing the mobility across subpopulations. The results presented here take a step further in offering insight into the fundamental mechanisms controlling the spreading of infectious diseases and other contagion processes across spatially-structured communities.

Keywords

Mathematical epidemiology; metapopulation model; reaction-diffusion process; contagion process; infectious disease

*Corresponding author. Tel: +39 0116603555. Fax: +39 0116600049. Address: Institute for Scientific Interchange (ISI), Viale Settimio Severo 65, Torino 10133, Italy., duygu.balcan@isi.it (Duygu Balcan), a.vespignani@neu.edu (Alessandro Vespignani).

1. Introduction

In recent years, reaction-diffusion processes have been used as a successful modeling framework to approach a wide array of systems that, along with the usual chemical and physical phenomena [1, 2], includes epidemic spreading [3–10], human mobility [6–9], and information and social contagion processes [11–16]. This paradigm is extremely useful in the case of populations characterized by a highly fragmented environment in which the population is structured and localized in relatively isolated discrete *patches* or subpopulations connected by mobility of individuals. In this case, the spatial structure of populations is known to play a key role in the system's evolution and the reaction-diffusion dynamics is integrated in a metapopulation modeling scheme in which different subpopulations are coupled together by the mobility or migration patterns of individuals [17–20]. Classic metapopulation dynamics focuses on the processes of local extinction, recolonization and regional persistence [21, 22] as the outcome of the interplay between migration processes and population dynamics, and has been successfully applied to understand the epidemic dynamics of spatially-structured populations with well-defined social units (e.g., families, villages, towns, cities, regions) connected through individuals' mobility [3–9, 23–28]. The metapopulation dynamics of infectious diseases has generated a wealth of models and results that consider both mechanistic approaches that take the movement of individuals explicitly into account [9, 29–34] and effective coupling approaches wherein the diffusion process is expressed as a force of infection coupling different subpopulations [6, 8, 35–39]. Recently, the metapopulation approach has been implemented in data-driven computational models for the large-scale analysis of the geographical spreading of infectious diseases [28, 40–45].

In large-scale systems, the metapopulation approach amounts to a particle-network description in which each subpopulation populated by a certain number of individuals is connected to a set of other subpopulation by mobility flows. The particle-network framework has stimulated the broadening of reaction-diffusion models in order to deal with complex network substrates and complex mobility schemes, which has in turn allowed for the uncovering of new and interesting dynamical behaviors as well as providing a rationale for the understanding of the emerging critical points underpinning some interesting characteristics of techno-social systems [46–49]. In particular, it has been shown that along the local epidemic threshold, which depends only on the disease parameters and is responsible for the epidemic outbreak within each subpopulation, structured populations may exhibit a global invasion threshold [47, 48, 50–52] that determines whether the metapopulation system is globally invaded by the contagion process. This novel threshold depends on the mobility rates and patterns of individuals and cannot be uncovered by continuous deterministic models as it is related to the stochastic effects of the reaction-diffusion process that describe the contagion process.

Metapopulation epidemic models, especially at the mechanistic level, are based on the spatial structure of the environment and the detailed knowledge of transportation infrastructures and movement patterns. However, the recent accumulation of large amounts of human mobility data [53–59] from the scale of single individuals to that of entire populations presents us with new challenges related to the high level of predictability and recurrence [60–62] found in the mobility and diffusion patterns in real data. For instance, commuting mobility – denoted by recurrent bidirectional flows among locations – dominates the human mobility network at the scale of census areas defined by major urban areas by one order of magnitude [59]. Highly predictable or recurrent mobility patterns do not find an easy representation in the particle-network framework as the framework is based on reaction-diffusion processes that in most cases exploit Markovian diffusion properties [47, 48, 63]. The description of mobility processes with memory and their importance in

epidemic processes have been put forward in detail in the seminal paper of Ref. [7]. The effect of recurrent and predictable mobility patterns of individuals on the onset of the global invasion behavior of contagion processes is just recently being studied both analytically and numerically [64–67].

Here we develop a framework based on a time-scale separation technique to analyze the behavior of contagion and spreading processes on a network of locations where individuals have memory of their location of origin. We focus on the prototypical example of the spreading of biological agents in populations characterized by bidirectional commuting patterns. We assume that individuals of a subpopulation will visit any one of the connected subpopulations with a per capita diffusion rate σ. As we aim at modeling commuting processes in which individuals have a memory of their location of origin, displaced individuals return to their original subpopulation with a per capita return rate τ. The mobility parameters σ and τ influence the probability that individuals carrying infection or information will export the contagion process to nearby subpopulations. If the diffusion rate approaches zero then the probability of contagion of neighboring subpopulations goes to zero as there are no occasions for the carriers of the process to visit them. On the other hand if the return rate is very high then the visiting time of individuals in neighboring populations is so short that they do not have time to spread the contagion in the visited subpopulations. This implies the presence of a transition [47, 48, 50–52] between a regime in which the contagion process may invade a macroscopic fraction of the network and a regime in which it is limited to a few subpopulations. The presented results extend and generalize the analysis of [47, 48, 64, 65] and we include in the analytical treatment the heterogeneity of the subpopulation network and find an explicit expression for the threshold separating a regime in which the spreading phenomenon affects a macroscopic fraction of the system and a regime in which only a few locations are affected. The invasion threshold depends on the mobility parameters, providing guidance on how to control disease spreading by constraining mobility processes. The results are confirmed by mechanistic Monte Carlo simulations for the infection dynamics in synthetic metapopulation systems in which each single individual is tracked in time to account for the discreteness of the processes involved. Heterogeneous connectivity patterns among subpopulations are modeled and different values of the parameters involved are considered to validate the theoretical results. The theoretical approach presented in this paper extends and generalizes the results presented in Ref. [64] opening the path to the inclusion of more complicated mobility or interaction schemes and at the same time provides a general framework that may be used not just as an interpretative framework. Understanding the effect of mobility and interaction patterns on the global spreading of contagion processes can then be used to enhance or suppress spreading by adjusting the basic parameters of the system in the appropriate ways.

The paper is organized as follows. Section 2 introduces the basic formalism for recurrent mobility patterns and the time-scale separation approximation that defines mixing subpopulations. Section 3 generalizes the formalism to the case of complex subpopulation networks by using a mean-field degree-block variables description equivalent to a mean-field description that includes the network heterogeneity. Section 4 incorporates the disease spreading into the mobility processes. Stochastic effects and discrete descriptions of the processes are considered with a tree-like approximation for the analysis of the invasion dynamics at the level of subpopulations. The effects of diffusion properties on the invasion dynamics are analyzed and related to the existence of an invasion epidemic threshold for the metapopulation system. In Section 5 we report extensive mechanistic Monte Carlo simulations which confirm the analytical findings of the previous sections.

2. Mobility processes with memory and commuting networks

In order to describe the mobility process induced by the commuting pattern of people among subpopulations let us consider a metapopulation system with V distinct subpopulations, each of which has a population size N_i ($i = 1, \ldots, V$). The subpopulations form a network in which each subpopulation i is connected to a set of other subpopulations $v(i)$. The edge connecting two subpopulations i and j indicates the presence of a flux of commuters. We assume that individuals in subpopulation i will visit any one of the connected subpopulations with a per capita diffusion rate σ_i. As we aim at modeling commuting processes in which individuals have memory of their residence, displaced individuals return to their original subpopulation with a per capita rate τ_i. The simple case of two connected subpopulations is shown in Fig. 1. In the case of each subpopulation i connected with k_i other subpopulations we assume that each commuting individual can visit a specific subpopulation j with rate σ_{ij} with the obvious condition that $\Sigma_{j\in v(i)}\,\sigma_{ij} = \sigma_i$. The rate σ_{ij} accounts for different attractiveness of each subpopulation and may depend on variables such as the population N_j, the actual distance, and other demographic or economic factors. We will provide two basic cases in which the rate σ_{ij} depends on the number of individuals at origin and destination subpopulations in analogy with the classic gravity law [68, 69] used in transportation sciences. For the sake of mathematical simplicity, however, we neglect the impact of actual distance between the origin and destination subpopulations, which is a key ingredient of gravity laws used to describe mobility fluxes among spatially structured populations.

At any moment in time, each member of subpopulation i is either in the subpopulation of residence or outside and visiting one of the neighboring subpopulations j. By using the approach developed in [7, 9], we may group the members of i according to the location in which they are actually present at a given time t, $N_{ii}(t)$ and $N_{ij}(t)$ with $j \in v(i)$ (see Fig. 1). The rate equations for the population sizes of different subgroups can be readily written by explicitly taking into account the mobility rates along the edges of the subpopulation network. This system of rate equations has a characteristic relaxation time that can be obtained by solving the differential equations, as reported in A. In particular, it is possible to show under the general assumption of $\sigma_i \ll \tau_i$ that the relaxation characteristic time is τ_i^{-1} and that the mixing subpopulations reads as

$$N_{ii}=\frac{N_i}{1+\sigma_i/\tau_i} \quad \text{and} \quad N_{ij}=\frac{N_i}{1+\sigma_i/\tau_i}\sigma_{ij}/\tau_i. \tag{1}$$

This implies that in the regime $\sigma_i \ll \tau_i$, $\{N_{ij}(t)\}$ represent a small perturbation to the overall subpopulation size N_i.

From the above considerations it appears that most of the system complexity is encoded in the subpopulation connectivity matrices that determine the σ_{ij} and the corresponding populations at the equilibrium (see A). A wide range of networks that represent activity and interaction patterns of individuals [70, 71], transportation fluxes, and population movements on local and global scales [53–59] have been found to exhibit complex features encoded in large-scale heterogeneity, self-organization, and other properties typical of complex systems. In particular, air-travel patterns between urban areas [54, 55] and commuting patterns in the scale of intra- and inter-urban areas [53, 56–59] have been shown to exhibit connectivity patterns varying over orders of magnitude. In order to provide empirical evidence of the connectivity patterns observed in real commuting networks we analyze the commuting fluxes in several countries. In Fig. 2 we display the cumulative distributions of the number of connections per administrative unit and the flux of commuters on each

connection in the United States and France. The networks exhibit important variability in the number of connections per geographical area, spanning three to four orders of magnitude. Similarly, the number of commuters on each connection is highly heterogeneous, distributed in a wide range of four to six orders of magnitude. These properties, often mathematically encoded in a heavy-tailed probability distribution, have been shown to have important consequences for dynamical processes on contact networks [72–77] as well as on subpopulation networks [46–48], altering the threshold behavior and the associated dynamical phase transition. It has been shown [47, 48] in particular that topological fluctuations in subpopulation networks favor global spreading of infectious diseases by lowering the global epidemic threshold and suppressing the phase transition at the infinite size limit of the network.

3. Mobility processes with memory in heterogeneous networks

In order to gain analytic insight into the case of subpopulation networks with highly heterogeneous connectivity patterns we rely on the assumption of statistical equivalence of subpopulations with similar degree. This is a mean-field approximation that considers all the subpopulations with same degree as statistically equivalent, thus allowing the introduction of degree-block variables that depend only upon the subpopulation degree. While this is an obvious approximation to the system description, it has been successfully applied to many dynamical processes on complex networks [46–48].

Imagine a random subpopulation network with given degree distribution $P(k)$ and denote the number of subpopulations with k connections by V_k. By using the statistical equivalence of subpopulations with the same degree k, we can express the average population in each node with degree k as a degree-block variable,

$$\overline{N}_k \equiv \frac{1}{V_k} \sum_{1|k_i=k} N_i, \tag{2}$$

where the sum runs over all the subpopulations with degree k, and attribute it to all such nodes. The above degree-block variable defines a mean-field approximation within each degree class, while relaxing the overall homogeneity assumption valid in homogeneous systems.

In the following we will consider that the degree-block variable $\bar{N}_k$ has a prescribed functional form identifying the stationary distribution of the population within the system, which does not change over time. In other words we assume that any mobility time scale is much shorter than the demography time scale of the population. For the sake of analytical calculation we will assume the functional form

$$\overline{N}_k = \overline{N} \frac{k}{\langle k \rangle}, \tag{3}$$

where $\bar{N} = \Sigma_k \bar{N}_k P(k)$ is the average number of individuals per node in the subpopulation network. This expression is obtained as the stationary population distribution in the case of simple random diffusive processes where the diffusion rate of individuals along each link leaving a node of degree k has the form $1/k$ [48]. Moreover, the empirical data from various sources suggest similar population scaling as a function of the connectivity to other subpopulations [45, 54, 78]. While the population $\bar{N}_k$ of each node is assumed to be fixed,

each individual in a subpopulation of degree k is assumed to commute to a neighboring subpopulation of degree k' at rate $\sigma_{kk'}$ and to return to the origin of her trip at rate τ_k. We thus also assume that the only dependence for the mobility rates is through the degrees of the nodes.

Following the same approach in the previous section we partition the individuals of a subpopulation with degree k according to classes identified by the degree of the subpopulation in which they are present at a given time t. Namely we define $\bar{N}_{kk}(t)$ as the average number of individuals of any subpopulation k present in the subpopulation k and $\bar{N}_{kk'}(t)$ as the average number of individuals of any subpopulation k present in the subpopulation k'. It is worth noting that $\bar{N}_{kk'}(t)$ is also a degree-block variable evaluated over all the connections $k \to k'$. The rate equations defining the commuting dynamics among subpopulations can be formalized by using the variables $\bar{N}_{kk}(t)$ and $\bar{N}_{kk'}(t)$. Considering Eq. (3) for $\bar{N}_k$, the set of differential equations leads to the equilibrium solutions

$$\overline{N}_{kk} = \frac{\overline{N}k}{\langle k \rangle (1+\rho_k)}, \tag{4}$$

$$\overline{N}_{kk'} = \frac{\overline{N}k\sigma_{kk'}}{\langle k \rangle (1+\rho_k)\tau_k}, \tag{5}$$

as detailed in B. In the above expressions ρ_k is the ratio of per capita total diffusion rate to return rate of residents of a subpopulation with degree k,

$$\rho_k \equiv \frac{\sigma_k}{\tau_k}. \tag{6}$$

The expressions (4–5) are valid for the stationary states and are good approximations to the system description under the general assumptions that $\sigma_k \ll \tau_k$ and that the system can equilibrate in a time interval larger than τ_k^{-1}. Analogously to the calculation originally presented by Keeling and Rohani [9], these expressions allow us to consider that the subpopulation k has an effective number of individuals $\bar{N}_{kk'}$ in contact with the individuals of the neighboring subpopulation k' in a quasi-stationary state that is reached whenever the time scale of a second dynamical process that we study is larger than τ_k^{-1}. This is extremely important in the case of disease dynamics, for which when the time scale of the disease is large enough compared to τ_k^{-1} one can generalize the quasi-stationary state expressions for all population compartments and obtain the effective force of infection or the number of individuals exposed to infection in the neighboring subpopulations. In the next section we will lever on this approximation to obtain explicit analytical conditions for the occurrence of epidemic outbreaks involving a finite fraction of subpopulations in the thermodynamic limit, the so-called ‘invasion threshold’ [47, 48, 50–52].

4. Epidemic spreading and the invasion threshold

Here we want to consider that an infectious disease has been introduced in one or a tiny number of subpopulations. For the sake of analytical simplicity we assume the usual susceptible-infectious-recovered (SIR) model [77] for the disease. The SIR compartmental

model classifies at any time t each individual by one of the disease compartments: susceptible (S); infectious (I); recovered (R). Susceptible individuals acquire infection in the case of contact with an infectious individual at a per capita rate β. An infectious individual permanently recovers at a rate μ, and from then on is immune to the disease and is not contagious. In a completely susceptible subpopulation, the spreading of the disease is determined by the basic reproduction ratio R_0, which is defined as the average number of secondary (infectious) cases generated by the introduction of a single typical infectious individual [77]. Assuming homogeneous mixing of the population, the basic reproduction ratio is given by $R_0 = \beta/\mu$. When $R_0 > 1$ the infectious disease spreads inside the subpopulation and causes an appreciable fraction of the population to acquire the infection during its entire progression.

In the case of metapopulation models, however, the spreading of the disease is determined not only by the local parameter R_0, but by the diffusion process and the topology of the underlying network as well. This can be readily understood in the case of no mobility across subpopulations since the disease cannot invade other subpopulations and therefore remains constrained to a single subpopulation. In the case of a very small diffusion of individuals the disease generally dies out, depleting the susceptible pool of individuals before it can generate an epidemic in neighboring subpopulations. On the contrary, in the case of sustained mobility across subpopulations many infectious individuals can export the disease to other subpopulations thus generating a global outbreak in the metapopulation system. This implies that there must be a second reproductive number at the subpopulation level R_* that depends on the mobility parameters and defines a threshold for the epidemic invasion of a finite fraction of subpopulations [47, 48, 50–52]. Here we want to investigate the dependence of R_* on the commuting dynamics of individuals, thus taking into account the non-Markovian nature of the mobility process.

Imagine to introduce an infectious disease to a single subpopulation of degree k and size $\bar{N}_k$. Given that the local threshold condition is satisfied ($R_0 > 1$), the epidemic will spread globally in the metapopulation systems if each infected subpopulation is able to trigger the start of an epidemic in at least one other subpopulation. This amounts to the derivation of a subpopulation reproductive number that expresses the average number of infected subpopulations generated by a typical infected subpopulation in a fully susceptible metapopulation system. We may describe the early stage of the disease spreading at the level of the metapopulation system as a branching process [47, 48, 50, 79, 80]. The process starts with a set of initially infected subpopulations $\{ D_k^0\}$, each of which during the course of the epidemic transmits the disease to a set of its non-infected neighboring subpopulations defining the set $\{ D_k^1\}$ at the next generation. Let us denote the number of diseased subpopulations with degree k at the nth generation by D_k^n and derive the relation between subsequent generations D_k^{n-1} and D_k^n. If we consider the case that we are just above the local epidemic threshold, $R_0 - 1 \ll 1$, and that there are no correlations between the degrees of connected nodes, then it is possible to show that

$$D_k^n = (R_0 - 1)\frac{kP(k)}{\langle k \rangle}\sum_{k'} D_{k'}^{n-1}(k' - 1)\lambda_{k'k}, \tag{7}$$

where $\lambda_{k'k}$ is the number of infectious individuals which can be generated and sent by a diseased subpopulation (source) with degree k' to a non-diseased subpopulation (target) with degree k and we made use of the fact that the probability of triggering an outbreak in the

non-diseased subpopulation is $1 - R_0^{-\lambda_{k'k}}$ [81]. (see C). In order to write an explicit form of the above expression we need to explicitly define $\lambda_{k'k}$. This quantity can be expressed in terms of the number of individuals living in one of the subpopulations and visiting the other by

$$\lambda_{k'k}=(\overline{N}_{k'k}+\overline{N}_{kk'})\alpha, \tag{8}$$

where α is the fraction of individuals that is affected by the disease by the end of the SIR epidemic and can be approximated by $\alpha \simeq 2(R_0 - 1)/R_0^2$ for $R_0 \simeq 1$ [77]. The first term on the right-hand side of $\lambda_{k'k}$ accounts for the total visits of infectious people from the source to the target subpopulation, while the second term accounts for the visits of individuals from the target to the source subpopulation during which they acquire infection and carry the disease back to their origin. If the time scale of the disease in each individual μ^{-1} is much larger than the time scale of the visits to neighbors τ^{-1} (i.e., $\mu^{-1} \gg \tau^{-1}$), we can then use the stationary solutions for $\bar{N}_{k'k}$ and $\bar{N}_{kk'}$ in Eq. (5). Furthermore we are assuming that there are no restrictions on the mobility of infectious people. This is generally not true even though for many diseases like influenza [82, 83] it is well known that a large fraction of clinical cases keep their regular habits as well as infectious asymptomatic cases. For other diseases this may not be a realistic assumption, but the following calculations can be performed by considering exposed or latent individuals as the disease carriers as long as the characteristic time of these disease states is larger than the mobility time scale.

In the following we will assume that the return rate τ_k is independent of the subpopulation degree, i.e., $\tau_k = \tau$. This implies that the time scale of commuting is basically the same across all the subpopulations, a reasonable assumption as the return rate is mostly defined by universal features such as work hours per day. For the commuting rates we instead assume the form

$$\sigma_{kk'}=\alpha\frac{\overline{N}_{k'}}{k\sum_{k'}\overline{N}_{k'}P(k'|k)}, \tag{9}$$

where the denominator corresponds to the total average population of subpopulations in the neighborhood. This term just represents a normalization factor so that the total commuting rate per capita is the same $\sigma_k = \sigma$ across all the subpopulations. The above relation states that the total number of individuals commuting between subpopulations k and k' is proportional to the product of the two populations. This is a simple case of the gravity law [68, 69] used in transportation studies. We use this form for the sake of mathematical simplicity, but more complicated functions of populations can be considered both analytically and numerically. We should note however that more realistic models consider also the actual distances between subpopulations [59, 84] in order to describe mobility fluxes among spatially structured populations.

Using all the assumptions above, the dynamical behavior of the iterative system (7) is determined by the branching ratio

$$R_* = \frac{2\overline{N}(R_0 - 1)^2 \rho}{R_0^2(1+\rho)} f(\langle k\rangle, \langle k^2\rangle, \langle k^3\rangle, \langle k^4\rangle), \tag{10}$$

where $\rho \equiv \sigma/\tau$ is the ratio of commuting to return rate, as shown in C. In the above expression f is a function only of the moments of the degree distribution of the subpopulation network,

$$f \equiv \frac{1}{\langle k\rangle\langle k^2\rangle}\left[\langle k^3\rangle - \langle k^2\rangle + (\langle k^4\rangle - \langle k^3\rangle)^{1/2}(\langle k^2\rangle - \langle k\rangle)^{1/2}\right]. \tag{11}$$

The infectious disease will spread globally in the metapopulation system only if $R_* > 1$. This is equivalent to defining a subpopulation reproductive number [47, 48, 50–52] that in structured metapopulation systems is the average number of infected subpopulations generated by a typical infected subpopulation in a fully susceptible metapopulation system. Thus, by setting $R_* = 1$, we can define an epidemic threshold relation for the mobility ratio ρ,

$$\rho_c = \frac{1}{2\overline{N}(1 - R_0^{-1})^2 f(\langle k\rangle, \langle k^2\rangle, \langle k^3\rangle, \langle k^4\rangle) - 1}, \tag{12}$$

below which the infection remains confined to a small number of subpopulations. In an infinite metapopulation system the threshold is defined rigorously and the fraction of infected subpopulations is zero below the threshold and finite only if the mobility parameters set the ratio ρ above the threshold value.

The threshold value is defined for the ratio between the rates characterizing the mobility process. We can therefore obtain two different threshold conditions on the mobility dynamics if we fix one of the two parameters σ and τ, and let the other parameter free. The threshold relation for σ is $\sigma_c = \rho_c\, \tau$,

$$\sigma_c = \frac{\tau}{2\overline{N}(1 - R_0^{-1})^2 f(\langle k\rangle, \langle k^2\rangle, \langle k^3\rangle, \langle k^4\rangle) - 1}. \tag{13}$$

This intuitively states that the rates of diffusion to nearby subpopulations has to be large enough ($\sigma > \sigma_c$) to guarantee the spreading of the disease. Interestingly, however, we can also define the threshold relation for τ^{-1}, $\tau_c^{-1} = \rho_c/\sigma$,

$$\tau_c^{-1} = \frac{\sigma^{-1}}{2\overline{N}(1 - R_0^{-1})^2 f(\langle k\rangle, \langle k^2\rangle, \langle k^3\rangle, \langle k^4\rangle) - 1}, \tag{14}$$

that tells us that the global spreading of the disease can be achieved by reducing the return rates of individuals – in other words by extending the visit times of individuals in nearby

subpopulations ($\tau^{-1} > \tau_c^{-1}$). However, this last condition breaks down when τ^{-1} gets much larger and becomes comparable to the disease time scale, thus breaking the time-scale separation assumption used here. In Fig. 3, we assume μ^{-1} very large, and draw the phase diagram separating the global invasion regime from the extinction regime in the σ-τ^{-1} plane, with the global epidemic threshold curve defined by the relation $R_* = 1$.

Another very interesting feature of the above threshold value ρ_c is the explicit effect of the network topology encoded in the moments of degree distribution $\langle k \rangle$, $\langle k^2 \rangle$, etc. As has been already observed in the Markovian diffusion case, the heterogeneity of the network favors the global spread of the epidemic by lowering the threshold value. Indeed, for heavy-tailed degree distributions $P(k) \sim k^{-\gamma}$ with $\gamma > 1$ and $k_{\min} \leq k \leq k_{\max}$, the nth moment scales as $k_{\max}^{n+1-\gamma}$ if $n \geq \gamma - 1$ and $k_{\max} \gg k_{\min}$. This means that for $n \geq \gamma - 1$, the nth moment of the degree distribution tends to diverge at the infinite size limit of the network, as in this limit $k_{\max} \to \infty$, virtually reducing the threshold to zero. In particular, if the exponent $\gamma < 5$, then $f(\langle k \rangle, \langle k^2 \rangle, \langle k^3 \rangle, \langle k^4 \rangle)$ tends to diverge in the limit of infinite network size, which in turn pushes the threshold value ρ_c to zero, as reported in D. Even at finite size, however, the threshold value is generally smaller the higher the network heterogeneity is (see D for various examples). In Fig. 3 we compare the epidemic threshold curves in the space of the mobility parameters for a Poisson network and a heavy-tailed network with degree distribution $P(k) \sim k^{-2.1}$.

While the previous R_* and threshold conditions have been obtained in the case of a very simple $\sigma_{kk'}$ resulting in $\sigma_k = \sigma$ independent of the subpopulation index, very similar expressions can be obtained in more complicated mobility schemes. For instance, we can consider the case

$$\sigma_{kk'} = \sigma \frac{\overline{N}_{k'}}{\overline{N}_k + k \sum_{k'} \overline{N}_{k'} P(k'|k)} \tag{15}$$

that assumes that the per capita mobility rate is rescaled by the number of individuals in the subpopulation [9], thus leading to σ_k that decreases as $\overline{N}_k$ increases. This behavior accounts for the effect introduced by large subpopulation sizes; the overall per capita commuting rate outside of the subpopulation generally decreases in large populations as individuals tend to commute internally. In this case we can repeat the above calculations, finally obtaining

$$R_* = \frac{2\overline{N}(R_0 - 1)^2 \rho}{R_0^2 [\langle k \rangle^2 + \langle k^2 \rangle \langle k \rangle (1 - \rho)]} \left[\langle k^3 \rangle - \langle k^2 \rangle + (\langle k^4 \rangle - \langle k^3 \rangle)^{1/2} (\langle k^2 \rangle - \langle k \rangle)^{1/2} \right]. \tag{16}$$

The above expression recovers global epidemic threshold conditions very close to those obtained previously, supporting the general robustness of the presented results. In Fig. 4 we draw the phase diagram separating the global invasion regime from the extinction regime in the σ-τ^{-1} plane, with the global epidemic threshold curve defined by the relation $R_* = 1$ for the case of Poisson and heavy-tailed networks. Also, in this case the global epidemic threshold is a function of the mobility parameters and the network topology.

It is worth stressing again that the previous expressions are derived under the assumption of equivalence of degree-block variables, valid only at the limit in which a small fraction of the subpopulations in the system is affected and in which $R_0 - 1 \ll 1$ and $\tau^{-1} \ll \mu^{-1}$. However,

it is extremely relevant that metapopulation systems intrinsically have two epidemic thresholds in the case of mobility processes with memory. The emergence of a global epidemic is first constrained by the intrinsic epidemic threshold within each subpopulation, $R_0 > 1$. If the epidemic process satisfies this condition, then each time an infectious individual seeds an epidemic within a subpopulation there is a finite probability that a macroscopic fraction of the population will be affected by the outbreak. While this condition guarantees the intra-population spreading of the epidemic, the inter-population spreading is controlled by the coupling among subpopulations as quantified by the rates of the commuting dynamics. The global invasion threshold condition $R_* > 1$ provides an estimate of the diffusion and return rates of individuals above which the epidemic is able to affect a macroscopic fraction of subpopulations defining the metapopulation system.

5. Stochastic Simulations

In the following we provide extensive numerical simulations to support the theoretical picture described above. We present in detail the mechanistic numerical simulations where each single individual is tracked in time, during both the infection dynamics and the diffusion processes, and the synthetic subpopulation networks. We report results from Monte Carlo simulations in a variety of different cases and compare them with the analytical findings.

5.1. Computational model and synthetic subpopulation networks

We construct the network of V subpopulations from a pool of $N\bar{V}$ people. Population size N_i assigned to each subpopulation i is chosen at random from a multinomial distribution with probability proportional to k_i, which ensures that the metapopulation system obeys Eq. (3). In order to compare with theoretical calculations, the subpopulation network structure has been generated by wiring the subpopulations according to two different random graph topologies:

- Erdős-Rényi graphs [85] have been synthetized by assigning a link between each pair of nodes with probability $\langle k \rangle/(V-1)$, where $\langle k \rangle$ is a prescribed average node degree;
- Networks with power-law degree distribution, $P(k) \sim k^{-\gamma}$ with $k_{\min} \leq k \leq k_{\max}$, have been generated by an uncorrelated configuration model [86, 87]. All the scale-free networks have been generated by setting $\gamma = 2.1$ and $k_{\min} = 2$.

For the sake of comparison, the average degree of Erdős-Rényi graphs has been set to that of scale-free networks.

As detailed in E, we adopt mechanistic numerical simulations in which each individual is tracked in time during both the commuting and the infection dynamics. The rate σ_{ij} at which a resident of subpopulation i commutes to a neighboring subpopulation $j \in \nu(i)$ assumes Eq. (15):

$$\sigma_{ij} = \sigma \frac{N_j}{N_i + \sum_{\ell \in \nu(i)} N_\ell}. \tag{17}$$

Each resident in subpopulation i leaves her origin and visits subpopulation j with probability $\sigma_{ij}\,\Delta t$, where Δt is the time interval considered. A commuter in subpopulation j returns back to her permanent subpopulation i with probability $\tau \Delta t$. Inside each subpopulation we consider an SIR epidemic model in which each individual is classified by one of the discrete disease states at any point in time. The rate at which a susceptible person in subpopulation i

acquires the infection, the so-called 'force of infection' λ_i, is determined by interactions with infectious individuals. The force of infection $\lambda_i(t)$ acting on each susceptible individual in subpopulation i at time t has been assumed to follow the mass-action principle

$$\lambda_i(t)=\beta\frac{I_i^*(t)}{N_i^*(t)}, \tag{18}$$

where β is the transmission rate of infection and $I_i^*(t)/N_i^*(t)$ is the prevalence of infectious individuals in the subpopulation. Given the force of infection $\lambda_i(t)$ in subpopulation i, each person in the susceptible compartment (S) contracts the infection with probability $\lambda_i(t)\Delta t$ and enters the infectious compartment (I). Each infectious individual permanently recovers with probability $\mu\Delta t$ and enters the recovered compartment (R).

5.2. Numerical results

The epidemic invasion threshold at the metapopulation level is determined by reproduction ratio R_0, commuting ratio ρ, and the architecture of the commuting networks. In the following we present extensive numerical simulations in order to demonstrate these dependencies and verify the analytical result of Eq. (16).

Simulations have been initialized with $I(0) = 10$ infectious individuals, seeded randomly in a single subpopulation of degree $k_{\min}$, while the rest of the population is assumed to be susceptible to infection. Since we aim at determining the global invasion threshold, we have let the metapopulation system progress until the infection dies out. In the results we present here all the realizations resulting in at least one diseased subpopulation have contributed to the statistical analysis unless stated otherwise. For each set of parameters we have generated at least 2, 000 system realizations. Since the subpopulation networks and dynamical processes on them are subject to fluctuations, we have sampled at least 10–20 network realizations and 100–200 dynamical realizations on each of them.

Let us first turn our attention to the effects of ρ and R_0 on the global attack rate as shown in Fig. 5. We can clearly see the dependence of the critical value of ρ on R_0, consistent with the analytical result of Eq. (16). The lower the basic reproduction ratio the higher the commuting ratio needed for the infection to spread successfully at the metapopulation level. Another observation is that if ρ is far from its critical value (below or above), then the global attack rate is not altered by further changes in ρ. Hence we can conclude that the attack rate is determined only by local parameters in subpopulation systems coupled with commuting.

In order to uncover the dependence of invasion threshold on the individual mobility rates numerically, we have simulated epidemics on homogeneous and heterogeneous subpopulation networks. In Fig. 6 we display the variation in epidemic size as a function of σ and τ^{-1}. Consistent with the phase diagrams of Fig. 4, the smaller the commuting rate the longer the duration of visits needed to enable the infection to invade an appreciable fraction of subpopulations. We can also easily notice the one order of magnitude difference in the critical values of σ and τ^{-1} as the degree distribution of subpopulation networks changes.

We have stressed the importance of network topology on the spreading of infectious agents at the metapopulation level. In order to verify the analytical result of Eq. (16), in Figs. 7–9 we compare homogeneous and heterogeneous subpopulation networks. In Fig. 7 we display the distribution of epidemic sizes as a function of ρ. We are performing stochastic simulations and, as it is well known also for the single population, the threshold condition $R_* > 1$ on the subpopulation reproductive number is not a sufficient condition for an

outbreak. If we start with a single infected subpopulation, stochastic fluctuations may lead to the extinction of the epidemic with a probability that is larger the more R_* approaches the threshold from above. By considering the same expression obtained in the single population case and by using a coarse grained perspective on the subpopulation spreading the probability of observing a global outbreak is $1 - R_*^{-1}$ [81]. In Fig. 8 we show that the probability of having a macroscopic outbreak is null below the values close to the global threshold observed in the epidemic size plot. Above the threshold the probability increase slowly as we start in all simulations with a single diseased subpopulation. The probability of outbreak approaches a step function at the threshold by increasing the initial number of diseased subpopulations. For the completeness of comparisons we also display the average epidemic size as a function of ρ in Fig. 9. One readily notices an increase by more than one order of magnitude in the critical value of ρ as we move from a heavy-tailed to a Poisson degree distribution. The heterogeneity in the degree distribution favors the spreading process at the metapopulation level, leading to $\rho_c \to 0$ at the infinite size limit.

6. Conclusions

In this paper we have set a mathematical framework to investigate the conditions of global epidemic invasion in the case of subpopulations coupled with recurrent mobility patterns. On one hand we have extended the mathematical framework of degree-block variables [47, 48] to gain insight into the impact of a non-Markovian mobility process on epidemic extinction/persistence, while on the other hand we have extended the time-scale separation approach of Ref. [9] to complex settings in which subpopulations exhibit heterogeneous demographic and mobility properties.

The basic reproduction ratio is responsible for an epidemic to take off within a single subpopulation. Our analytical approach however derives the expression of a global invasion threshold parameter that depends on the disease parameters, the mobility parameters and the network architecture. Once the network structure and the disease dynamics are fixed, there exists a critical value of the mobility ratio below which the disease dies out before reaching an appreciable fraction of subpopulations. This indicates the importance of both of the commuting and return rates. In other words, the time individuals spend in temporary subpopulations is as important as the rate at which they visit a neighboring subpopulation. The network architecture also has substantial impact on the spreading dynamics. We have shown that the heterogeneity in the commuting matrices favors global spreading by lowering the critical values of the threshold parameters. In principle, the epidemic threshold is suppressed at the infinite size limit of the subpopulation network. Given that real commuting networks are highly heterogeneous, this has important consequences on the dynamical processes and explains the ineffectiveness of mobility restrictions in the containment of emergent infectious diseases [88].

Although our theoretical contribution provides a first theoretical framework for the analysis of global threshold phenomena in systems with recurrent mobility patterns, there are still theoretical and practical issues to address. We have assumed that the mobility rates only depend at most upon the population sizes of the origin and destination subpopulations. In realistic cases, however, the rates may also depend on other social and economic factors as well as the geographical distance between communities. We have also considered a very simplistic case of the dynamics within each subpopulation, ignoring possible intra-population heterogeneities such as different levels of susceptibility and infectiousness [89], or heterogeneous mixing among individuals [90, 91]. All these factors could introduce the need for more sophisticated theoretical approaches. While most of the studies in defining epidemic thresholds have been focused on single populations, it is clear that more attention has to be devoted to the study of spreading in structured populations. In this case the

understanding of the invasion threshold is crucial to the analysis of large-scale spreading across communities and subpopulations. Accurate data on human mobility patterns have been increasingly available in recent years, enabling us to gain insight into the statistical laws governing their properties. Given that the spreading of human infectious diseases is only possible through movements of individuals across different geographical regions, it is crucial to explore the consequences of different types of movement patterns for the dynamical processes like epidemic spread. The theoretical framework presented here is thus very important for the development of realistic computational approaches in the study of structured metapopulation systems as well, which are good candidates for capturing fundamental aspects of interacting human communities.

A. Time-scale separation and stationary population

Mobility and infection dynamics detailed in previous sections have been assumed to operate at different time scales. In particular we have restricted our calculations to a regime in which the commuting is much faster than the infection dynamics. In the following we consider the temporal progression of subpopulation sizes and evaluate the relaxation time to an equilibrium configuration.

The rate equations governing the spatial distributions $N_{ii}(t)$ and $N_{ij}(t)$ of residents of subpopulation i can be readily written by explicitly taking into account the diffusion rates along the edges of the subpopulation network. For each subpopulation i the diffusion equations describing the system dynamics are

$$\partial_t N_{ii} = -\sigma_i N_{ii}(t) + \tau_i \sum_{\ell \in \nu(i)} N_{i\ell}(t), \tag{19}$$

$$\partial_t N_{ij} = \sigma_{ij} N_{ii}(t) - \tau_i N_{ij}(t). \tag{20}$$

The first equation evaluates the variation of $N_{ii}(t)$ as the net balance of individuals diffusing away or returning to the subpopulation i according to the rates σ_i and τ_i, respectively. The second equation considers the change in the population size $N_{ij}(t)$ as the net balance of individuals of subpopulation i leaving to and returning from subpopulation j. Using the relation for the total population size of subpopulation i,

$$N_i = N_{ii}(t) + \sum_{j \in \nu(i)} N_{ij}(t), \tag{21}$$

we can derive the closed expression

$$\partial_t N_{ii}(t) + (\tau_i + \sigma_i) N_{ii}(t) = N_i \tau_i. \tag{22}$$

The solution of this first order differential equation is

$$N_{ii}(t)=\frac{N_i}{1+\sigma_i/\tau_i}+C_{ii}\exp\left[-(\tau_i+\sigma_i)t\right], \tag{23}$$

where the constant C_{ii} is determined by the initial condition $N_{ii}(0)$, leading to

$$N_{ii}(t)=\frac{N_i}{1+\sigma_i/\tau_i}+\left(N_{ii}(0)-\frac{N_i}{1+\sigma_i/\tau_i}\right)\exp\left[-(\tau_i+\sigma_i)t\right]. \tag{24}$$

We can similarly solve the differential equation for the population size of the residents of subpopulation i present at subpopulation j, obtaining

$$\begin{aligned}N_{ij}(t)=&\frac{N_i}{1+\sigma_i/\tau_i}\sigma_{ij}\\ &/\tau_i-\frac{\sigma_{ij}}{\sigma_i}\left(N_{ij}(0)-\frac{N_i}{1+\sigma_i/\tau_i}\right)\exp\left[-(\tau_i+\sigma_i)t\right]\\ &+\left[N_{ii}(0)-\frac{N_i}{1+\sigma_i/\tau_i}\sigma_{ij}/\tau_i+\frac{\sigma_{ij}}{\sigma_i}\left(N_{ii}(0)-\frac{N_i}{1+\sigma_i/\tau_i}\right)\right]\exp(-\tau_i t).\end{aligned} \tag{25}$$

The relaxation times of $N_{ii}(t)$ and $N_{ij}(t)$ to their equilibrium configurations are thus $(\tau_i + \sigma_i)^{-1}$ and τ_i^{-1}, respectively. The former term is dominated by τ_i^{-1} if the relation $\tau_i \gg \sigma_j$ holds. In the case of commuting, $\sigma_i = \Sigma_{i\in\nu(i)}\,\sigma_{ij}$, which equals the daily total commuting rate per resident of subpopulation i. Such the rate is always smaller than 1 since only a fraction of the local population commutes, and it is typically much smaller than the typical return rate $\tau_i \simeq 3 - 10\ \text{day}^{-1}$ of such visits. Therefore the relaxation characteristic time can be safely approximated by τ_i^{-1}. Hence for $t \gg \tau_i^{-1}$ we can approximate the population sizes $N_{ii}(t)$ and $N_{ij}(t)$ with their equilibrium values, and recover the Eq. (1) reported in the main text. This approximation, originally introduced by Keeling and Rohani [9], allows us to consider each subpopulation i as having an effective number of individuals N_{ij} in contact with the individuals of the neighboring subpopulation j. In practice, this is equivalent to consider an infinite time-scale separation between the commuting time scale and the other time scales in the problem (e.g., infection dynamics).

In the case of influenza-like illnesses, the typical time-scale separation between τ_i and the compartmental transition rates is close to one order of magnitude or even larger. Eq. (1) can then be generalized in the time-scale separation regime to all traveling compartments X obtaining the general expressions

$$X_{ii}(t)=\frac{X_i(t)}{(1+\sigma_i/\tau_i)} \quad \text{and} \quad X_{ij}(t)=\frac{X_i(t)}{(1+\sigma_i/\tau_i)}\sigma_{ij}/\tau_i, \tag{26}$$

while $X_{ii}(t) = X_i(t)$ and $X_{ij}(t) = 0$ for all the other compartments which are restricted from traveling. While the approximation holds exactly only in the limit $\tau_i \to \infty$, it is good enough as long as τ_i is much larger than the typical transition rates of the disease dynamics.

B. Stationary population in heterogeneous networks

In our mean-field description the total number of individuals within a subpopulation of degree k is given by

$$\overline{N}_k=\overline{N}_{kk}(t)+k\sum_{k'}\overline{N}_{kk'}(t)P(k'|k), \tag{27}$$

where the sum in the second term accounts for the average total number of residents of subpopulation k visiting any of its neighbors. This term is proportional to the number of neighbors k and the average number of residents $\bar{N}_{kk'}$ traveling on each connection $k \rightarrow k'$. Thus the sum is performed over all the possible connections through the degrees weighted by the conditional probability $P(k'|k)$ that represents the conditional probability that any given edge departing from a node of degree k is pointing to a node of degree k' [92]. The rate equations defining the commuting dynamics among subpopulations can be defined using the variables $\bar{N}_{kk}(t)$ and $\bar{N}_{kk'}(t)$ as

$$\partial_t\overline{N}_{kk}=-\sigma_k\overline{N}_{kk}+\tau_k k\sum_{k'}\overline{N}_{kk'}P(k'|k), \tag{28}$$

$$\partial_t\overline{N}_{kk'}=\sigma_{kk'}\overline{N}_{kk}-\tau_k\overline{N}_{kk'}, \tag{29}$$

where $\sigma_k = k\Sigma_{k'}\sigma_{kk'}P(k'|k)$ expresses the per capita diffusion rate for individuals in a subpopulation of degree k. In this case we also have considered that the diffusion rate $\sigma_{kk'}$ is averaged over all possible k' according to the probability $P(k'|k)$ that a given edge departing from a node of degree k is pointing to a node of degree k'. The above set of equations leads to the equilibrium condition

$$\overline{N}_{kk'}=\frac{\sigma_{kk'}}{\tau_k}\overline{N}_{kk}, \tag{30}$$

and using the expression of Eq. (27) we obtain the following equilibrium expressions

$$\overline{N}_{kk}=\frac{\overline{N}_k}{1+\sigma_k/\tau_k}, \tag{31}$$

$$\overline{N}_{kk'}=\frac{\overline{N}_k}{1+\sigma_k/\tau_k}\sigma_{kk'}/\tau_k. \tag{32}$$

Considering in the above expressions the functional form of Eq. (3) for $\bar{N}_k$ leads to Eqs. (4–5).

C. Branching process and epidemic invasion threshold

The expression for the number of infected subpopulations with degree k at the nth generation reads as

$$D_k^n = \sum_{k'} D_{k'}^{n-1}(k'-1)\left[(1-R_0^{-\lambda_{k'k}}\right]P(k/k') \times \prod_{m=0}^{n-1}\left(1-\frac{D_k^m}{V_k}\right), \tag{33}$$

where each diseased subpopulation of degree k' at the $n-1$th generation may seed its $k'-1$ non-infected neighbors (all of its neighbors minus the one from which it got the infection). Here $\lambda_{k'k}$ is the number of infectious seeds that are introduced into a fully susceptible population, and the probability of observing an outbreak in this case is $1-R_0^{-\lambda_{k'k}}$ [81]. The probability that the neighboring subpopulation has not already been infected in the earlier generations is $\prod_{m=0}^{n-1}(1-D_k^m/V_k)$. In the early stage of the epidemic we can assume that $\prod_{m=0}^{n-1}(1-D_k^m/V_k) \simeq 1$. We will also consider the case that we are just above the local epidemic threshold, $R_0 - 1 \ll 1$, so that the outbreak probability can be approximated by

$$1-R_0^{-\lambda_{k'k}} \simeq \lambda_{k'k}(R_0-1). \tag{34}$$

In the following we will ignore the degree correlation between neighboring nodes. In this case the conditional probability $P(k'|k)$ does not depend on the originating node, i.e., $P(k'|k) = k'P(k')/\langle k\rangle$ [92]. This relation simply states that any edge has a probability of pointing to a node with degree k' that is proportional to the degree of the node to which it points. By using this form of the conditional probability we obtain Eq. (7).

In order to write an explicit form of the above expression we need to explicitly define $\lambda_{kk'}$. If the time scale of the disease in each individual μ^{-1} is much larger than the time scale of the visits to neighbors τ^{-1} (i.e., $\mu^{-1} \gg \tau^{-1}$), we can then use the stationary solutions for $\bar{N}_{kk'}$ and $\bar{N}_{k'k}$ in Eq. (5). If we also consider the approximate expression for α in the SIR case for $R_0 \simeq 1$, then we obtain

$$\lambda_{kk'} = \frac{2\overline{N}(R_0-1)}{R_0^2\langle k\rangle}\left[\frac{k\sigma_{kk'}}{(1+\rho_k)\tau_k} + \frac{k'\sigma_{k'k}}{(1+\rho_{k'})\tau_{k'}}\right]. \tag{35}$$

If we substitute this expression for $\lambda_{k'k}$ in Eq. (7) we obtain the explicit expression

$$D_k^n = \frac{2\overline{N}(R_0-1)^2}{R_0^2\langle k\rangle^2}kP(k)\sum_{k'}D_{k'}^{n-1}(k'-1)\times\left[\frac{k'\sigma_{k'k}}{(1+\rho_{k'})\tau_{k'}} + \frac{k\sigma_{kk'}}{(1+\rho_k)\tau_k}\right] \tag{36}$$

that depends only on the precise definition of the commuting rates $\sigma_{kk'}$ and τ_k. In the following we will assume that $\tau_k = \tau$ and $\sigma_k = \sigma$ in all the subpopulations. For the commuting rate $\sigma_{kk'}$ on each connection $k \to k'$ we instead assume the form of Eq. (9). In this case by

defining $\rho \equiv \sigma/\tau$ we can write the basic branching process equations at the metapopulation level as

$$D_k^n = \frac{2\overline{N}(R_0-1)^2\rho}{R_0^2\langle k^2\rangle\langle k\rangle(1+\rho)} kP(k) \times \sum_{k'} D_{k'}^{n-1}(k'-1)(k+k'). \tag{37}$$

In order to write a closed form of the above iterative process we introduce the variables $\Theta_0^n \equiv \sum_k D_k^n(k-1)$ and $\Theta_1^n \equiv \sum_k D_k^n(k-1)k$ whose next generation equations are defined as

$$\mathbf{\Theta^n} = G\mathbf{\Theta^{n-1}} \quad \text{with} \quad \mathbf{\Theta^n} = \begin{pmatrix} \Theta_0^n \\ \Theta_1^n \end{pmatrix}, \tag{38}$$

where the matrix G is

$$G = c\begin{pmatrix} \langle k^3\rangle - \langle k^2\rangle & \langle k^2\rangle - \langle k\rangle \\ \langle k^4\rangle - \langle k^3\rangle & \langle k^3\rangle - \langle k^2\rangle \end{pmatrix} \tag{39}$$

and c is a constant, defined by the following expression:

$$c = \frac{2\overline{N}(R_0-1)^2\rho}{R_0^2\langle k^2\rangle\langle k\rangle(1+\rho)}. \tag{40}$$

The dynamical behavior of the system is determined by the largest eigenvalue R_* of the matrix G, leading to Eq. (10).

D. Impact of heterogeneity on epidemic invasion

Remember that Eq. (11) encodes the dependence of global epidemic invasion threshold R_* on the topology of subpopulation networks. In the following we consider three different uncorrelated random network topologies.

Heavy-tailed degree distribution—We would here like to turn our attention to the scaling of $f(\langle k\rangle, \langle k^2\rangle, \langle k^3\rangle, \langle k^4\rangle)$ with system size for heavy-tailed degree degree distributions $P(k) \sim k^{-\gamma}$ with $\gamma > 1$ and $k_{\min} \le k \le k_{\max}$. For very large system sizes, and in the case that $1 < \gamma < 2$, f scales as

$$f \sim \frac{\langle k^3\rangle + \langle k^4\rangle^{1/2}\langle k^2\rangle^{1/2}}{\langle k\rangle\langle k^2\rangle} \sim k_{\max}^{\gamma-1}. \tag{41}$$

In the case that $2 < \gamma < 3$, the second moment $\langle k^2\rangle$ in the denominator and higher moments in the numerator dominate, leading to the scaling relation:

$$f \sim \frac{\langle k^3\rangle + \langle k^4\rangle^{1/2}\langle k^2\rangle^{1/2}}{\langle k^2\rangle} \sim k_{\max}. \tag{42}$$

In the range $3 < \gamma < 4$, only the third moment $\langle k^3\rangle$ in the numerator dominates, thus

$$f \sim \langle k^3\rangle \sim k_{\max}^{4-\gamma}. \tag{43}$$

In the range $4 < \gamma < 5$, only the fourth moment $\langle k^4\rangle$ in the numerator dominates, leading to

$$f \sim \langle k^4\rangle^{1/2} \sim k_{\max}^{(5-\gamma)/2}. \tag{44}$$

The above expressions state that for any heavy-tailed degree distribution with exponent $\gamma < 5$, $f(\langle k\rangle, \langle k^2\rangle, \langle k^3\rangle, \langle k^4\rangle)$ tends to diverge in the limit of infinite network size, which in turn pushes the threshold value ρ_c to zero. If $\gamma > 5$, then $f(\langle k\rangle, \langle k^2\rangle, \langle k^3\rangle, \langle k^4\rangle)$ has a finite value.

Poisson degree distribution—Now let us consider a Poisson degree distribution with average degree λ. By using the generating function, $M_k(\tau) = e^{\lambda(e^t-1)}$, we can calculate the nth moment as

$$\langle k^n\rangle = \frac{d^n M_k(t)}{dt^n}\Big|_{t=0}. \tag{45}$$

We only need to know the first four moments in order to calculate f, which are $\langle k\rangle = \lambda$,

$$\begin{aligned}\langle k^2\rangle &= \lambda+\lambda^2,\\ \langle k^3\rangle &= \lambda+3\lambda^2+\lambda^3,\\ \langle k^4\rangle &= \lambda+7\lambda^2+6\lambda^3+\lambda^4.\end{aligned} \tag{46}$$

Then f is

$$f(\langle k\rangle, \langle k^2\rangle, \langle k^3\rangle, \langle k^4\rangle) = \frac{\lambda+2+(\lambda+4)^{1/2}(\lambda+1)^{1/2}}{\lambda+1}. \tag{47}$$

The global invasion threshold parameter in this case is

$$R_* = \frac{2\overline{N}(1-R_0^{-1})^2[\lambda+2+(\lambda+4)^{1/2}(\lambda+1)^{1/2}]}{(1+\rho^{-1})(1+\lambda)}. \tag{48}$$

Finally the critical mobility ratio is given by

$$\rho_c = \frac{\lambda+1}{2\overline{N}(1-R_0^{-1})^2[\lambda+2+(\lambda+4)^{1/2}(\lambda+1)^{1/2}] - \lambda - 1}. \tag{49}$$

Equal degrees—If all the nodes have the same degree λ, then the nth moment of degree distribution is simply $\langle k^n \rangle = \lambda^n$. In this case, f is

$$f(\langle k\rangle, \langle k^2\rangle, \langle k^3\rangle, \langle k^4\rangle) = \frac{2(\lambda-1)}{\lambda}. \tag{50}$$

The global epidemic threshold parameter R_* is then

$$R_* = \frac{4\overline{N}(1-R_0^{-1})^2(\lambda-1)}{(1+\rho^{-1})\lambda}. \tag{51}$$

Notice what this expression tells us about the system: For any $\lambda \leq 1$, global epidemic spread is not possible (as in this regime $R_* \leq 0$). If $\lambda > 1$, then the critical value of mobility ratio ρ_c is

$$\rho_c = \frac{\lambda}{4\overline{N}(1-R_0^{-1})^2(\lambda-1) - \lambda}. \tag{52}$$

If we assume that ρ is fixed, then we can also calculate the critical value of λ above which the epidemic will spread globally:

$$\lambda_c = \frac{4\overline{N}(1-R_0^{-1})^2}{4\overline{N}(1-R_0^{-1})^2 - (1+\rho^{-1})}. \tag{53}$$

E. Infection and commuting dynamics

Since all the individuals with the same three indices (X, i, j) are identical in terms of the dynamical processes, we are going refer to the number of such individuals at time t by $X_{ij}(t)$. Then, by definition, the instantaneous compartment size $X_j^*(t)$ in subpopulation j can be expressed as

$$X_j^*(t) = X_{jj}(t) + \sum_{\ell \in \nu(j)} X_{\ell j}(t), \tag{54}$$

and the total number of individuals as $N_j^* = \sum_X X_j^*$. The number of individuals in each compartment X with a residence in i and present in j is subject to discrete and stochastic dynamical processes defined by disease and transport operators. The disease operator $\mathcal{D}_j$

represents the change due to the compartment transition induced by the infection dynamics, and the transport operator Ω_X represents the variation due to mobility.

The term $\mathcal{D}_j$ can be written as a combination of a set of transitions $\{\mathcal{D}_j(X, Y)\}$, where $\mathcal{D}_j(X, Y)$ represents the number of transitions from compartment X to Y and is simulated as an integer random number extracted from a multinomial distribution. Then the change due to infection dynamics reads as

$$\mathcal{D}_j(X)=\sum_Y[\mathcal{D}_j(Y,X)-\mathcal{D}_j(X,Y)]. \tag{55}$$

As a concrete example let us consider the temporal change in the infectious compartment. There is only one possible transition from the compartment, which is to the recovered compartment. The number of transitions is extracted from the binomial distribution

$$\Pr^{\text{Binom}}(I_{ij}(t), p_{I_{ij}\to R_{ij}}), \tag{56}$$

which is determined by the transition probability

$$p_{I_{ij}\to R_{ij}}=\mu\Delta t, \tag{57}$$

and the number of individuals in the compartment I_{ij} (τ) (its size). This transition causes a reduction in the size of the compartment. The increase in the compartment size is due to the transitions from the susceptible to infectious compartment. This is also a random number extracted from the binomial distribution

$$\Pr^{\text{Binom}}(S_{ij}(t), p_{S_{ij}\to I_{ij}}), \tag{58}$$

given by the chance of contagion

$$p_{S_{ij}\to I_{ij}}=\lambda_j(t)\Delta t, \tag{59}$$

and the number of attempts equal to the number of susceptibles S_{ij} (t). After extracting these numbers from the appropriate distributions, we can calculate the total change $\mathcal{D}_j(I)$ in the infectious compartment as

$$\mathcal{D}_j(I)=\mathcal{D}_j(S,I)-\mathcal{D}_j(I,R). \tag{60}$$

Transport operator Ω_X expresses the total change in compartment sizes due to the commuting of permanent residents of subpopulation i back and forth. The variation in X_{ij} can be decomposed into $\Omega_x^{\rightarrow}(i, j)$ and $\Omega_x^{\leftarrow}(j, i)$ as

$$\Omega_X = \Omega_X^{\rightarrow}(i, j) - \Omega_X^{\leftarrow}(j, i). \tag{61}$$

The first term $\Omega_X^{\rightarrow}(i, j)$ represents an increase that is caused by the departing residents of subpopulation i to visit subpopulation j. The $\Omega_X^{\rightarrow}(i, j)$ is a random number extracted from the multinomial distribution

$$\Pr^{\text{Multinom}}(X_{ii}(t), \{p_{X_{ii}\to X_{i\ell}} | \ell \in \nu(i)\}), \tag{62}$$

determined by the probability of commuting to subpopulation j

$$p_{X_{ii}\to X_{ij}} = \sigma_{ij}\Delta t, \tag{63}$$

and the number of such trails $X_{ii}(t)$. The second term $\Omega_X^{\leftarrow}(j, i)$ corresponds to a reduction in X_{ij} and is due to the return trips from subpopulation j to permanent subpopulation i. The $\Omega_X^{\leftarrow}(j, i)$ is also a random number extracted from the binomial distribution

$$\Pr^{\text{Binom}}(X_{ij}(t), p_{X_{ij}\to X_{ii}}), \tag{64}$$

given by the probability of returning home

$$p_{X_{ij}\to X_{ii}} = \tau\Delta t, \tag{65}$$

and the size of the compartment X_{ij} (τ). We have assumed that the infection does not alter people's behavior, i.e., all the compartments are identical in their mobility. Recognize that the stochastic state variables $\{S_{ij}(\tau), I_{ij}(\tau), R_{ij}(\tau)\}$ define a multivariate Markov chain [93–95] in which the present state of the system is determined only by the state of the system in the previous time step.

Acknowledgments

We would like to thank Vitaly Belik and Dirk Brockmann for sharing their results on a closely related work. We would also like to thank Chiara Poletto and Vittoria Colizza for interesting discussions during the preparation of this manuscript. This work has been partially funded by the NIH R21-DA024259 award and the DTRA-1-0910039 award to AV; The work has also been partly sponsored by the Army Research Laboratory and was competed under Cooperative Agreement Number W911NF-09-2-0053. The views and conclusions contained in this document are those of the authors and should not be interpreted as representing the official policies, either expressed or implied, of the Army Research Laboratory or the U.S. Government.

References

1. Marro, J.; Dickman, R. Nonequilibrium Phase Transitions in Lattice Models. Cambridge University Press; Cambridge: 1999.
2. van Kampen, NG. Stochastic Processes in Physics and Chemistry. North-Holland; Amsterdam: 1981.

3. Anderson RM, May RM. Spatial, temporal and genetic heterogeneity in host populations and the design of immunization programs. IMA J Math Appl Med Biol. 1984; 1:233–266. [PubMed: 6600104]
4. May RM, Anderson RM. Spatial heterogeneity and the design of immunization programs. Math Biosci. 1984; 72:83–111.
5. Bolker BM, Grenfell T. Chaos and biological complexity in measles dynamics. Proc R Soc London B. 1993; 251:75–81.
6. Bolker BM, Grenfell T. Space persistence and dynamics of measles epidemics. Philos Trans R Soc London B. 1995; 348:309–320. [PubMed: 8577828]
7. Sattenspiel L, Dietz K. A structured epidemic model incorporating geographic mobility among regions. Math Biosci. 1995; 128:71–91. [PubMed: 7606146]
8. Lloyd AL, May RM. Spatial heterogeneity in epidemic models. J Theor Biol. 1996; 179:1–11. [PubMed: 8733427]
9. Keeling MJ, Rohani P. Estimating spatial coupling in epidemiological systems: a mechanistic approach. Ecol Lett. 2002; 5:20–29.
10. Watts D, Muhamad R, Medina DC, Dodds PS. Multiscale resurgent epidemics in a hierarchical metapopulation model. Proc Natl Acad Sci USA. 2005; 102:11157–11162. [PubMed: 16055564]
11. Rapoport A. Spread of information through a population with socio-structural bias: I. assumption of transitivity. Bull Math Biol. 1953; 15:523–533.
12. Goffman W, Newill VA. Generalization of Epidemic Theory: An Application to the Transmission of Ideas. Nature. 1964; 204:225–228. [PubMed: 14212412]
13. Goffman W. Mathematical Approach to the Spread of Scientific Ideas – The History of Mast Cell Research. Nature. 1966; 212:449–452. [PubMed: 5339138]
14. Dietz K. Epidemics and Rumours: A Survey. J of Royal Stat Soc A. 1967; 130:505–528.
15. Tabah AN. Literature Dynamics: Studies on Growth, Diffusion, and Epidemics. Ann Rev Inform Sci Technol. 1999; 34:249–286.
16. Daley, DJ.; Gani, J. Epidemic Modeling: An Introduction. Cambridge University Press; Cambridge: 2000.
17. Hanski, I.; Gilpin, ME. Metapopulation Biology: Ecology, Genetics, and Evolution. Academic Press; San Diego: 1997.
18. Tilman, D.; Kareiva, P. Spatial Ecology. Princeton University Press; Princeton: 1997.
19. Bascompte, J.; Solé, RV. Modeling Spatiotemporal Dynamics in Ecology. Springer; New York: 1998.
20. Hanski, I.; Gaggiotti, OE. Ecology Genetics and Evolution of Metapopulations. Elsevier, Academic Press; Amsterdam, New York: 2004.
21. Levins R. Some demographic and genetic consequences of environmental heterogeneity for biological control. Bull Entomol Soc Am. 1969; 15:237–240.
22. Levins R. Extinction, Lecture Notes in Mathematics. 1970; 2:75–107.
23. Hethcote HW. An immunization model for a heterogeneous population. Theor Popul Biol. 1978; 14:338–349. [PubMed: 751264]
24. May RM, Anderson RM. Population biology of infectious diseases part II. Nature. 1979; 280:455–461. [PubMed: 460424]
25. Grenfell BT, Harwood J. (Meta)population dynamics of infectious diseases. Tree. 1997; 12:395–399. [PubMed: 21238122]
26. Grenfell BT, Bolker BM. Cities and villages: infection hierarchies in a measles metapopulation. Ecol Lett. 1998; 1:63–70.
27. Ferguson NM, Keeling MJ, Edmunds WJ, Gani R, Grenfell BT, Anderson RM. Planning for smallpox outbreaks. Nature. 2003; 425:681–685. [PubMed: 14562094]
28. Riley S. Large-scale transmission models of ingectious disease. Science. 2007; 316:1298–1301. [PubMed: 17540894]
29. Baroyan OV, Genchikov LA, Rvachev LA, Shashkov VA. An attempt at large-scale influenza epidemic modelling by means of a computer. Bull Int Epidemiol Assoc. 1969; 18:22–31.

30. Rvachev LA, Longini IM. A mathematical model for the global spread of influenza. Math Biosci. 1985; 75:3–22.
31. Grais RF, Hugh Ellis J, Glass GE. Assessing the impact of airline travel on the geographic spread of pandemic influenza. Eur J Epidemiol. 2003; 18:1065–1072. [PubMed: 14620941]
32. Longini IM. A mathematical model for predicting the geographic spread of new infectious agents. Math Biosci. 1988; 90:367–383.
33. Flahault A, Valleron A-J. A method for assessing the global spread of HIV-1 infection based on air-travel. Math Popul Stud. 1991; 3:1–11.
34. Ruan S, Wang W, Levin S. The effect of global travel on the spread of SARS. Math Biosci Eng. 2006; 3:205–218. [PubMed: 20361819]
35. Earn DJD, Rohani P, Grenfell BT. Persistence chaos and synchrony in ecology and epidemiology. Proc R Soc London B. 1998; 265:7–10.
36. Rohani P, Earn DJD, Grenfell BT. Opposite patterns of synchrony in sympatric disease metapopulations. Science. 1999; 286:968–971. [PubMed: 10542154]
37. Keeling MJ. Metapopulation moments: coupling, stochasticity and persistence. J Anim Ecol. 2000; 69:725–736.
38. Park AW, Gubbins S, Gilligan CA. Extinction times for closed epidemics: the effects of host spatial structure. Ecol Lett. 2002; 5:747–755.
39. Vázquez A. Epidemic outbreaks on structured populations. J Theor Biol. 2007; 245:125–129. [PubMed: 17097683]
40. Grais RF, Ellis JH, Kress A, Glass GE. Modeling the spread of annual influenza epidemics in the US: the potential role of air travel. Health Care Manage Sci. 2004; 7:127–134.
41. Cooper BS, Pitman RJ, Edmunds WJ, Gay NJ. Delaying the international spread of pandemic influenza. PLoS Med. 2006; 3:e12. [PubMed: 16435888]
42. Hollingsworth TD, Ferguson NM, Anderson RM. Will travel restrictions control the international spread of pandemic influenza? Nat Med. 2006; 12:497–499. [PubMed: 16675989]
43. Hufnagel L, Brockmann D, Geisel T. Forecast and control of epidemics in a globalized world. Proc Natl Acad Sci USA. 2004; 101:15124–15129. [PubMed: 15477600]
44. Colizza V, Barrat A, Barthélemy M, Valleron A-J, Vespignani A. Modeling the worldwide spread of pandemic influenza: Baseline case and containment interventions. PLoS Medicine. 2007; 4:e13. [PubMed: 17253899]
45. Colizza V, Barrat A, Barthélemy M, Vespignani A. The role of the airline transportation network in the prediction and predictability of global epidemics. Proc Natl Acad Sci USA. 2006; 103:2015–2020. [PubMed: 16461461]
46. Colizza V, Pastor-Satorras R, Vespignani A. Reaction-diffusion processes and metapopulation models in heterogeneous networks. Nat Phys. 2007; 3:276–282.
47. Colizza V, Vespignani A. Invasion Threshold in Heterogeneous Metapopulation Networks. Phys Rev Lett. 2007; 99:148701. [PubMed: 17930732]
48. Colizza V, Vespignani A. Epidemic modeling in metapopulation systems with heterogeneous coupling pattern: Theory and simulations. J Theor Biol. 2008; 251:450–467. [PubMed: 18222487]
49. Vespignani A. Predicting the behavior of tecno-social systems. Science. 2009; 325:425–428. [PubMed: 19628859]
50. Ball F, Mollison D, Scalia-Tomba G. Epidemics with two levels of mixing. Ann Appl Probab. 1997; 7:46–89.
51. Cross P, Lloyd-Smith JO, Johnson PLF, Wayne MG. Duelling timescales of host movement and disease recovery determine invasion of disease in structured populations. Ecol Lett. 2005; 8:587–595.
52. Cross P, Johnson PLF, Lloyd-Smith JO, Wayne MG. Utility of R_0 as a predictor of disease invasion in structured populations. J R Soc Interface. 2007; 4:315–324. [PubMed: 17251146]
53. Chowell G, Hyman JM, Eubank S, Castillo-Chavez C. Scaling laws for the movement of people between locations in a large city. Phys Rev E. 2003; 68:066102.
54. Barrat A, Barthélemy M, Pastor-Satorras R, Vespignani A. The architecture of complex weighted networks. Proc Natl Acad Sci USA. 2004; 101:3747–3752. [PubMed: 15007165]

55. Guimerá R, Mossa S, Turtschi A, Amaral LAN. The worldwide air transportation network: anomalous centrality, community structure, and cities' global roles. Proc Natl Acad Sci USA. 2005; 102:7794–7799. [PubMed: 15911778]
56. Brockmann D, Hufnagel L, Geisel T. The scaling laws of human travel. Nature. 2006; 439:462–465. [PubMed: 16437114]
57. Patuelli R, Reggiani R, Gorman SP, Nijkamp P, Bade F-J. Network analysis of commuting flows: A comparative static approach to German data. Networks Spatial Econ. 2007; 7:315–331.
58. González MC, Hidalgo CA, Barabási A-L. Understanding individual human mobility patterns. Nature. 2008; 453:779–782. [PubMed: 18528393]
59. Balcan D, Colizza V, Gonçalves B, Hu H, Ramasco JJ, Vespignani A. Multiscale mobility networks and the spatial spreading of infectious diseases. Proc Natl Acad Sci USA. 2009; 106:21484–21489. [PubMed: 20018697]
60. Wang P, González MC. Understanding spatial connectivity of individuals with non-uniform population density. Phil Trans R Soc A. 2009; 367:3321–3329. [PubMed: 19620127]
61. Song C, Qu Z, Blumm N, Barabási A-L. Limits of Predictability in Human Mobility. Science. 2010; 327:1018–1021. [PubMed: 20167789]
62. Song C, Koren T, Wang P, Barabási A-L. Modelling the scaling properties of human mobility. Nat Phys. 2010; 6:818–823.
63. Ni S, Weng W. Impact of travel patterns on epidemic dynamics in heterogeneous spatial metapopulation networks. Phys Rev E. 2009; 79:016111.
64. Balcan D, Vespignani A. Phase transitions in contagion processes mediated by recurrent mobility patterns. Nat Phys. 2011; 7:581–586. [PubMed: 21799702]
65. Belik V, Geisel T, Brockmann D. Natural human mobility patterns and spatial spread of infectious diseases. Phys Rev X. 2011; 1:011001.
66. Danon L, House T, Keeling MJ. The role of routine versus random movements on the spread of disease in Great Britain. Epidemics. 2009; 1:250–258. [PubMed: 21352771]
67. Keeling MJ, Danon L, Vernon MC, House TA. Individual identity and movement networks for disease metapopulations. Proc Natl Acad Sci USA. 2010; 107:8866–8870. [PubMed: 20421468]
68. Erlander, S.; Stewart, NF. The Gravity Model in Transportation Analysis. VSP; Utrecht: 1990.
69. de Dios Ortúzar, J.; Willumsen, LG. Modelling Transport. Wiley; Chichester: 2001.
70. Liljeros F, Edling CR, Amaral LAN, Stanley HE, Aberg Y. The web of human sexual contacts. Nature. 2001; 411:907–908. [PubMed: 11418846]
71. Schneeberger A, Mercer CH, Gregson SAJ, Ferguson NM, Nyamukapa CA, Anderson RM, Johnson AM, Garnett GP. Scale-free networks and sexually transmitted diseases. Sexually Transmitted Dis. 2004; 31:380–387.
72. Pastor-Satorras R, Vespignani A. Epidemic spreading in scale-free networks. Phys Rev Lett. 2001; 86:3200–3203. [PubMed: 11290142]
73. Pastor-Satorras R, Vespignani A. Epidemic dynamics and endemic states in complex networks. Phys Rev E. 2001; 63:066117.
74. Moreno Y, Pastor-Satorras R, Vespignani A. Epidemic outbreaks in complex heterogeneous networks. Eur Phys J B. 2002; 26:521–529.
75. Lloyd AL, May RM. How viruses spread among computers and people. Science. 2001; 292:1316–1317. [PubMed: 11360990]
76. Barthélemy M, Barrat A, Pastor-Satorras R, Vespignani A. Dynamical patterns of epidemic outbreaks in complex heterogeneous networks. J Theor Biol. 2005; 235:275–288. [PubMed: 15862595]
77. Keeling, MJ.; Rohani, P. Modeling Infectious Diseases in Humans and Animals. Princeton University Press; Princeton: 2008.
78. Colizza V, Barrat A, Barthélemy M, Vespignani A. The modeling of global epidemics: stochastic dynamics and predictability. Bull Math Biol. 2006; 68:1893–1921. [PubMed: 17086489]
79. Harris, TE. The theory of branching processes. Dover Publications; 1989.
80. Vázquez A. Polynomial growth in age-dependent branching processes with diverging reproductive number. Phys Rev Lett. 2006; 96:038702. [PubMed: 16486783]

81. Bailey, NT. The Mathematical Theory of Infectious Diseases. Macmillan; New York: 1975.
82. Longini IM, Halloran ME, Nizam A, Yang Y. Containing Pandemic Influenza with Antiviral Agents. Am J Epidemiol. 2004; 159:623–633. [PubMed: 15033640]
83. Carrat F, Vergu E, Ferguson NM, Lemaitre M, Cauchemez S, Leach S, Valleron A-J. Time Lines of Infection and Disease in Human Influenza: A Review of Volunteer Challenge Studies. Am J Epidemiol. 2008; 167:775–785. [PubMed: 18230677]
84. Viboud C, Bjørnstad ON, Smith DL, Simonsen L, Miller MA, Grenfell BT. Synchrony, Waves, and Spatial Hierarchies in the Spread of Influenza. Science. 2006; 312:447–451. [PubMed: 16574822]
85. Erdős P, Rényi A. On random graphs. Publ Math. 1959; 6:290–297.
86. Molloy M, Reed B. The Size of the Largest Component of a Random Graph on a Fixed Degree Sequence. Combinatorics, Probab Comput. 1998; 7:295–306.
87. Catanzaro M, Bogunã M, Pastor-Satorras R. Generation of uncorrelated random scale-free networks. Phys Rev E. 2005; 71:027103.
88. Bajardi P, Poletto C, Ramasco JJ, Tizzoni M, Colizza V, Vespignani A. Human Mobility Networks, Travel Restrictions, and the Global Spread of 2009 H1N1 Pandemic. PLoS ONE. 2011; 6:e16591. [PubMed: 21304943]
89. Miller JC. Spread of infectious disease through clustered populations. J R Soc Interface. 2009; 6:1121–1134. [PubMed: 19324673]
90. Meyers LA, Pourbohloul B, Newman MEJ, Skowronski DM, Brunham RC. Network theory and SARS: predicting outbreak diversity. J Theor Biol. 2005; 232:71–81. [PubMed: 15498594]
91. Lloyd-Smith JO, Schreiber SJ, Kopp PE, Getz WM. Super-spreading and the effect of individual variation on disease emergence. Nature. 2005; 438:355–359. [PubMed: 16292310]
92. Barrat, A.; Barthélemy, M.; Vespignani, A. Dynamical Processes on Complex Networks. Cambridge University Press; Cambridge: 2008.
93. Gani J, Jerwood D. Markov chain methods in chain binomial epidemic models. Biometrics. 1971; 27:591–603. [PubMed: 5116573]
94. Longini, IM. Chain binomial models. In: Armitage, P.; Colton, T., editors. The Encyclopedia of Biostatistics. Vol. 1. Wiley; NY: 1998. p. 593-597.
95. Halloran, ME.; Longini, IM.; Struchiner, CJ. Design and Analysis of Vaccine Studies, Statistics for Biology and Health. Springer; New York: 2010. Binomial and stochastic transmission models; p. 63-84.

Highlights

- We model the contagion spreading mediated by recurrent mobility patterns.
- We characterize analytically a phase transition between two regimes of spreading.
- We derive the threshold values in mobility rates that ensure the global spreading.

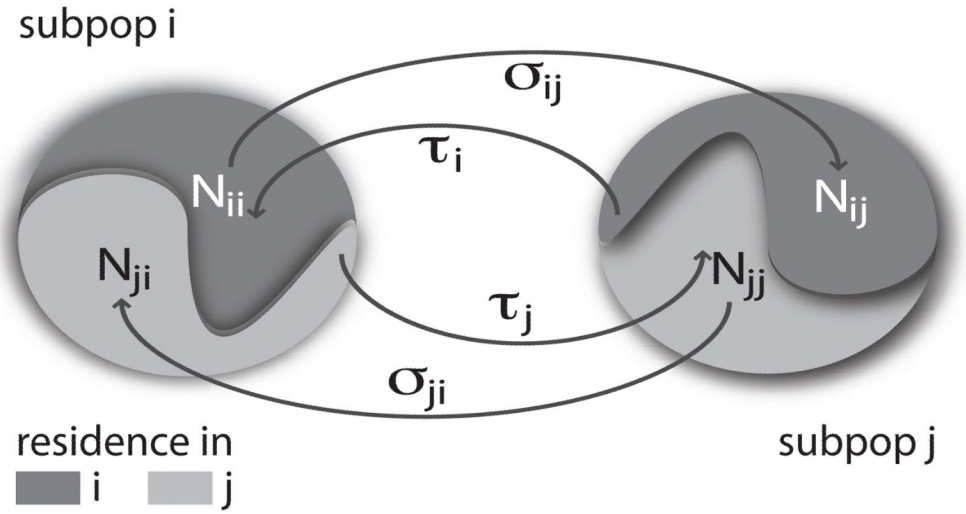

Figure 1.
Illustration of commuting and subdivision of population. At any time each subpopulation is occupied by its own residents plus visitors from its neighbors. For instance, the population in subpopulation *i* is divided between individuals who reside and are present in the subpopulation (N_{ii}) and those who are residents in subpopulation *j* but present in subpopulation *i*(N_{ji}). Different classes of people move between connected subpopulations along the edges at the rates shown.

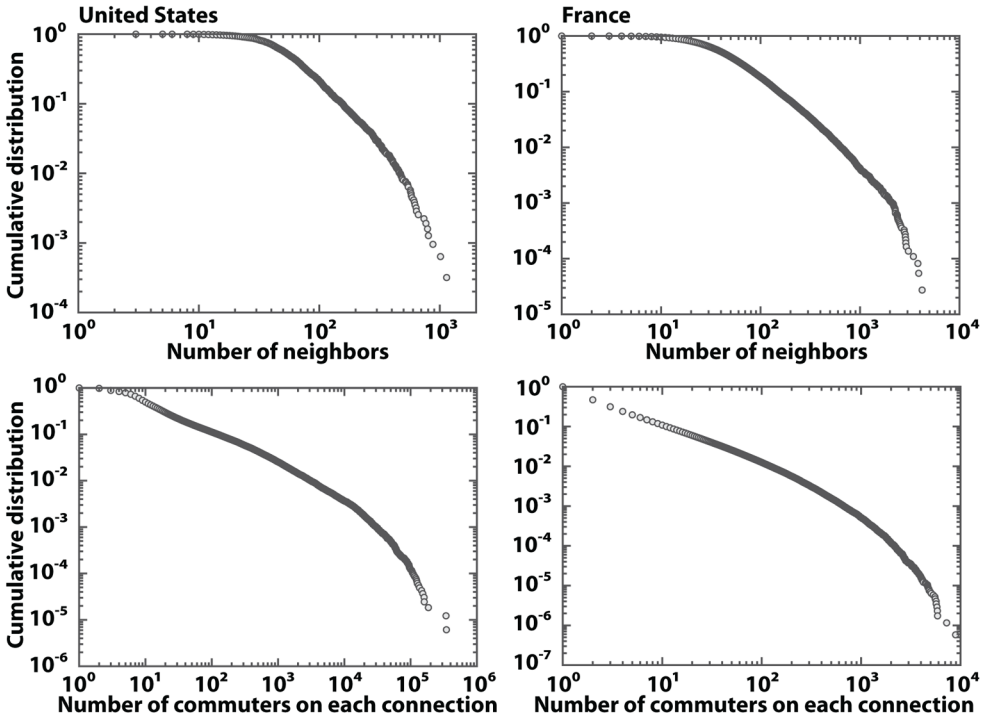

Figure 2.
Statistical properties of commuting networks in the United States and France. Cumulative distributions of the number of connections per administrative unit and the number of commuters on each connection are displayed. The networks are highly heterogeneous in the number of connections per geographical area as well as in the flux of individuals on each connection.

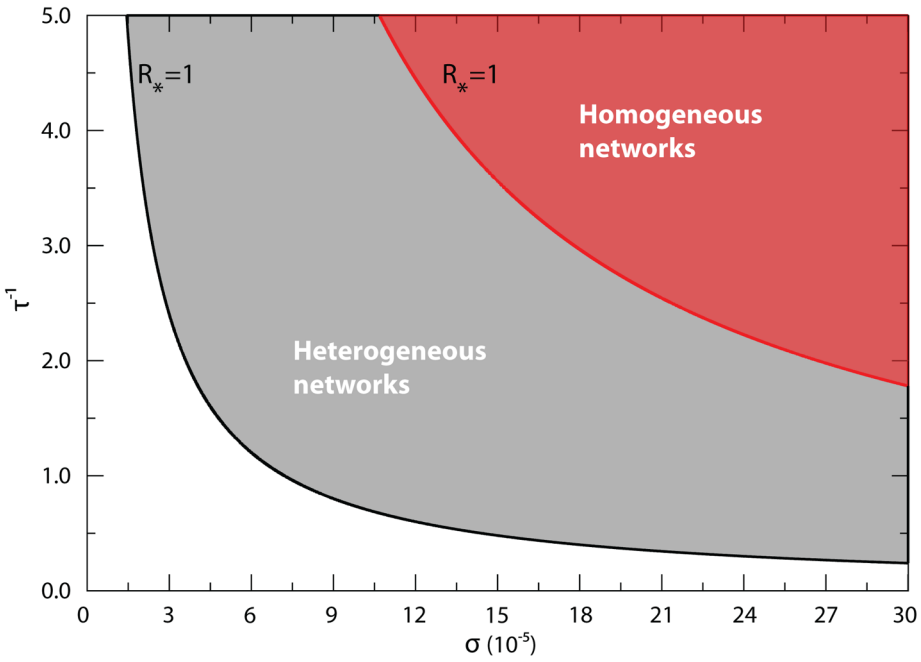

Figure 3.

Phase diagram on the σ-τ^{-1} plane for the case of Eq. (9) in heterogeneous and homogeneous subpopulation networks. The phase diagram separating the global invasion from the extinction regime is shown on the σ-τ^{-1} plane assuming Eq. (9) for commuting rates. The solid lines correspond to the solution $R_* = 1$ of Eq. (10), above which the infection spreads at the metapopulation level as indicated by the shaded areas. We can easily see an increase of about one order of magnitude in the critical values of mobility rates as we switch from a heavy-tailed to a Poisson degree distribution. Networks have the same average degree and contain $V = 10^4$ subpopulations in which the heavy-tailed network assumes $P(k) \sim k^{-2.1}$. Each subpopulation accommodates a degree-dependent population of $\bar{N}_k = \bar{N}k/\langle k \rangle$ individuals with $\bar{N} = 10^4$. Moreover the disease is characterized by $R_0 = 1.25$ and $\mu^{-1} = 15$ *days*.

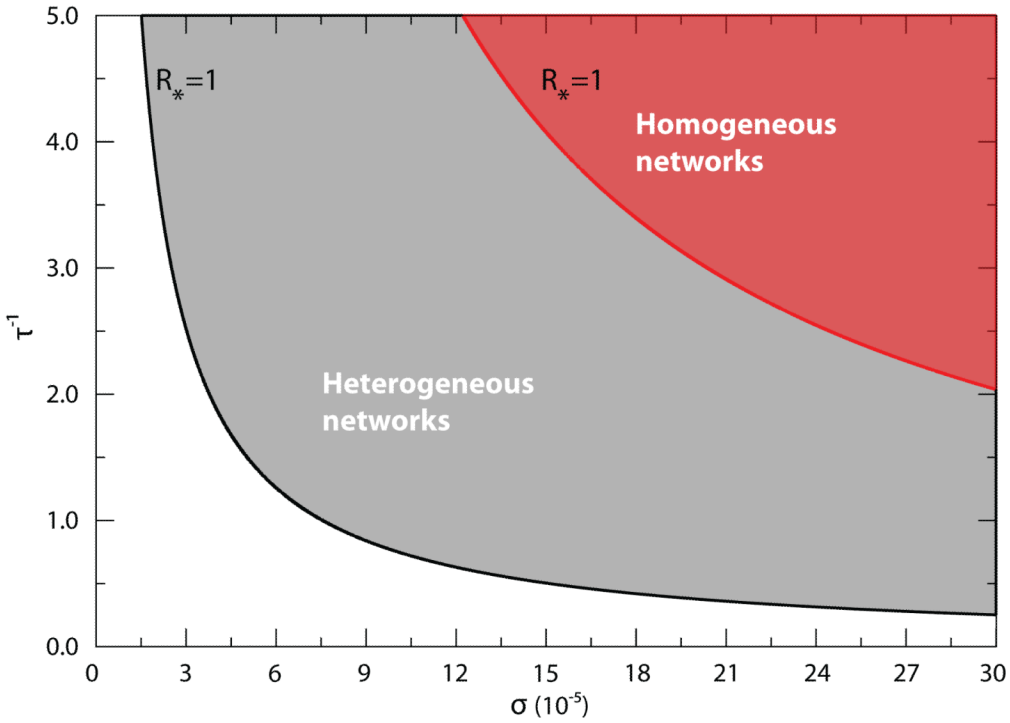

Figure 4.

Phase diagram on the σ-τ^{-1} plane for the case of Eq. (15) in heterogeneous and homogeneous subpopulation networks. The phase diagram separating the global invasion from the extinction regime is shown on the σ-τ^{-1} plane assuming Eq. (15) for commuting rates. The solid lines correspond to the solution $R_* = 1$ of Eq. (16), above which the infection spreads at the metapopulation level as indicated by the shaded areas. The diagrams should be compared with Fig. 3.

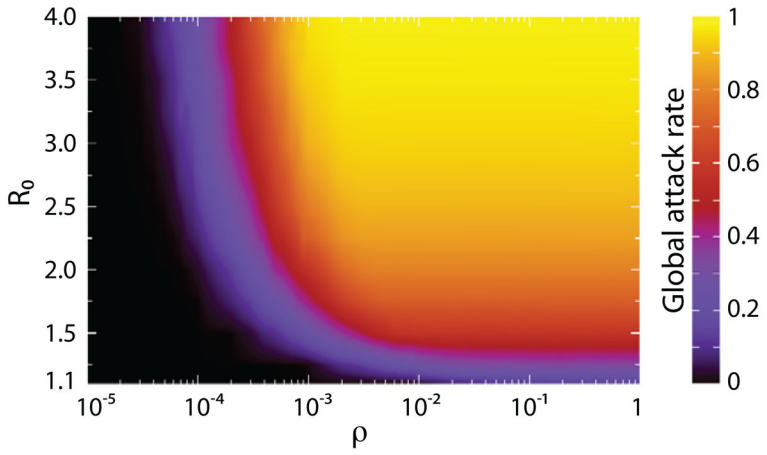

Figure 5.

Average global attack rate as a function of ρ and R_0 in homogeneous subpopulation networks. The figure codes with color the average fraction of individuals infected by the outbreak in the space of ρ and R_0. For smaller values of R_0 larger values of ρ are needed for the infection to spread at the global scale. Once ρ is well below or above its critical value its precise value does not affect the attack rate. In order to vary ρ, we have fixed the return rate τ at 1/*day* and changed the value of the commuting rate σ. Networks are made of $V = 10^4$ subpopulations, each of which accommodates a degree-dependent population of $N_k = Nk/\langle k \rangle$ individuals with $N = 104$. The infectious period is set to $\mu^{-1} = 3$ *day*.

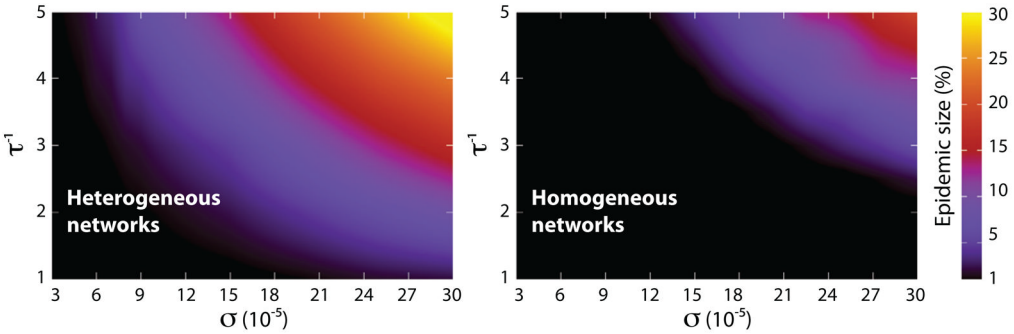

Figure 6.

Average epidemic size as a function of σ and τ^{-1} in heterogeneous (left) and homogeneous (right) subpopulation networks. Each figure shows via a color map the average percentage of subpopulations affected by the outbreak in the space of σ and τ^{-1}. The lower the commuting rate the longer the visiting time is needed for the infection to spread to a finite fraction of subpopulations. The figure should be compared with the phase diagrams of Fig. 4. Networks are composed of $V = 10^4$ subpopulations, each of which with a degree-dependent population of $\bar{N}_k = \bar{N}k/\langle k \rangle$ residents with $\bar{N} = 10^4$. Disease is characterized by $R_0 = 1.25$ and $\mu^{-1} = 15$ *day*.

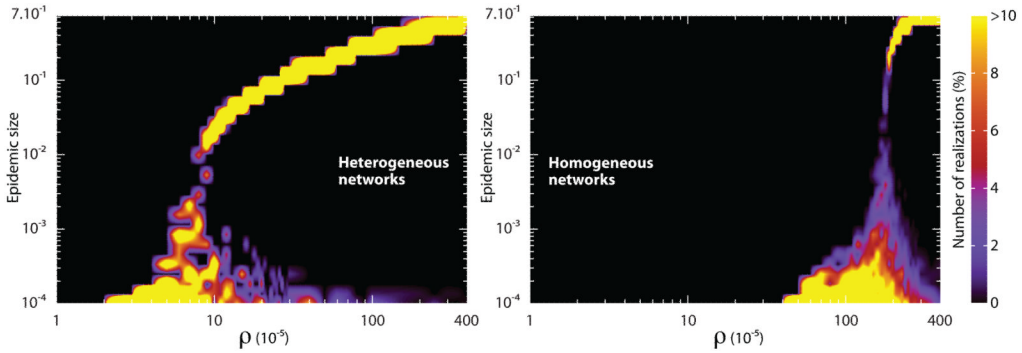

Figure 7.

Distribution of epidemic sizes in heterogeneous (left) and homogeneous (right) subpopulation networks as a function of ρ. Each color map shows the probability of observing a fraction of subpopulations affected by the outbreak as a function of commuting ratio. Each figure readily shows a critical point of ρ, below which most of the epidemics are confined to a few number of subpopulations. Above the critical point most of the realizations affect almost the entire set of subpopulations. In order to highlight the differences with respect to network topology, realizations resulting in epidemics confined to less than ten subpopulations have been excluded from the analysis of size distributions. The critical value of the commuting ratio differs more than one order of magnitude as we switch from a heavy-tailed to a Poisson degree distribution. Since the precise value of the return rate does not alter the results, we have set $\tau^{-1} = 1$ *day* and changed σ in order to vary ρ. Both networks contain $V = 10^5$ subpopulations, each of which accommodates a degree dependent population of $N_k = Nk/\langle k \rangle$ inhabitants with $N = 10^3$. The disease is characterized by $R_0 = 1.5$ and $\mu^{-1} = 5$ *day*.

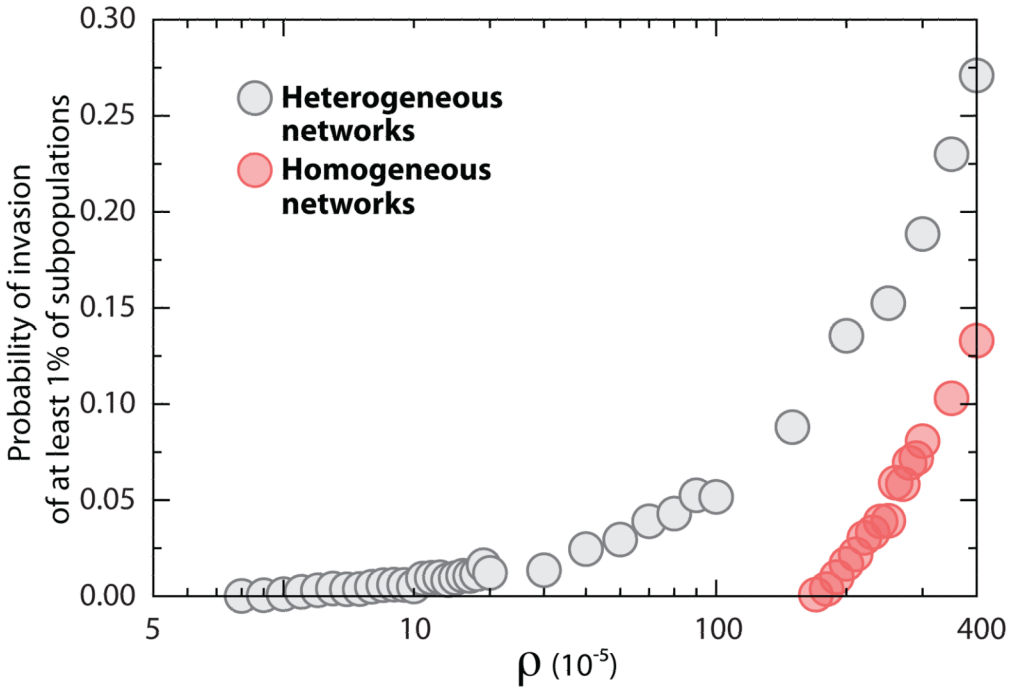

Figure 8.

Probability of invasion of at least 1% of the subpopulations in heterogeneous and homogeneous networks as a function of ρ. Figure shows the fraction of realizations in which at least 1% of the subpopulations are invaded by the epidemic process as a function of commuting ratio. For clarity, the data points corresponding to zero values of probability have been excluded from the figure. We have set $\tau^{-1} = 1$ *day* and changed σ in order to vary ρ. Both networks contain $V = 10^5$ subpopulations, each of which accommodates a degree dependent population of $\bar{N}_k = \bar{N}k/\langle k \rangle$ inhabitants with $N = 10^3$. The disease is characterized by $R_0 = 1.5$ and $\mu^{-1} = 5$ *day*.

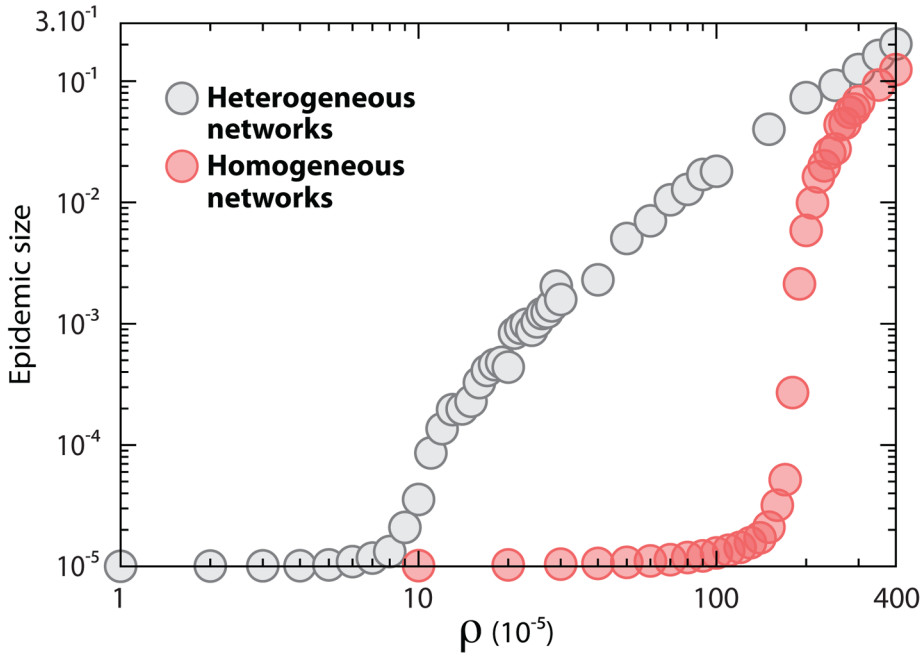

Figure 9.

Average epidemic size in heterogeneous and homogeneous subpopulation networks as a function of ρ. Figure displays the average fraction of subpopulations affected by the outbreak as a function of commuting ratio. We have set $\tau^{-1} = 1$ *day* and changed σ in order to vary ρ. Both networks contain $V = 10^5$ subpopulations, each of which accommodates a degree dependent population of $N_k = Nk/\langle k \rangle$ inhabitants with $N = 10^3$. The disease is characterized by $R_0 = 1.5$ and $\mu^{-1} = 5$ *day*.