Available online at www.sciencedirect.com
ScienceDirect
Journal of Theoretical Biology 251 (2008) 450–467
Journal of Theoretical Biology
www.elsevier.com/locate/yjtbi

Epidemic modeling in metapopulation systems with heterogeneous coupling pattern: Theory and simulations

Vittoria Colizza[a,*], Alessandro Vespignani[b,c]

[a]Complex Networks Lagrange Laboratory (CNLL), Institute for Scientific Interchange (ISI), Torino, Italy
[b]School of Informatics and Biocomplexity Institute, Indiana University, Bloomington, IN 47408, USA
[c]Institute for Scientific Interchange (ISI), Torino, Italy

Received 25 June 2007; received in revised form 25 September 2007; accepted 26 November 2007
Available online 28 January 2008

Abstract

The spatial structure of populations is a key element in the understanding of the large-scale spreading of epidemics. Motivated by the recent empirical evidence on the heterogeneous properties of transportation and commuting patterns among urban areas, we present a thorough analysis of the behavior of infectious diseases in metapopulation models characterized by heterogeneous connectivity and mobility patterns. We derive the basic reaction–diffusion equations describing the metapopulation system at the mechanistic level and derive an early stage dynamics approximation for the subpopulation invasion dynamics. The analytical description uses a homogeneous assumption on degree block variables that allows us to take into account arbitrary degree distribution of the metapopulation network. We show that along with the usual single population epidemic threshold the metapopulation network exhibits a global threshold for the subpopulation invasion. We find an explicit analytic expression for the invasion threshold that determines the minimum number of individuals traveling among subpopulations in order to have the infection of a macroscopic number of subpopulations. The invasion threshold is a function of factors such as the basic reproductive number, the infectious period and the mobility process and it is found to decrease for increasing network heterogeneity. We provide extensive mechanistic numerical Monte Carlo simulations that recover the analytical finding in a wide range of metapopulation network connectivity patterns. The results can be useful in the understanding of recent data driven computational approaches to disease spreading in large transportation networks and the effect of containment measures such as travel restrictions.

Keywords: Metapopulation models; Epidemic spreading; Complex networks

1. Introduction

The metapopulation modeling approach is an essential theoretical framework used in population ecology, genetics and adaptive evolution to describe population dynamics whenever the spatial structure of populations is known to play a key role in the system's evolution (Hanski and Gilpin, 1997; Hanski and Gaggiotti, 2004; Tilman and Kareiva, 1997; Bascompte and Solé, 1998). Metapopulation models rely on the basic assumption that the system under study is characterized by a highly fragmented environment in which the population is structured and localized in relatively isolated discrete *patches* or subpopulations connected by some degree of migration. Classic metapopulation dynamics focuses on the processes of local extinction, recolonization and regional persistence (Levins, 1969, 1970), as the outcome of the interplay between migration processes among unstable local populations and population dynamics (e.g. birth and death rates, competition and predations). This paradigm is extremely useful also in the case of infectious diseases, and can be applied to understand the epidemic dynamics of spatially structured populations with well-defined social units (e.g. families, villages, city locations, towns, cities, regions) connected through individual mobility (Hethcote, 1978; May and

*Corresponding author at: Institute for Scientific Interchange (ISI), Torino, Italy.
E-mail addresses: vcolizza@isi.it (V. Colizza), alexv@indiana.edu (A. Vespignani).

0022-5193/$ - see front matter © 2007 Elsevier Ltd. All rights reserved.
doi:10.1016/j.jtbi.2007.11.028

Anderson, 1979; Anderson and May, 1984; May and Anderson, 1984; Bolker and Grenfell, 1993, 1995; Keeling and Rohani, 2002; Lloyd and May, 1996; Grenfell and Harwood, 1997; Grenfell and Bolker, 1998; Ferguson et al., 2003; Riley, 2007). The arrival of the infection in any subpopulation and its epidemic evolution are determined by the coupling generated by the mobility processes among subpopulations. The metapopulation dynamics of infectious diseases has generated a wealth of models and results considering both mechanistic approaches taking explicitly into account the movement of individuals (Baroyan et al., 1969; Rvachev and Longini, 1985; Longini, 1988; Flahault and Valleron, 1991; Sattenspiel and Dietz, 1995; Keeling and Rohani, 2002; Grais et al., 2003; Ruan et al., 2006) and effective coupling approaches where the diffusion process is expressed as a force of infection coupling different subpopulations (Bolker and Grenfell, 1995; Lloyd and May, 1996; Earn et al., 1998; Rohani et al., 1999; Keeling, 2000; Park et al., 2002; Vázquez, 2007). Recently, the metapopulation approach is being revamped in computational approaches for the large-scale forecast of infectious disease spreading (Grais et al., 2004; Hufnagel et al., 2004; Colizza et al., 2006a, 2007a; Cooper et al., 2006; Hollingsworth et al., 2006; Riley, 2007).

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. The increasing computational power and informatics advances are beginning to lift the constraints limiting the collection of large spatiotemporal data on human behavior and demography, finally allowing for the formulation of realistic data driven models. On the other hand, networks which trace the activities and interactions of individuals, social patterns, transportation fluxes and population movements on a local and global scale (Liljeros et al., 2001; Schneeberger et al., 2004; Barrat et al., 2004; Guimerá et al., 2005; Chowell et al., 2003) have been analyzed and found to exhibit complex features encoded in large-scale heterogeneity, self-organization and other properties typical of complex systems (Albert and Barabási, 2002; Dorogovtsev and Mendes, 2003; Newman, 2003; Pastor-Satorras and Vespignani, 2004). In particular, a wide range of societal and technological networks exhibits highly heterogeneous topologies. The airport network among cities (Barrat et al., 2004; Guimerá et al., 2005), the commuting patterns in inter- and intra-urban areas (Chowell et al., 2003; Barrett et al., 2000; De Montis et al., 2007), and several info-structures (Pastor-Satorras and Vespignani, 2004) are indeed characterized by networks whose nodes, representing the elements of the system, have a wildly varying degree, i.e. the number of connections to other elements. These topological fluctuations are mathematically encoded in a heavy-tailed degree distribution $P(k)$, defined as the probability that any given node has degree k, and have been found to have a large impact on epidemic phenomena on complex contact patterns (Anderson and May, 1992; Pastor-Satorras and Vespignani, 2001a, b; Moreno et al., 2002; Lloyd and May, 2001; Barthélemy et al., 2005).

Motivated by the above findings we provide here the analysis of the behavior of epidemic models in metapopulation systems with heterogeneous connectivity patterns. In order to have a mechanistic description of the system, we derive the deterministic reaction–diffusion equations describing the evolution of the epidemic in the metapopulation systems. The heterogeneity of the network is taken explicitly into account by partitioning subpopulations according to degree block variables. By using a homogenous approximation for subpopulations with the same degree, it is possible to provide explicit results for the system behavior expressed as functions of the moments of the degree distribution of the substrate networks. In order to account for the discreteness of the system and microscopic fluctuations in the diffusion processes we derive also coarse-grained equations for the invasion dynamics at the subpopulation level. The system is characterized by the standard (i.e. single population) epidemic threshold and by a global invasion threshold providing the condition for the infection of a macroscopic number of subpopulations. The first threshold defines the usual reproductive number $R_0 > 1$ that is just a function of the disease parameters while the second threshold defines a subpopulations reproductive number $R_* > 1$ that depends also on the diffusion rate of individuals among subpopulations (Ball et al., 1997; Cross et al., 2005, 2007). We find an explicit analytic expression in the limit $R_0 \gtrapprox 1$ for the invasion threshold that is found to depend also on the network heterogeneity. The larger is the network heterogeneity and the smaller is the diffusion rate that guarantees the invasion of a finite fraction of subpopulations. This result provides a framework for the understanding of the evidence collected on the interplay between travel and global spread of infectious diseases (Viboud et al., 2006) and the poor effectiveness of travel restrictions in the containment of epidemics (Cooper et al., 2006; Hollingsworth et al., 2006; Colizza et al., 2007a). Finally, the analytic results are confirmed by mechanistic Monte Carlo simulations for the infection dynamics in the metapopulation system, 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 paper is organized as follows. Section 2 introduces the metapopulation epidemic model on a heterogeneous network of connections among subpopulations. Two different kinds of mobility processes are introduced in Section 3 to analyze the stationary diffusion properties of the system. Section 4 incorporates the mobility processes analyzed in the previous section into a metapopulation epidemic model. Stochastic effects and discrete description of the processes are considered with a tree-like approximation for the analysis of the invasion dynamics at the level of

the subpopulations. The effect 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 the behavior of the system above the invasion threshold is studied by mechanistic reaction–diffusion equations using a deterministic degree block variables representation. Finally, in Section 6 we report extensive mechanistic Monte Carlo simulations which confirm the analytical findings of the previous sections.

2. Metapopulation mechanistic model as a microscopic reaction–diffusion process

Metapopulation models describe spatially structured interacting subpopulations, such as city locations, urban areas, or defined geographical regions (Hanski and Gaggiotti, 2004; Grenfell and Harwood, 1997). Individuals within each subpopulation are divided into classes denoting their state with respect to the modeled disease (Anderson and May, 1992)—such as infected, susceptible, immune, etc.—and the compartment dynamics accounts for the possibility that individuals in the same location may get into contact and change their state according to the infection dynamics. The interaction among subpopulations is the result of the movement of individuals from one subpopulation to the other. It is clear that the key issue in such a modeling approach is how accurately we can describe the commuting patterns or traveling of people. In many instances even complicate mechanistic patterns can be accounted for by effective couplings expressed as a force of infection generated by the infectious individuals in subpopulation j on the individuals in subpopulation i (Bolker and Grenfell, 1995; Lloyd and May, 1996; Earn et al., 1998; Rohani et al., 1999; Keeling, 2000; Park et al., 2002). More realistic descriptions are provided by explicit mechanistic approaches which include the detailed rate of traveling/commuting obtained from data or from empirical fit to gravity law models (for a recent reference, see Viboud et al., 2006), accompanied by the associated mixing subpopulations N_{ij} denoting the number of individuals of the subpopulation i present in the subpopulation j (Keeling and Rohani, 2002; Sattenspiel and Dietz, 1995; Ruan et al., 2006).

A simplified mechanistic approach uses a markovian assumption in which at each time step the movement of individuals is given according to a matrix d_{ij} that expresses the rate at which an individual in the subpopulation i is traveling to the subpopulation j in the unit time. The markovian character is in the fact that we do not label individuals according to their original subpopulation (e.g. *home* in a commuting pattern framework) and at each time step the same traveling probability applies to all individuals in the subpopulation without having memory of their origin. This approach is extensively used for very large populations in the case the traffic w_{ij} among subpopulations is known by stating that $d_{ij} \sim w_{ij}/N_j$. Several modeling approaches to the large-scale spreading of infectious disease (Baroyan et al., 1969; Rvachev and Longini, 1985; Longini, 1988; Flahault and Valleron, 1991; Grais et al., 2003, 2004; Hufnagel et al., 2004; Colizza et al., 2006a, b, 2007a) use this mobility process based on transportation networks for which it is now possible to obtain detailed data.

In their simplest formulation markovian mechanistic mobility processes are equivalent to the classic reaction–diffusion processes used in many physical, chemical and biological processes (Marro and Dickman, 1999; van Kampen, 1981; Murray, 2005). The reaction–diffusion framework (Colizza et al., 2007b) considers that the occupation numbers N_i of each subpopulation can have any integer value, including $N_i = 0$, that is, void nodes with no individuals. The total population of the metapopulation system is $N = \sum_i N_i$ and each individual diffuses along the edges with a diffusion coefficient d_{ij} that depends on the node degree, subpopulation size and/or the mobility matrix. A sketch of the metapopulation model which shows the different scales of the system is shown in Fig. 1. The system is composed of a network substrate connecting subpopulations over which individuals diffuses. Each subpopulation is represented by a node i of the network. We consider that each node i is connected to other k_i nodes according to its degree resulting in a network with degree distribution $P(k)$ and distribution moments $\langle k^\alpha \rangle = \sum_k k^\alpha P(k)$.

In the case of large metapopulation systems with a high level of heterogeneity the analytical description of the metapopulation model in terms of specific features of each single subpopulation is extremely complicate. In the following we propose an analytical framework that uses degree block variables to obtain the dynamical equations describing the system's behavior, relying on the empirical evidence pointing to a statistical equivalence of subpopulations having the same degree.

2.1. Metapopulation networks with heterogeneous topology

In the real world, the network specifying the coupling between different subpopulations is in many cases very heterogeneous. Examples can be drawn from several transportation infrastructures, commuting data and census information (Chowell et al., 2003; Barrett et al., 2000; Barrat et al., 2004; Guimerá et al., 2005; De Montis et al., 2007). A particularly relevant one in the field of epidemic modeling is given by the airline transportation network. In this case the coupling is provided by the number of passengers traveling on a given route connecting two airports, thus yielding a transfer of individuals between the corresponding urban areas. For instance, Barrat et al. (2004) reports a detailed study of the International Air Transport Association[1] database which contains the complete list of world commercial airport pairs connected by direct flights. Moreover, to each direct flight connection

[1]IATA, International Air Transport Association, http://www.iata.org/.

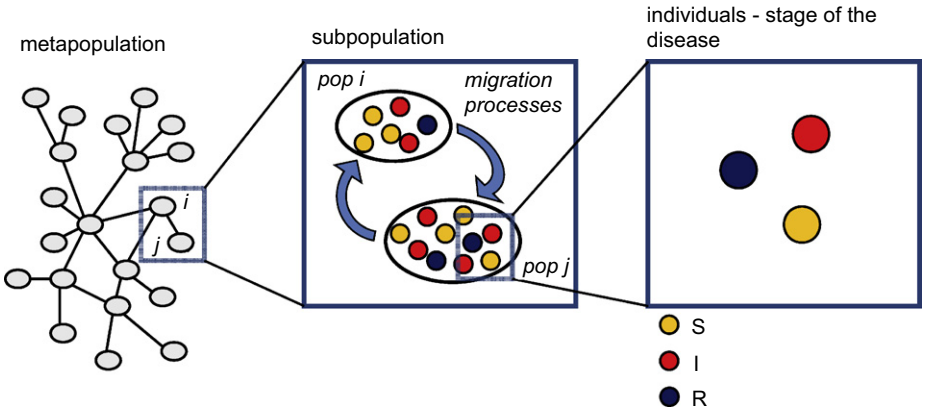

Fig. 1. Schematic representation of a metapopulation model. The system is composed of a heterogeneous network of subpopulations or patches, connected by migration processes. Each patch contains a population of individuals who are characterized with respect to their stage of the disease (e.g. susceptible, infected, removed), and identified with a different color in the picture. Individuals can move from a subpopulation to another on the network of connections among subpopulations.

between airports j and ℓ is assigned a weight $w_{j\ell}$ which corresponds to the number of available seats or passengers on the given route. The obtained network displays high levels of heterogeneity both in the connectivity pattern and in the traffic capacities, as revealed by the broad distributions of the number of connections of each airport, of the travel flows between connected airports and of the traffic in terms of number of passengers handled by each airport (Barrat et al., 2004). These results have been confirmed on the US subnetwork along different years and considering both market and segment traffic data[2] and analogous results are recovered by analyzing commuting patterns data, intra-city traffic among locations and several other data sets concerning the movements of people and goods (Chowell et al., 2003; Barrett et al., 2000; Barrat et al., 2004; Guimerá et al., 2005; De Montis et al., 2007). In Fig. 2 we report the degree and weight probability distributions in some examples of these networks. In many cases we find heavy-tailed distributions varying overseveral orders of magnitude. For instance, the airline traffic among different urban areas in the world shows a probability distribution $P(w)$—where w is the traffic on a single connection—varying over six orders of magnitude (see Fig. 2B).

In addition, it is also possible to find some general statistical laws relating the traffic and the degree of each node in the network. In general the behavior of the average weight along the connection between two subpopulations with degree k and k' is a function of their degree

$$\langle w_{kk'}\rangle = w_0(kk')^{\theta}, \tag{1}$$

where w_0 and the exponent θ depend on the specific system (e.g. $\theta \simeq 0.5$ in the world-wide air transportation network (Barrat et al., 2004)). A related quantity is the total average traffic per unit time T_k of the subpopulations with degree k that behaves as

$$T_k = Ak^{(1+\theta)}. \tag{2}$$

[2]BTS, Bureau of Transportation Statistics, http://www.bts.org/.

Here the proportionality constant A and the exponent $1+\theta$ are defined by the sum rule $T_k = k\sum_{k'} P(k'|k)w_0(kk')^{\theta}$ that must be satisfied on average. Here $P(k'|k)$ represents the conditional probability that any given edge departing from a node of degree k is pointing to a node of degree k'. In the following we will consider the case of uncorrelated networks in which the conditional probability does not depend on the originating node, i.e. $P(k'|k) = k'P(k')/\langle k\rangle$ (Dorogovtsev and Mendes, 2003; Pastor-Satorras et al., 2001). This relation simply states that any edge has a probability to point to a node with degree k' that is proportional to the degree of the node. By using this form for $P(k'|k)$, we obtain the proportionality constant yielding $A = \langle k^{1+\theta}\rangle w_0/\langle k\rangle$. It is important to stress that the above relations defines a statistical equivalence of the subpopulations of degree k. In the following we want to address the questions of how the large-scale complex features (scale invariance, extreme heterogeneity, unbounded fluctuations) of interaction and communication networks affect the behavior of metapopulation models by defining a mechanistic approach based on the master equations describing the disease dynamics as a microscopic reaction—diffusion process. The equivalence assumption that will be used throughout the rest of this paper is crucial in order to carry out the analytical treatment of the model. While this assumption is indeed recovered in several data sets, exceptions and fluctuations have been noted that would require more complicate calculation schemes.

3. Mobility processes and diffusion properties in heterogeneous networks

In order to tackle the description of metapopulation models at the mechanistic level, let us first analyze the simple diffusion process of a global population of N individuals who diffuse in a network made of V nodes, each representing a subpopulation. Each node i stores a number N_i of individuals as defined in Section 2. In order to take into account the topological fluctuations of the

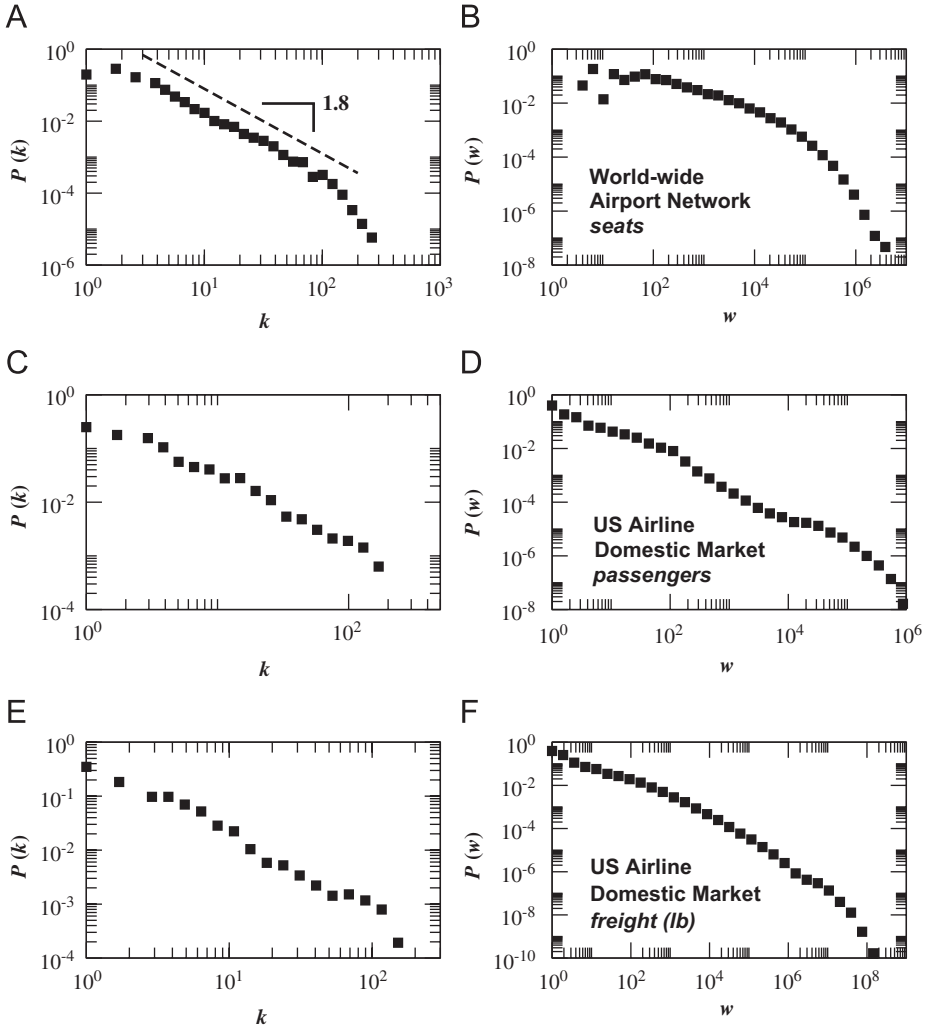

Fig. 2. Degree (left column) and weight (right column) probability distributions for three different data sets of the movement of people and goods: (A), (B) world-wide airport network, where the weight represents the number of seats on the flights connecting two different airports (source: IATA); (C), (D) US Air Domestic Market, where the weight represents the number of passengers flying on a given origin–destination itinerary (source: BTS); (E), (F) US Air Domestic Market for freight transportation, where the weight represents the amount of freight (expressed in lb) transported from an origin airport to its final destination airport (source: BTS). All data sets show heavy-tailed distributions both in the number of connections and in the amount of people/good transported.

network we have to explicitly consider the presence of nodes with a widely fluctuating degree k. A more convenient representation of the system is therefore provided by the quantities defined in terms of the degree k

$$N_k = \frac{1}{V_k} \sum_{i|k_i=k} N_i, \tag{3}$$

where V_k is the number of nodes with degree k and the sums run over all nodes i having degree k_i equal to k. The degree block variable N_k is therefore representing the average number of individuals in all subpopulations with degree k. This representation corresponds to a homogeneous approximation that does not take into account possible variations within subpopulations having the same degree k, assuming they are statistically equivalent. While the degree-block description is clearly an approximation, as discussed in the previous section, several data sets provide empirical evidence for this assumption. In addition, as for similar homogenous assumptions in single population models, it allows for an explicit analytic solution of the metapopulation model while taking into account the network heterogeneity. Naturally, realistic models would need to include specific features of each subpopulations, introducing the heterogeneities that distinguish subpopulations of the same degree.

Let us assume a general framework in which the individuals move from a subpopulation with degree k to another with degree k' with a diffusion rate $d_{kk'}$ that depends on the degrees of the origin and destination subpopulations. The rate at which individuals leave a subpopulation with degree k is then given by $p_k = k\sum_{k'} P(k'|k)d_{kk'}$. In the following, we will first write the equations for the dynamics of the individuals under this

generic type of diffusion, and then address two specific diffusion rates to find the stationary solutions.

The dynamics of individuals is simply represented by a mean-field dynamical equation expressing the variation in time of the subpopulations $N_k(t)$ in each degree block. This can be easily written as

$$\partial_t N_k(t) = -p_k N_k(t) + k \sum_{k'} P(k'|k) d_{k'k} N_{k'}(t). \quad (4)$$

The first rhs term of the equation just considers that only a fraction of particles p_k moves out of the node per unit time. The second term instead accounts for the particles diffusing from the neighbors into the node of degree k. This term is proportional to the number of links k times the average number of particles coming from each neighbors. This is equal to average over all possible degrees k' the fraction of particles moving on that edge $d_{k'k}N_{k'}(t)$ according to the conditional probability $P(k'|k)$. By assuming uncorrelated networks, the dynamical rate equation (4) for the subpopulation sizes reads as

$$\partial_t N_k(t) = -p_k N_k(t) + \frac{k}{\langle k \rangle} \sum_{k'} k' P(k') d_{k'k} N_{k'}(t). \quad (5)$$

In the following subsections we solve the previous set of equations for different diffusion processes that consider diffusion rates depending on the traffic of each node or on the population size of each subpopulation.

3.1. Traffic dependent mobility rates

Here we assume that the rate at which an individual leaves a given subpopulation is independent of its degree k, $p_k = p \ \forall k$. If we also assume homogeneous diffusion along any given connection, individuals have the same probability to move along anyone of the links departing from the node at which they are located. In this case the diffusion rate along any given link of a node with degree k will be simply equal to

$$d_{kk'} = p/k. \quad (6)$$

This is obviously not the case in a wide range of real systems where the extreme heterogeneity of traffic is well documented (see Subsection 2.1). A more realistic process therefore considers the movement of individuals to be proportional to the traffic intensity along a given edge. This is simply obtained by defining a heterogeneous diffusion rate for any given individual to go from a subpopulation of degree k to a subpopulation of degree k' as

$$d_{kk'} = p \frac{w_0 (kk')^\theta}{T_k}. \quad (7)$$

This relation states that the diffusion rate p is still constant in each subpopulation but the individuals move on each connection in a proportion dependent on the actual traffic on the connection. The denominator $T_k = Ak^{(1+\theta)}$ provides the correct normalization in order to ensure that by summing over all k edges departing from the node the overall diffusion rate is p.

By using the expression of Eq. (7) for $d_{kk'}$ and imposing $p_k = p$, the dynamical rate equation (4) for the subpopulation sizes reads as

$$\partial_t N_k(t) = -pN_k(t) + pk^{(1+\theta)} \frac{w_0}{A\langle k \rangle} \sum_{k'} P(k') N_{k'}(t). \quad (8)$$

The stationary solution $\partial_t N_k(t) = 0$ does not depend upon the diffusion rate p that just fixes the time scale at which the equilibrium is reached and has the solution

$$N_k = k^{(1+\theta)} \frac{w_0}{A\langle k \rangle} \bar{N}, \quad (9)$$

where $\bar{N} = \sum_k P(k) N_k(t)$ represents the average subpopulation size. The explicit form of the normalization constant $A = \langle k^{1+\theta} \rangle w_0 / \langle k \rangle$ finally provides the explicit stationary solution

$$N_k = \frac{k^{(1+\theta)}}{\langle k^{(1+\theta)} \rangle} \bar{N}. \quad (10)$$

The above solution states that the population of each node scales with the node degree in the stationary limit. The above behavior is simply the effect of the diffusion process that brings a large number of individuals in well connected, high traffic nodes, thus showing the impact of network's topological (i.e. dependence on k) and traffic (i.e. dependence on θ) fluctuations on the subpopulation size behavior. When $\theta = 0$ we recover the homogeneous diffusion case in which $d_{kk'} = d_k = p/k$, obtaining

$$N_k = \frac{k}{\langle k \rangle} \bar{N}. \quad (11)$$

In this case the subpopulation size is just fixed from topological fluctuations and the exponent θ clearly appears as the parameter that takes into account the traffic fluctuations. It is worth remarking that in this framework, the subpopulation size as a function of the degree is constrained by the diffusion processes, a feature that has not to be expected in real systems where the population size of local patches can be considered as an independent variable. On the other hand, the degree dependence is close to those observed in real systems where in several cases it is possible to find a relation $N_k \sim k^\phi$ with $0.5 \leqslant \phi \leqslant 1.5$ (Colizza et al., 2006a, b). We can thus consider the obtained stationary state as a first approximation to the real case and use the exponent θ to explore different levels of heterogeneity.

3.2. Population dependent mobility rates

In a more general perspective, it is important to have the possibility of considering the population sizes N_k as independent variables. This is indeed the case of many metapopulation models in which the diffusion process represents the large-scale travel of individuals between subpopulations, e.g. via air travel. In this framework the

number of people traveling from a subpopulation to the other in a unitary time scale is a defined number w_{ij} and the number of traveling individuals is independent from the population size N_i. This amounts to state that each individual in the subpopulation has a diffusion rate $\sum_j w_{ij}/N_i$, where $\sum_j w_{ij}$ is the total number of people traveling out of city i in the unitary time scale. In other words, the diffusion rate of each individual is inversely proportional to the population size. Empirical evidence obtained from the study of the world-wide air transportation network (Barrat et al., 2004) showed the symmetry of the network both in the directionality and in the travel fluxes. Therefore the condition $w_{ij} = w_{ji}$ is satisfied in large-scale real transportation networks and one obtains

$$\partial_t N_i = \sum_j (w_{ji} - w_{ij}) = 0, \tag{12}$$

so that any initial conditions for the population size satisfy the stationary state. In the degree block variable representation we can recover the above condition by considering a diffusion rate for each particle of the form $p_k = T_k/N_k$. The diffusion rate on any given edge from a subpopulation of degree k to a subpopulation of degree k' is therefore given by

$$d_{kk'} = \frac{w_0 (kk')^\theta}{N_k} \tag{13}$$

and the degree block diffusion equations read in the uncorrelated networks case as

$$\partial_t N_k(t) = -T_k + k^{(1+\theta)} w_0 \frac{\langle k^{1+\theta} \rangle}{\langle k \rangle}. \tag{14}$$

Since we know that by normalization $T_k = k^{(1+\theta)} w_0 \langle k^{1+\theta} \rangle / \langle k \rangle$, we recover by definition the solution $\partial_t N_k(t) = 0$ that allows any stationary value distribution N_k. Differently from the results obtained in the previous subsection, where each individual has the same diffusion rate p of leaving a subpopulation, Eq. (12) shows that a population dependent diffusion process does not fix the subpopulation size, which can be given as a parameter of the model.

4. Epidemic spreading and the invasion threshold

In order to explore the epidemic behavior in metapopulation models, the disease dynamics needs to be explicitly considered inside each subpopulation. In the following we will consider the standard compartmentalization approach in which individuals exist in a certain number of discrete states such as susceptible, infected or permanently recovered (Anderson and May, 1992). The paradigmatic epidemiological model one can consider is the susceptible-infected-removed (SIR) model (Anderson and May, 1992; Murray, 2005), where the total number of individuals N_j in the subpopulation j is partitioned in the compartment $S_j(t)$, $I_j(t)$ and $R_j(t)$ denoting the number of susceptible, infected and recovered individuals at time t, respectively. By definition it follows $N_j = S_j(t) + I_j(t) + R_j(t)$. The disease transmission is described in an effective way. The probability that a susceptible individual acquires the infection from any given neighbor in an infinitesimal time interval $\mathrm{d}t$ is $\beta\,\mathrm{d}t$, where β defines the disease *transmissibility*. At the same time, infected vertices are cured and become recovered with probability $\mu\,\mathrm{d}t$. Individuals thus run stochastically through the susceptible → infected → recovered transitions, hence the name of the model. The SIR model assumes that recovered individuals are basically removed from the system, they do not participate anymore to the disease dynamics, due to their death or acquired immunization. Another popular model takes into account the possibility that infected individuals are again susceptible with probability $\mu\,\mathrm{d}t$. In this case individuals thus run stochastically through the cycle susceptible → infected → susceptible, defining the so-called SIS model. The SIS model is mainly used as a paradigmatic model for the study of infectious diseases leading to an endemic state with a stationary and constant value for the prevalence of infected individuals, i.e. the degree to which the infection is widespread in the population.

A basic parameter in the analysis of a single population epidemic outbreaks is the basic reproductive number R_0, which counts the number of secondary infected cases generated by a primary infected individual (Anderson and May, 1992). Under the assumption of the homogeneous mixing of the population the basic reproductive number is defined as

$$R_0 = \frac{\beta}{\mu}. \tag{15}$$

It is straightforward to see from Eq. (15) that in the single population case any epidemic will spread across a non-zero fraction of the population only for $R_0 > 1$. In this case the epidemic is able to generate a number of infected individuals larger than those who recover, leading to an increase in the overall number of infectious individuals $I(t)$. The previous considerations lead to the definition of a crucial epidemiological concept—the epidemic threshold (Anderson and May, 1992). Indeed, if the spreading rate is not large enough to allow a reproductive number larger than 1 (i.e. $\beta > \mu$), the epidemic outbreak will not affect a finite portion of the population and will die out in a finite amount of time.

At the metapopulation level, however, the epidemic behavior on the global scale is determined also by the diffusion process of individuals. In particular, the effects due to the finite size of subpopulations and the stochastic nature of the diffusion might have a crucial role in the problem of resurgent epidemics, extinction and eradication (Ball et al., 1997; Cross et al., 2005; Watts et al., 2005; Vázquez, 2007; Cross et al., 2007). Therefore it is important to consider the discrete nature of individuals. Indeed, each subpopulation may or may not transmit the infection to another subpopulation it is in contact with, depending on the occurrence or not of the travel event of at least one

infected individual to the non-infected subpopulation during the entire epidemic evolution. The spreading process across subpopulations will therefore occur with a probability that is related to the diffusion rate of individuals and the total number of individuals who will experience the infection (Ball et al., 1997; Cross et al., 2005, 2007). In the case of epidemic processes with $R_0<1$ the epidemic will die with probability 1 and is not going to spread across subpopulations. In a model like the SIS model, if $R_0>1$ the number of infected individuals reaches a stationary state and the epidemic will eventually spread to different subpopulations, since locally endemic. In a model such as the SIR model, however, the epidemic within each subpopulation generates a finite fraction of infectious individuals in a given amount of time and even if $R_0>1$ the diffusion rate must be large enough to ensure the timely diffusion of infected individuals to other subpopulations of the metapopulation system, before the local epidemic outbreak dies out. This is captured by the definition of a new predictor of disease invasion, R_*, regulating the number of subpopulations that become infected from a single initially infected subpopulation, i.e. the analogous of the reproductive number at the subpopulation level (Ball et al., 1997; Cross et al., 2005, 2007).

This effect would not be captured by a continuous deterministic description that would allow any fraction pI of diffusing infected individual to inoculate the virus in a subpopulation not yet infected (Ruan et al., 2006). In certain conditions this fraction pI may be a number smaller than 1 that just represents a mean-field average value. This is a common feature of deterministic continuous approximations that allow the infection to persist and diffuse via "nano-individuals" that are not capturing the discrete nature of the real systems. For this reason, in the next section we will use an approach working at the level of subpopulations that allows to take into account effectively the fluctuations inherent to the diffusion process and the outbreak extinction probability.

4.1. Global invasion threshold in homogeneous metapopulation networks

Let us consider a metapopulation system in which the initial condition is provided by a single introduction in a subpopulation of degree k and size N_k, given $R_0>1$. While the stochastic nature of the process may lead in some cases to the extinction of the process, as R_0 is above the epidemic threshold the epidemic will affect a finite fraction of the population with non-zero probability. In the case of a macroscopic outbreak in a closed population the total number of infected individuals during the evolution of the epidemic will be equal to αN_k where α depends on the specific disease model used and the disease parameter values. Each infected individual stays in the infectious state for an average time μ^{-1} equal to the inverse of the recovery rate, during which it can travel to the neighboring subpopulation of degree k' with rate $d_{kk'}$. To a first approximation we can therefore consider that the number of new seeds that may appear into a connected subpopulation of degree k' during the duration of the subpopulation epidemic is given by

$$\lambda_{kk'} = d_{kk'}\frac{\alpha N_k}{\mu}. \tag{16}$$

In this perspective we can consider the metapopulation model in a coarse-grained view (see Fig. 3) and provide a characterization of the invasion dynamics at the level of the subpopulations, translating epidemiological and demographic parameters into Levins-type metapopulation parameters of extinction and invasion rate. Let us define D_k^0 as the number of *diseased* subpopulation of degree k at generation 0, i.e. those which are experiencing an outbreak at the beginning of process. Each infected subpopulation during the course of the outbreak will seed the infection in neighboring subpopulations defining the set D_k^1 of infected subpopulations at the following generation and so on. This corresponds to a basic branching process (Harris, 1989;

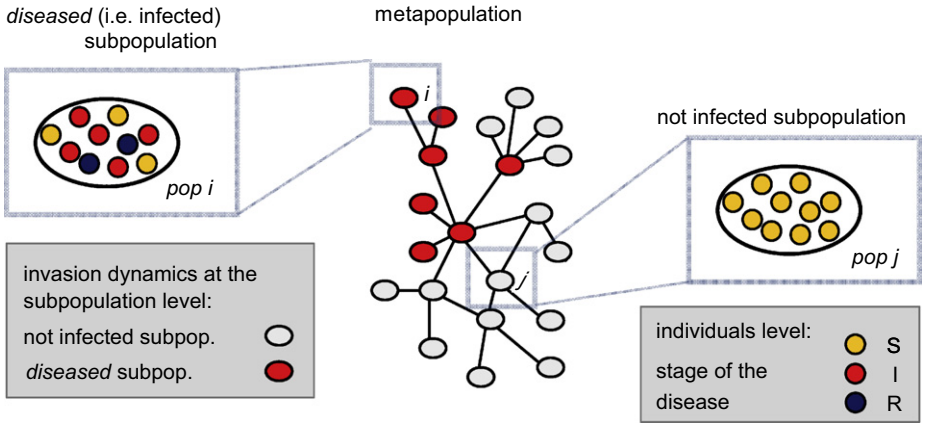

Fig. 3. Schematic representation of the invasion dynamics at the level of the subpopulations. The metapopulation system can be considered in a coarse-grained perspective as a network where each node represents a subpopulation which can be infected (i.e. *diseased*) if it is reached by the virus as carried by the infected individuals diffusing on the system.

Ball et al., 1997; Vázquez, 2006) where the nth generation of infected subpopulations of degree k is denoted by D_k^n.

In the early stage of the epidemics we assume that the number of subpopulations affected by an outbreak (with $R_0>1$) is small and we can therefore study the evolution of the number of diseased subpopulations by using a tree-like approximation relating D_k^n with D_k^{n-1}. Let us first analyze the case of a metapopulation system in the form of a homogeneous random graph in which each subpopulation has the same degree $k=\bar{k}$ and population $\bar{N}$. In this case we can drop the subscript index k (all subpopulations being equal) and obtain that

$$D^n = D^{n-1}(\bar{k}-1)\left[1-\left(\frac{1}{R_0}\right)^{\lambda_{\bar{k}\bar{k}}}\right]\left(1-\frac{D^{n-1}}{V}\right). \tag{17}$$

This equation assumes that each infected subpopulation of the $(n-1)$th generation, D^{n-1}, will seed with infected individuals a number of subpopulations during the course of the outbreak that depends on the product of the number of neighbor subpopulations minus the one which originally transmitted the disease, $\bar{k}-1$, times the probability that the subpopulation is not already seeded by infected individuals, $(1-D^{n-1})/V$, and the probability that the new seeded subpopulation will experience an outbreak, i.e. $(1-R_0^{-\lambda_{\bar{k}\bar{k}}})$ (Bailey, 1975). The last expression stems from the probability of extinction $P_{ext}=1/R_0$ given by a seed of a single infectious individual (Bailey, 1975). The simplest case of homogeneous diffusion of individuals $d_{\bar{k}}=p/\bar{k}$ yields $\lambda_{\bar{k}\bar{k}}=p\bar{N}\alpha\mu^{-1}/\bar{k}$. In order to obtain an explicit result we will consider in the following that $R_0-1\ll 1$, so that the system is poised very close to the epidemic threshold. In this limit we can approximate the outbreak probability as

$$\left[1-\left(\frac{1}{R_0}\right)^{\lambda_{\bar{k}\bar{k}}}\right] \simeq \lambda_{\bar{k}\bar{k}}(R_0-1), \tag{18}$$

and assuming that at the early stage of the epidemic $D^{n-1}/V\ll 1$ we obtain

$$D^n = p\bar{N}\alpha\mu^{-1}\frac{\bar{k}-1}{\bar{k}}(R_0-1)D^{n-1}. \tag{19}$$

This relation states that the number of subpopulations affected by an outbreak will increase only if the quantity

$$R_* = p\bar{N}\alpha\mu^{-1}\frac{\bar{k}-1}{\bar{k}}(R_0-1)>1. \tag{20}$$

This relation defines the *global invasion threshold*, i.e. the subpopulation reproductive number R_* that is the analogous of the basic reproductive number R_0 in structured metapopulation models. From the above expression it is possible to write the threshold condition on the mobility rate

$$p\bar{N} \geqslant \frac{\bar{k}}{\bar{k}-1}\frac{\mu}{\alpha}(R_0-1)^{-1} \tag{21}$$

that fixes the threshold in the diffusion rate of individuals for the global spread of the epidemic in the metapopulation systems. In other words, this equation states that there is a minimum rate for the traveling of individuals in order to ensure that on average each subpopulation can seed more than one neighboring subpopulations. As it has been pointed out by Cross et al. (2007) we find that other factors such as the infectious period and the mobility process are as much important as R_0 in the spread of epidemics in structured populations. The constant α is larger than 0 for any $R_0>1$, and in the SIR case for R_0 close to 1 it can be approximated by (Murray, 2005):

$$\alpha \simeq 2\frac{\mu}{\beta}\left(1-\frac{\mu}{\beta}\right)=\frac{2(R_0-1)}{R_0^2}, \tag{22}$$

yielding the mobility threshold for the SIR model

$$p\bar{N} \geqslant \frac{\bar{k}}{\bar{k}-1}\frac{\mu R_0^2}{2(R_0-1)^2}. \tag{23}$$

The above condition readily tells us that the closer to the epidemic threshold is the epidemic in the single subpopulation and the larger it has to be the traveling rate in order to sustain the global spread into the metapopulation model. We therefore find that we can define two different thresholds in a homogeneous metapopulation model. The first one is the local epidemic threshold $R_0>1$ within each subpopulation and the second one $R_*>1$ represents the global invasion threshold defining the traveling rate of individuals according to Eq. (23). It is important to stress that when R_0 increases the small R_0-1 expansions are no longer valid and the invasion threshold is obtained only in the form of a complicate implicit expression.

4.2. Global invasion threshold in metapopulation networks with traffic dependent mobility rates

The calculation for the global threshold becomes more complicated in the case of heterogeneous metapopulation networks. In this case Eq. (17) has to include also the degree and population heterogeneities yielding:

$$D_k^n = \sum_{k'} D_{k'}^{n-1}(k'-1)\lambda_{k'k}(R_0-1)P(k|k')\left(1-\frac{D_k^{n-1}}{V_k}\right). \tag{24}$$

This expression considers that each subpopulation of degree k' will seed the infection in a number $k'-1$ of subpopulations corresponding to the number of neighboring subpopulations minus the one which originally transmitted the infection, the probability $P(k|k')$ that each of the k' neighboring populations has degree k, and the probability to observe an outbreak in the seeded population, where as before we considered the limit $R_0-1\ll 1$ to obtain the outbreak probability as $\lambda_{k'k}(R_0-1)$. As in the previous case of the homogeneous network, we consider the early stage of the epidemic in which $(1-D_k^{n-1}/V_k)\simeq 1$. In addition we assume that degree correlations can be

neglected and $P(k|k') = kP(k)/\langle k\rangle$ obtaining

$$D_k^n = \frac{kP(k)}{\langle k\rangle}(R_0 - 1)\sum_{k'} D_{k'}^{n-1}(k'-1)\lambda_{k'k}. \quad (25)$$

The behavior of the above expression depends on the specific form of $\lambda_{k'k}$ that is determined by the diffusion rate $d_{kk'}$. Let us first consider the heterogeneous diffusion rate of Eq. (7) that gives

$$\lambda_{k'k} = \frac{p\langle k\rangle}{\langle k^{(1+\theta)}\rangle}\frac{\alpha}{\mu}\frac{(k)^\theta}{k'}N_{k'} = \frac{p\langle k\rangle}{\langle k^{(1+\theta)}\rangle^2}\frac{\alpha}{\mu}(kk')^\theta\bar{N}, \quad (26)$$

where in the last expression we have considered that the size of each population of degree k is given by the stationary diffusion process according to Eq. (10). The equation describing the generation of infected subpopulations is therefore read as

$$D_k^n = (R_0 - 1)\frac{k^{1+\theta}P(k)}{\langle k^{1+\theta}\rangle^2}\frac{p\bar{N}\alpha}{\mu}\sum_{k'} D_{k'}^{n-1}k'^\theta(k'-1). \quad (27)$$

By defining $\Theta^n = \sum_{k'} D_{k'}^n k'^\theta(k'-1)$ the last expression can be conveniently written in the iterative form

$$\Theta^n = (R_0 - 1)\frac{\langle k^{2+2\theta}\rangle - \langle k^{1+2\theta}\rangle}{\langle k^{1+\theta}\rangle^2}\frac{p\bar{N}\alpha}{\mu}\Theta^{n-1} \quad (28)$$

that allows the increasing of infected subpopulations and a global epidemic in the metapopulation process only if

$$R_* = (R_0 - 1)\frac{\langle k^{2+2\theta}\rangle - \langle k^{1+2\theta}\rangle}{\langle k^{1+\theta}\rangle^2}\frac{p\bar{N}\alpha}{\mu} > 1. \quad (29)$$

The subpopulation reproductive number is therefore an increasing function of the network heterogeneity that plays a role in the spread of the pathogen across subpopulations. In the case of an SIR epidemic within each subpopulation, the threshold on the mobility rate is provided by the expression

$$p\bar{N} \geqslant \frac{\langle k^{1+\theta}\rangle^2}{\langle k^{2+2\theta}\rangle - \langle k^{1+2\theta}\rangle}\frac{\mu R_0^2}{2(R_0-1)^2} \quad (30)$$

that differs from the homogeneous case for a correction factor depending on the topology of the network. Noticeably, the ratio $\langle k^{1+\theta}\rangle^2/(\langle k^{2+2\theta}\rangle) - \langle k^{1+2\theta}\rangle)$ is extremely small in heavy-tailed networks and it is vanishing in the limit of infinite network size. This implies that the heterogeneity of the metapopulation network is favoring the global spread of epidemics by lowering the global invasion threshold.

4.3. Global invasion threshold in metapopulation networks with population dependent mobility rates

As a final case let us consider the realistic framework in which the diffusion rate of individuals is proportional to the ratio between traveling people and population size, i.e. $p_k = T_k/N_k$. In Section 3.2 we have seen that this case corresponds to the mean-field assumption in metapopulation models coupled by traveling fluxes, leading to a stationary state in which the population N_k is stationary and independent on the diffusion process. Here $\lambda_{kk'} = w_0(kk')^\theta\alpha\mu^{-1}$ and by using the approximations considered in the previous cases the basic equation for the number of infected subpopulations reads as

$$D_k^n = (R_0 - 1)\frac{k^{1+\theta}P(k)}{\langle k\rangle}\frac{w_0\alpha}{\mu}\sum_{k'} D_{k'}^{n-1}k'^\theta(k'-1). \quad (31)$$

Also in this case by using the auxiliary function $\Theta^n = \sum_{k'} D_{k'}^n k'^\theta(k'-1)$ we obtain the recursive relation

$$\Theta^n = (R_0 - 1)\frac{\langle k^{2+2\theta}\rangle - \langle k^{1+2\theta}\rangle}{\langle k\rangle}\frac{w_0\alpha}{\mu}\Theta^{n-1}, \quad (32)$$

yielding for the global invasion the condition

$$R_* = (R_0 - 1)\frac{\langle k^{2+2\theta}\rangle - \langle k^{1+2\theta}\rangle}{\langle k\rangle}\frac{w_0\alpha}{\mu} > 1. \quad (33)$$

In the case of an SIR model for the intra-population disease we obtain

$$w_0 \geqslant \frac{\langle k\rangle}{\langle k^{2+2\theta}\rangle - \langle k^{1+2\theta}\rangle}\frac{\mu R_0^2}{2(R_0-1)^2}, \quad (34)$$

where also in this case the mobility threshold is lowered by the topological fluctuations of the network as the more heterogeneous is the metapopulation network and the smaller is the ratio $\langle k\rangle/(\langle k^{2+2\theta}\rangle - \langle k^{1+2\theta}\rangle)$. In this respect it is worth remarking that in principle in an infinite network with heavy tails the mobility threshold is vanishing as the ratio $\langle k\rangle/(\langle k^{2+2\theta}\rangle - \langle k^{1+2\theta}\rangle) \to 0$. In an infinite network, however, the above equations should be rewritten in terms of the density of diseased subpopulations in order to avoid the pathological divergence of some terms. Finally, it is interesting to notice that the effect of the network heterogeneity on the subpopulation reproductive number is similar to that of the contact pattern heterogeneity on the basic reproductive number (Anderson and May, 1992; Pastor-Satorras and Vespignani, 2001a; Lloyd and May, 2001; Barthélemy et al., 2005), stressing even more the close analogy between the two metrics.

4.4. Local and global threshold in real-world cases

It is worth stressing that the previous expressions are approximate and valid only in the limit in which a small fraction of the populations in the system is affected and in which $R_0 - 1 \ll 1$. It is, however, extremely relevant that metapopulations systems have intrinsically two epidemic thresholds. 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, 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 rate of diffusing/traveling individuals. The global invasion threshold condition $R_* > 1$ provides an estimate of the rate of diffusion of individuals above which the epidemic is able to affect a macroscopic fraction of the subpopulations defining the meta-population network.

At this point, it is useful to draw some gross estimate of the critical population coupling w_0 as a function of R_0 and of a realistic value of μ. If we assume a very mild reproductive rate of about $R_0 \simeq 1.1$ and a value $\mu = \frac{1}{3}$ per unit time (1 day) as in many estimates for influenza strains, we obtain that w_0 per unit time must be larger than approximately 20 individuals per day in the case of homogeneous networks and one order of magnitude or more smaller in heterogeneous networks. This is a value met in most of the modern real transportation systems. For example, the world-wide air transportation network analyzed in Barrat et al. (2004) and briefly described in Section 2.1 is characterized by a topology whose degree distribution moments which appear in the expressions of the global threshold are given by: $\langle k \rangle \simeq 10$, and $\langle k^{2+2\theta} \rangle - \langle k^{1+2\theta} \rangle \simeq 7 \times 10^4$, given that $\theta \simeq 0.5$ (Barrat et al., 2004). Therefore the condition expressed in Eq. (34) states that an epidemic carried by air travelers would reach global proportion if the average number of travelers per day is larger than approximately 3×10^{-3}, a constraint which is met in the airport network where the average daily traffic on a given connection has a minimum corresponding to $\simeq 10^{-2}$ (see footnote 1). The result of this estimate is in agreement with recent studies on contingency planning for a possible influenza pandemic, which show that travel restrictions, reducing the rate at which individuals leave an infected region, would not be able to considerably slow down the global spread unless $>90\%$ or more effective (Hollingsworth et al., 2006; Cooper et al., 2006; Colizza et al., 2007a). Our analytical results show that the invasion threshold is extremely small in realistic situationsand traffic reduction of more than one order of magnitude are in order to bring the system below the threshold.

5. Epidemic behavior above the invasion threshold

Above the invasion threshold we can assume that with finite probability the epidemics will affect a macroscopic fraction of subpopulations. In this limit, the stochastic effect due to the diffusion can be neglected and it is possible to study the epidemic spreading in the system by the deterministic equations obtained from a mechanistic approach to the metapopulation model where the disease dynamics in each subpopulation can be viewed as a reaction process (Colizza et al., 2007b). In the case of the SIR scheme the dynamics is identified by the following set of reaction equations:

$$I + S \to 2I, \tag{35}$$

$$I \to R. \tag{36}$$

In the SIS case the second reaction is just replaced by the reaction $I \to S$. From the rate equations it is clear that the dynamics conserves the total number of individuals. Before the diffusion process, the I_j and S_j individuals belonging to the same subpopulation j react according to the Eqs. (35) and (36). In each node j the spontaneous process $I \to R$ simply consists in turning each I_j individual into an R_j individual with rate μ. This process accounts for the recovery of infected individuals from the disease. The process $I + S \to 2I$ is related to the dependence of the transmissibility on the population density. In general, in large populations it is customary to consider that each individual has a finite number of contacts per unit time. In this case the probability that a susceptible has a contact with an infectious individual is equal to the density of infectious individuals within the subpopulation j, i.e. I_j/N_j. If we consider a homogeneous mixing assumption within the population, the creation rate of infectious individuals will be provided by $\beta\Gamma_j$, where Γ_j is an interaction kernel of the form

$$\Gamma_j = \frac{I_j S_j}{N_j}. \tag{37}$$

It is natural also to consider different dependencies of the transmissibility with respect to the density, giving rise to different reaction kernels (Anderson and May, 1992; Colizza et al., 2007b). We will provide an analysis of the case of reaction kernels simulating population with internal network structure and fully connected populations in a forthcoming paper (Colizza et al., in preparation).

5.1. Deterministic reaction–diffusion rate equations

In order to provide the explicit equations describing the dynamical evolution of the metapopulation system, we generalize the basic mechanistic approach with degree block variables used in the previous section to the complete reaction–diffusion process. We take into account the topological fluctuations of the coupling networks by introducing the quantities:

$$I_k = \frac{1}{V_k} \sum I_j, \quad S_k = \frac{1}{V_k} \sum S_j, \tag{38}$$

which represent the average number of I and S individuals in subpopulations with degree k. Analogously, the reaction kernel in the homogeneous assumption is written for subpopulation in each degree block as $\Gamma_k = I_k S_k / N_k$. Again it is worth remarking that the degree block variables assume the statistical equivalence of subpopulations with the same degree k. While this approximation is in fair agreement with empirical analysis, real-world subpopulations have differences that the present analysis does not take into account.

At the end of the reaction–diffusion process the variation in the number of infected individuals in subpopulations of degree block k can be written as a discrete master equations with the form

$$I_k(t+\Delta t) - I_k(t) = W_k^+ - W_k^-, \tag{39}$$

where the terms W_k^+ and W_k^- identify the number of infected individuals that entered or left the class I_k because of both the disease dynamics and the diffusion process. Assuming the general framework in which the diffusion rate out of a given subpopulation depends on its degree, p_k, the depletion term W_k^- can be simply evaluated as

$$W_k^- = p_k I_k + (1-p_k)\mu I_k. \tag{40}$$

The depletion term is just the sum of the I_k individuals that diffuse away of the subpopulation (first term of the rhs) and the infected individuals that do not diffuse away from the subpopulation but have a transition to the class R_k (second term of the rhs). The positive term W_k^+ takes into account both the new infected individuals generated by the disease dynamics within the subpopulation and the infected individuals that diffuse from the neighboring subpopulations with diffusion rate $d_{kk'}$, and it is given by

$$W_k^+ = (1-p_k)\beta\Gamma_k + k\sum_{k'} P(k'|k) d_{k'k}[(1-\mu)I_{k'} + \beta\Gamma_{k'}]. \tag{41}$$

The first term on the rhs considers the newly generated infected individuals within the subpopulation with degree k. The factor $1-p_k$ takes into account only those individuals that do not diffuse out of the subpopulation. The second term accounts for all the infected individuals arriving because of the diffusion process from neighboring subpopulations. The factor k considers the presence of k neighboring subpopulations, each one contributing a fraction $d_{k'k}$ of its number of diffusing infected individuals which is given by the non-recovering plus the newly generated ones $(1-\mu)I_{k'} + \beta\Gamma_{k'}$. Finally, on each connection edge we have to average over the probability that the neighboring subpopulation has degree k', that is, given by the weighted sum over the conditional probability $P(k'|k)$. As for the simple diffusion in Section 3, we consider an infinitesimal time interval $\Delta t \to 0$ and divide both terms of Eq. (39) to obtain the following set of differential equations

$$\begin{aligned}\partial_t I_k = & -p_k I_k + (1-p_k)[-\mu I_k + \beta\Gamma_k] \\ & + k\sum_{k'} P(k'|k) d_{k'k}[(1-\mu)I_{k'} + \beta\Gamma_{k'}],\end{aligned} \tag{42}$$

where all the parameter combinations are now infinitesimal transition rates. By considering the uncorrelated case $P(k'|k) = k'P(k')/\langle k\rangle$, we obtain the following dynamical reaction-rate equations

$$\begin{aligned}\partial_t I_k = & -p_k I_k + (1-p_k)[-\mu I_k + \beta\Gamma_k] \\ & + \frac{k}{\langle k\rangle}\sum_{k'} k'P(k') d_{k'k}[(1-\mu)I_{k'} + \beta\Gamma_{k'}].\end{aligned} \tag{43}$$

Similar expressions can be written also for the evolution of S_k and R_k.

5.2. The early stage of the epidemic outbreak

An explicit solution to the previous equations can be obtained for the early stages of the epidemic. In this case we can imagine to have a very small density of infectious individuals in the metapopulation system so that in general we can neglect contributions of order I_k^2. In this setting the reaction kernel Γ_k can be approximated as

$$\Gamma_k = \frac{(N_k - I_k - R_k)I_k}{N_k} \simeq I_k, \tag{44}$$

where we have neglected all terms of order I_k^2 and considered that R_k is of the same order of I_k in the early stage of the dynamics. The simplification of the reaction kernel allows to write Eq. (43) for the evolution of the number of infectious individuals in the following form

$$\begin{aligned}\partial_t I_k = & -p_k I_k + (1-p_k)(\beta-\mu)I_k \\ & + \frac{k}{\langle k\rangle}\sum_{k'} k'P(k') d_{k'k}[(1-\mu+\beta)I_{k'}].\end{aligned} \tag{45}$$

Explicit solutions to the above set of equations for the early dynamics of infectious individuals in degree block k can be found by considering the specific diffusion processes already introduced in the previous section.

5.3. Traffic dependent mobility rates

By considering a uniform p and the expression for $d_{kk'}$ of Eq. (7) we obtain after some simple algebra the following dynamical reaction-rate equations

$$\partial_t I_k = -pI_k + (1-p)(\beta-\mu)I_k + p\frac{k^{(1+\theta)}}{\langle k^{(1+\theta)}\rangle}[(1-\mu+\beta)\bar{I}] \tag{46}$$

that depend only on the number of infectious individuals, and where $\bar{I} = \sum_{k'} P(k')I_{k'}$. A solution for the average number of infectious individual in the metapopulation system is obtained by averaging both terms of the equation over $P(k)$, obtaining

$$\partial_t \sum_k P(k) I_k = \partial_t \bar{I} = (\beta-\mu)\bar{I}, \tag{47}$$

where we have considered that $\sum_k P(k)k^{(1+\theta)} = \langle k^{(1+\theta)}\rangle$. This equation has the simple solution

$$\bar{I} = \bar{I}(0)\mathrm{e}^{(\beta-\mu)t}, \tag{48}$$

where $\bar{I}(0)$ is the initial number of infected individuals in the metapopulation system. It readily states that the overall average number of infectious individuals in the system can grow only if $\beta > \mu$, thus recovering the epidemic threshold condition $R_0 = \beta/\mu > 1$. The metapopulation system exhibits at the deterministic level an epidemic threshold condition equivalent to that of each single population that

sets the time scale for the whole system. Intuitively this is stating that if the epidemic is not able to proliferate in each local subpopulation, then it cannot produce a major outbreak at the metapopulation level.

It is possible to solve the early time behavior for all subpopulations of a given degree block $I_k(t)$ by plugging in the explicit solution of $\bar{I}(t)$ in Eq. (46). This yields

$$I_k(t) = A\frac{k^{1+\theta}}{\langle k^{1+\theta}\rangle}e^{(\beta-\mu)t} + C_k e^{[(1-p)(\beta-\mu)-p]t}, \tag{49}$$

where A and C_k are parameters fixed by the initial conditions. If we assume that the metapopulation system is seeded with a total number of I_0 initially infected individuals distributed homogeneously among subpopulations, i.e. $I_k(0) = I_0/V = \bar{I}(0)\ \forall k$, we obtain

$$A = \bar{I}(0) \quad \text{and} \quad C_k = \bar{I}(0)\left(1 - \frac{k^{1+\theta}}{\langle k^{1+\theta}\rangle}\right). \tag{50}$$

If the I_0 infected are distributed only in the k_0-block subpopulations, $I_k(0) = \delta_{k,k_0}\bar{I}(0)/P(k_0)$, the coefficients A and C_k assume the following values:

$$A = \bar{I}(0) \quad \text{and} \quad C_k = \bar{I}(0)\left(\frac{\delta_{k,k0}}{P(k_0)} - \frac{k^{1+\theta}}{\langle k^{1+\theta}\rangle}\right). \tag{51}$$

The above results show how the choice of initial conditions, whether if homogeneously distributed or locally distributed, affects the early stages of the epidemic outbreak inside subpopulations of different block k. The change in the initial stage of the disease evolution in subpopulations depending on the degree block is confirmed by the numerical results reported in Section 6.

5.4. Population dependent diffusion rate

If we consider a population dependent diffusion rate, $p_k = T_k/N_k$, the system's behavior will be given by plugging into the set of Eq. (43) the degree dependent diffusion rate p_k and the expression of the rate of diffusion on a link $k' \to k$, $d_{k'k} = w_0(k'k)^\theta/N_{k'}$. In the approximation of early stage dynamics and considering the normalization condition $T_k = k^{1+\theta}w_0\langle k^{1+\theta}\rangle/\langle k\rangle$, we obtain

$$\begin{aligned}\partial_t I_k = & -p_k I_k + (1-p_k)(\beta-\mu)I_k \\ & + \frac{k^{1+\theta}}{\langle k^{1+\theta}\rangle}(1+\beta-\mu)\Omega,\end{aligned} \tag{52}$$

where $\Omega = \sum_k P(k)p_k I_k$. Proceeding along the line followed in the previous section, the solution can be found for the early stage behavior of the average number of infectious individuals $\bar{I} = \sum P(k)I_k$ by averaging both terms of the equation over $P(k)$:

$$\partial_t \bar{I} = (\beta-\mu)\bar{I}, \tag{53}$$

yielding

$$\bar{I} = \bar{I}(0)e^{(\beta-\mu)t} \tag{54}$$

and thus recovering the epidemic threshold condition $R_0 = \beta/\mu > 1$ also in this case. The early stage behavior of the epidemic in the deterministic approximation for the whole system does not differ from what observed in the previous heterogeneous frame. However, the dependence of p_k on the population N_k does not allow to use the same methods adopted in the previous subsection to solve Eq. (52) in the early time behavior for a given degree block $I_k(t)$. The explicit solution of $I_k(t)$ therefore, requires additional investigation and future development to provide important insight into the initial stages of the outbreak. It is worth to mention that the solution for the population dependent diffusion rate is obtained under very general conditions for the size of the subpopulations N_k which can assume any value, without being set by the dynamics as it happens for traffic dependent mobility rates.

It is clear from the previous analysis that the deterministic equations are not capable to account for the invasion threshold as they consider the diffusion and reaction processes in a mean-field perspective that provides deterministic equations for the average values. Indeed, the deterministic diffusion process allows any fraction pI_k of infected individual to deterministically seed new populations, washing out the stochastic effects responsible for the invasion threshold. While these equations do not allow to capture the invasion threshold, they provide a good description of the system and its behavior across degree classes above the invasion threshold, as we will show in the next section by comparing the analytical results with stochastic simulations at the mechanistic level.

6. Mechanistic numerical simulations

Here we provide extensive numerical simulations to support the theoretical picture described above. We report results from Monte Carlo simulations in a variety of different cases and compare them with the analytical findings. We adopt mechanistic numerical simulations where each single individual is tracked in time, during both the infection dynamics and the diffusion processes. The system evolves following a stochastic microscopic dynamics and at each time step it is possible to monitor quantities which depend on the subpopulation—such as, e.g. the number $I_j(t)$ of infectious individuals in the subpopulation j at time t—and also averages over blocks of nodes—e.g. the average number $I_k(t)$ of infected in subpopulations with degree k—or over the whole system, $\bar{I}(t)$. In addition, it is possible to study the evolution of the epidemic by monitoring the invasion dynamics at the local population level, and therefore measure the number of diseased subpopulations at time t, $D(t)$. Given the stochastic nature of the dynamics, the experiment can be repeated with different realizations of the noise, different underlying graphs and different initial conditions. This approach is equivalent to the real evolution of epidemic processes in the generated networks and can be used to

validate the theoretical results obtained in the analytical approach.

The substrate network is given by an uncorrelated complex network generated with the uncorrelated configuration model (Catanzaro et al., 2005), based on the *Molloy–Reed* algorithm (Molloy and Reed, 1995) with an additional constraint on the possible maximum value of the degree in order to avoid inherent structural correlations. The algorithm is defined as follows. Each node i is assigned a degree k_i obtained from a given degree sequence $P(k)$ subject to the restriction $k_i < V^{1/2}$. Here we assume a power-law degree distribution, $P(k) \sim k^{-\gamma}$, with $\gamma = 2.1$ and 3. Links are then drawn to randomly connect pairs of nodes, respecting their degree and avoiding self-loops and multiple edges. Sizes of $V = 10^4$ and 10^5 nodes have been considered. Weights on the connections among subpopulations are defined following the statistical law found in real transportation system (see Eq. (1)). Therefore, the weight on the link between subpopulation i and subpopulation j is given by

$$w_{ij} = w_0 (k_i k_j)^{\theta}, \tag{55}$$

where k_i and k_j are the degrees of the subpopulations i and j, respectively. Here we fix $w_0 = 1$, whereas θ assumes different values, including $\theta = 0$ for uniform weights. This expression for the weights is then used to define the diffusion rates in the cases analytically investigated, as in Eqs. (7) and (13).

The dynamics proceeds in parallel and considers discrete time steps representing the unitary time scale τ of the process. The reaction and diffusion rates are therefore converted into probabilities and at each time step; the system is updated according to the following rules. (a) Infection dynamics: (i) The contagion process assumes that in each subpopulation j individuals homogeneously mix and have a finite number of contacts, so that the probability for a susceptible to contract a virus from an infected is proportional to the transmission rate and normalized to the subpopulation size, β/N_j. At each time step in the simulation, each susceptible is turned into an infectious with probability $1-(1-(\beta/N_j)\tau)^{I_j}$. (ii) At the same time, each infectious individual is subject to the recovery process and becomes recovered with probability $\mu\tau$. (b) After all nodes have been updated for the reaction, we simulate the diffusion process. Results shown in the following subsections refer to the traffic dependent diffusion rate.

6.1. Global and local threshold in heterogeneous metapopulation models

In the previous sections we have shown that along with the usual local epidemic threshold $R_0 > 1$, the stochasticity and discreteness of the metapopulation diffusion process induce an intrinsic invasion threshold $R_* > 1$—at the global level—which rules the invasion dynamics in the coarse-grained view of the system. This threshold determines whether the coupling between subpopulation is high enough in order to allow the virus to spread from one subpopulation to another and invade a finite portion of the whole system. Here we numerically investigate this phenomenon by studying a metapopulation model with heterogeneous structure ($P(k) \simeq k^{-\gamma}$) and varying the coupling force between subpopulations. Initially, let us consider that the diffusion rate of an individual on the heterogeneous metapopulation structure is locally independent of the degree of the subpopulation—i.e. $p_k = p$—and heterogeneous on the links departing from a given subpopulation, following Eq. (7). The rate of diffusion from a subpopulation i to a subpopulation j for each individual in any given compartment located in i is therefore given by

$$d_{ij} = p \frac{w_0 (k_i k_j)^{\theta}}{T_i}, \tag{56}$$

where $T_i = \sum_j w_{ij}$ represents the traffic in i. The simulations proceed according to the following procedure: at each time step, after the update for the local infection dynamics within the subpopulations (see previous section), each individual in any compartment in subpopulation i moves to a neighboring subpopulation j with rate d_{ij}. We analyze different values of R_0 by assuming $\mu = 0.2$ and 0.02, and varying the value of the transmission rate β. The simulations start with a localized initial condition given by the seeding of a subpopulation having degree k_0 with $I_0 = 10$ infected individuals. This allows to monitor the epidemic evolution in the metapopulation model and measure the final size of the epidemic, expressed in terms of the number or density of cases obtained in the whole system and the number of subpopulations experiencing an outbreak.

In Fig. 4 we analyze the behavior of $\bar{I}(t)$ and $D(t)$ for $R_0 = 3$, above the local threshold, and for two values of the diffusion rate $p = 0.5$ and 10^{-5} that poise the system below and above the invasion threshold, respectively. While in both cases the figure shows an increase of the value of $\bar{I}(t)$, the behavior of $D(t)$ is very different above and below the invasion threshold. Indeed, above the invasion threshold the number of affected subpopulations is increasing exponentially, while below the threshold the number of subpopulations remains small and goes to 0 in a finite time. The increase in $\bar{I}(t)$ instead is guaranteed also below the invasion threshold by the outbreak in the initial seeded population. On a longer time, however, $\bar{I}(t)$ keeps increasing only if the system is above the invasion threshold and new subpopulations are progressively infected.

While Fig. 4 provides a clear evidence of the two separate threshold mechanisms, a complete analysis of the system phase diagram is obtained by analyzing the behavior of the global metapopulation attack rate, defined as the total fraction of cases $R(\infty)/N$ at the end of the epidemic, as a function of both R_0 and p. In Fig. 5, we report the global attack rate surface in the $p-R_0$ space, and the 2D plots of the p and R_0 crosscuts. Fig. 5 clearly shows the effect of different couplings as expressed by the value of

p in reducing the final size of the epidemic at a given fixed value of R_0. The smaller the value of R_0, the higher the coupling needs to be in order for the virus to successfully invade a finite fraction of the subpopulations, in agreement with the analytic result of Eq. (30). This provides a clear illustration of the varying global invasion threshold as a function of the reproductive rate R_0. On the contrary, p-crosscuts show that whatever the value of p, $R_0 < 1$ does not allow the epidemic to spread.

Finally, it is possible to study the effect of the heterogeneity of the metapopulation structure. Fig. 6 shows the results obtained comparing a heterogeneous network characterized by a scale-free degree distribution $P(k) \sim k^{-2.1}$ with a homogeneous network having the same size $V = 10^5$ and same average degree. The presence of topological fluctuations leads to a smaller ratio $\langle k^{1+\theta} \rangle^2 / (\langle k^{2+2\theta} \rangle - \langle k^{1+2\theta} \rangle)$, thus lowering the value of the mobility threshold with respect to the homogeneous network.

6.2. Epidemics above the invasion threshold

Above the global invasion threshold $R_* > 1$, the epidemic process is guaranteed to invade a macroscopic fraction of

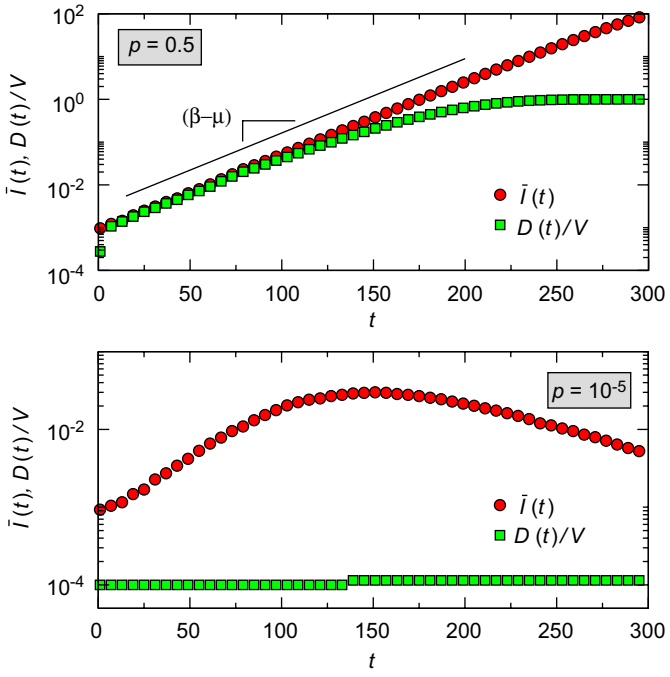

Fig. 4. Metapopulation system's behavior above and below the global threshold. Results refer to $R_0 = 3$, $\bar{N} = 10^3$ and $\theta = 0.5$. The epidemic is in both cases above the local threshold, leading to an exponential increase of $\bar{I}(t)$. Differences in the diffusion rate values (top: $p = 0.5$/unit time, bottom: $p = 10^{-5}$/unit time) show the effect of the global threshold on the number $D(t)$ of diseased subpopulations. $D(t)$ is normalized to the system size $V = 10^4$ for sake of visualization.

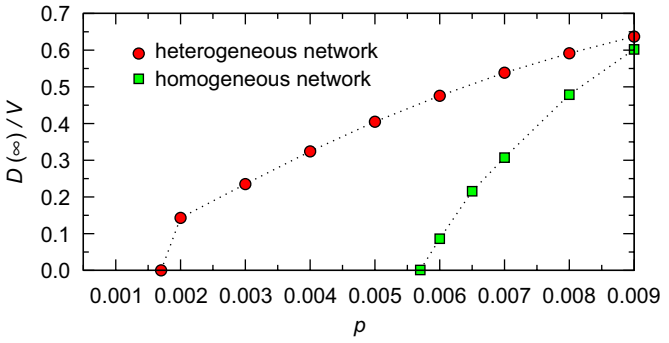

Fig. 6. Effect of metapopulation structure heterogeneity on the global epidemic threshold. The final fraction of diseased subpopulations $D(\infty)/V$ at the end of the global epidemic is shown as a function of the traveling diffusion rate p. A heterogeneous network with heavy-tailed degree distribution, $P(k) \sim k^{-2.1}$ is compared to a homogeneous network with poissonian $P(k)$ having the same size $V = 10^5$ and same average degree. Here $\theta = 0$.

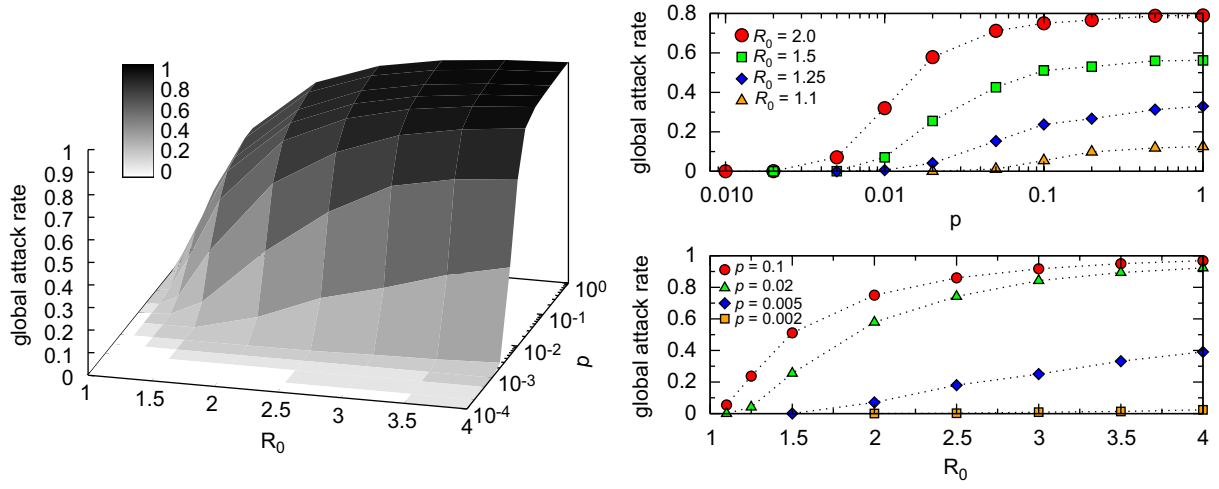

Fig. 5. Global threshold in a heterogeneous metapopulation system with traffic dependent diffusion rates. The left panel shows a 3D surface representing the value of the final epidemic size in the metapopulation system as a function of the local threshold R_0 and of the diffusion rate p. If R_0 approaches the threshold, larger values of the diffusion rate p need to be considered in order to observe a global outbreak in the metapopulation system. On the right, two plots showing the cross-sections of the 3D plot at fixed values of R_0 (top) and at fixed values of the traveling rate p (bottom).

subpopulations and it is possible to inspect the validity of the results obtained in Section 5 with the deterministic reaction–diffusion equations. A general conclusion is that the average number of infectious individuals in the system in the early stage of the epidemic dynamics grows as

$$\bar{I}(t) = \bar{I}(0)e^{(\beta-\mu)t}, \tag{57}$$

if the threshold condition, $R_0 = \beta/\mu > 1$ is satisfied. The early time behavior expressed in the above equation is also independent of the parameters related to the diffusion process among subpopulations, such as the homogeneous diffusion rate p and the exponent θ which governs the relation between weights and subpopulation degrees. The analytic result of Eq. (57) is confirmed in Fig. 7, where we show simulation results of the metapopulation epidemic model with traffic dependent diffusion rates. We consider systems of $V = 10^4$ subpopulations each of initial size $\bar{N} = 10^4$, connected through a heterogeneous network having degree distribution $P(k) \simeq k^{-3}$. The disease parameters assume the following values: $\beta = 0.04$ and $\mu = 0.02$, yielding $R_0 > 1$. The simulations are seeded with $\bar{I}(0) = 100$ infectious individuals, homogeneously distributed among subpopulations. Both homogeneous ($\theta = 0$) and heterogeneous ($\theta = 0.5$) diffusions are considered, as well as different values of the diffusion rate $p = 0.5, 0.75, 1.0$ per unit time.

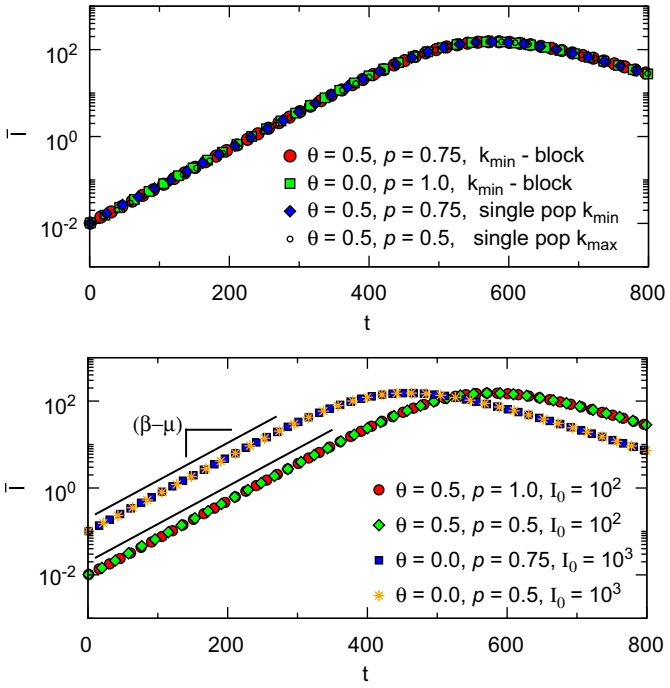

Fig. 7. Evolution in time of the average number of infectious individuals in a heterogeneous metapopulation system with traffic dependent diffusion rates. Top: the effects of the heterogeneity of diffusion (θ), of the rate of diffusion (p), and of the distribution of initially infected individuals in the system (homogeneously distributed in a k-degree block or concentrated in a single subpopulation) are compared and found to produce the same early stage behavior. Bottom: changes in the initial condition value $\bar{I}$ produce the same exponential increase in the metapopulation system behavior.

Results in Fig. 7 show that the early behavior of the average number of infectious individuals is independent of the values of θ and p, and of the location of the initial seed, whether if homogeneously distributed among the subpopulations of a given degree block or localized in a single subpopulation. All simulations show an exponential increase which confirms the analytic findings.

Numerical simulations also allow for the study of the dynamic behavior of infectious individuals in subpopulations of degree block k. The solutions obtained in Section 5.3 show a dependence of the early time behavior on the degree k of the subpopulation, pointing to a dynamics which switches on degree modes at different times. Results of the numerical simulations confirm this findings, as shown in Fig. 8. Here the disease parameters assume the same values as before, and the diffusion is governed by the values $\theta = 0.5$ and $p = 0.75$ for the numerical results reported in the top panel, whereas the effect of different values of θ is reported in the bottom panel. In order to see the effects of different initial conditions on the dynamical behavior of degree block k subpopulations (see Eqs. (50) and (51)), we seed the epidemics with (i) 10^2 infectious individuals homogeneously distributed among subpopulations, or with (ii) 10^2 infectious individuals localized in subpopulations of degree block k_0 (results in the top panel correspond to $k_0 = k_{max}$). While the global behavior $\bar{I}(t)$ is not affected by the choice of the initial conditions (see previous figure), the subpopulations experience outbreaks

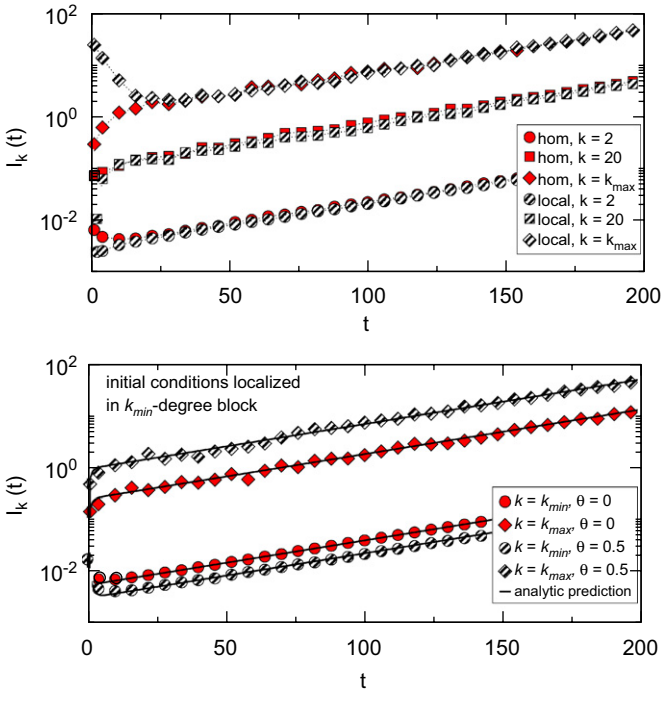

Fig. 8. Evolution in time of the number of infectious individuals in degree block k subpopulations. Top: changes in the initial conditions (homogeneous *vs.* localized in k_{max}) yield different behaviors in the early time dynamics of distinct degree blocks. Bottom: changes in the value of θ impact differently the time evolution of the average number of infected I_k in each degree block, whether if $k = k_0$ (i.e. k_{min}) or $k \neq k_0$ (i.e. $k = k_{max}$).

at different times, as brought and delayed by the diffusion dynamics. The system heterogeneity, as contained in the factor $k^{1+\theta}/\langle k^{1+\theta}\rangle$ of the explicit solution of $I_k(t)$, differentiates the evolution of the degree block subpopulations at short times. Numerical results obtained for the study of the effect of traffic heterogeneity (Fig. 8, bottom) are compared with the analytical findings of Section 5.3.

7. Conclusions and outlook

Here we have introduced an analytic framework in terms of degree block variables which allows to gain insights into the behavior of mechanistic metapopulation epidemic models which explicitly include demographic and mobility heterogeneities. The system is shown to display a local epidemic threshold which depends on the disease parameter values only and is responsible for the epidemic outbreak at the local scale, and a global epidemic threshold which determines the invasion dynamics of the subpopulations and depends critically on the disease parameters and the diffusion rates of the individuals. Changes in coupling between subpopulations are shown to have critical implications for disease extinction.

The results provide useful insights for the basic theoretical understanding of mechanistic epidemic models in complex environments, which can then be used to build more realistic data-driven large-scale computational approaches for real case scenarios and spatially targeted control measures. However, several key theoretical and practical issues are still to be addressed. Data on human dynamics at the local level, i.e. within any subpopulation, could push forward a more sophisticate theoretical framework for the local infection dynamics, to go beyond the homogeneous mixing assumption (Meyers et al., 2005; Lloyd-Smith et al., 2005). The behavior of metapopulation models characterized by complex internal structure in each patch is a major question for the theoretical epidemiology of the future. In addition, more realistic and detailed diffusion patterns should be included in order to better model the coupling terms by including non-Markov processes and introducing elements of memory in the system. Obviously this corresponds to the need for more accurate data on population behavior, such as fraction of commuters, probability of short/medium/long range travel, trip duration, and so on (Riley, 2007). Additional levels of heterogeneity can be also included in the diffusive patterns by introducing a dependence of the probability of diffusion on the stage of the disease. In many real cases, e.g. the severity of symptoms or hospitalization measures would prevent the diffusion out of a patch to a portion of the population. The impact of heterogeneities in traveling pattern of individuals depending on their infection state could provide additional insights fundamental to the study of global extinction and eradication. All these improvements and future directions would help filling the gap between the evidence from increasingly realistic epidemic models and their theoretical understanding.

Acknowledgments

We are grateful to Alain Barrat, Marc Barthelemy and Romualdo Pastor-Satorras for useful discussions during all stages of preparation of this work. A.V. is partially funded by the NSF award IIS-0513650. V.C and A.V. are partially funded by the CRT foundation through the Lagrange Project.

References

Albert, R., Barabási, A.-L., 2002. Statistical mechanics of complex networks. Rev. Mod. Phys. 74, 47–97.

Anderson, R.M., May, R.M., 1984. Spatial, temporal and genetic heterogeneity in host populations and the design of immunization programs. IMA J. Math. Appl. Med. Biol. 1, 233–266.

Anderson, R.M., May, R.M., 1992. Infectious Diseases in Humans. Oxford University Press, Oxford.

Bailey, N.T., 1975. The Mathematical Theory of Infectious Diseases., second ed., Hodder Arnold.

Ball, F., Mollison, D., Scalia-Tomba, G., 1997. Epidemics with two levels of mixing. Ann. Appl. Probab. 7, 46–89.

Baroyan, O.V., Genchikov, L.A., Rvachev, L.A., Shashkov, V.A., 1969. An attempt at large-scale influenza epidemic modelling by means of a computer. Bull. Int. Epidemiol. Assoc. 18, 22–31.

Barrat, A., Barthélemy, M., Pastor-Satorras, R., Vespignani, A., 2004. The architecture of complex weighted networks. Proc. Natl Acad. Sci. USA 101, 3747–3752.

Barrett, C.L., et al., 2000. TRANSIMS: Transportation Analysis Simulation System. Technical Report LA-UR-00-1725, Los Alamos National Laboratory.

Barthélemy, M., Barrat, A., Pastor-Satorras, R., Vespignani, A., 2005. Dynamical patterns of epidemic outbreaks in complex heterogeneous networks. J. Theor. Biol. 235, 275–288.

Bascompte, J., Solé, R.V., 1998. Modeling Spatiotemporal Dynamics in Ecology. Springer, New York.

Bolker, B.M., Grenfell, T., 1993. Chaos and biological complexity in measles dynamics. Proc. R. Soc. London B 251, 75–81.

Bolker, B.M., Grenfell, T., 1995. Space persistence and dynamics of measles epidemics. Philos. Trans. R. Soc. London B 348, 309–320.

Catanzaro, M., Boguña, M., Pastor-Satorras, R., 2005. Generation of uncorrelated random scale-free networks. Phys. Rev. E 71, 027103.

Chowell, G., Hyman, J.M., Eubank, S., Castillo-Chavez, C., 2003. Scaling laws for the movement of people between locations in a large city. Phys. Rev. E 68, 066102.

Colizza, V., Barrat, A., Barthélemy, M., Vespignani, A., 2006a. The role of the airline transportation network in the prediction and predictability of global epidemics. Proc. Natl Acad. Sci. USA 103, 2015–2020.

Colizza, V., Barrat, A., Barthélemy, M., Vespignani, A., 2006b. The modeling of global epidemics: stochastic dynamics and predictability. Bull. Math. Biol. 68, 1893–1921.

Colizza, V., Barrat, A., Barthélemy, M., Valleron, A.-J., Vespignani, A., 2007a. Modeling the worldwide spread of pandemic influenza: baseline case and containment interventions. PLoS Med. 4 (1), e13.

Colizza, V., Pastor-Satorras, R., Vespignani, A., 2007b. Reaction–diffusion processes and metapopulation models in heterogeneous networks. Nat. Phys. 3, 276–282.

Colizza, V., Pastor-Satorras, R., Vespignani, A., in preparation.

Cooper, B.S., Pitman, R.J., Edmunds, W.J., Gay, N.J., 2006. Delaying the international spread of pandemic influenza. PLoS Med. 3, e12.

Cross, P., Lloyd-Smith, J.O., Johnson, P.L.F., Wayne, M.G., 2005. Duelling timescales of host movement and disease recovery determine invasion of disease in structured populations. Ecol. Lett. 8, 587–595.

Cross, P., Johnson, P.L.F., Lloyd-Smith, J.O., Wayne, M.G., 2007. Utility of R_0 as a predictor of disease invasion in structured populations. J. R. Soc. Interface 4, 315–324.

De Montis, A., Barthélemy, M., Chessa, A., Vespignani, A. 2007. The structure of interurban traffic: a weighted network analysis. Environment Plann. B (doi:10.1068/b32128).

Dorogovtsev, S.N., Mendes, J.F.F., 2003. Evolution of networks: From biological nets to the Internet WWW. Oxford University Press, Oxford.

Earn, D.J.D., Rohani, P., Grenfell, B.T., 1998. Persistence chaos and synchrony in ecology and epidemiology. Proc. R. Soc. London B 265, 7–10.

Ferguson, N.M., Keeling, M.J., Edmunds, W.J., Gani, R., Grenfell, B.T., Anderson, R.M., 2003. Planning for smallpox outbreaks. Nature 425, 681–685.

Flahault, A., Valleron, A.-J., 1991. A method for assessing the global spread of HIV-1 infection based on air-travel. Math. Popul. Stud. 3, 1–11.

Grais, R.F., Ellis, J.H., Glass, G.E., 2003. Assessing the impact of airline travel on the geographic spread of pandemic influenza. Eur. J. Epidemiol. 18, 1065–1072.

Grais, R.F., Ellis, J.H., Kress, A., Glass, G.E., 2004. Modeling the spread of annual influenza epidemics in the US: the potential role of air travel. Health Care Manage. Sci. 7, 127–134.

Grenfell, B.T., Bolker, B.M., 1998. Cities and villages: infection hierarchies in a measles metapopulation. Ecol Lett. 1, 63–70.

Grenfell, B.T., Harwood, J., 1997. (Meta)population dynamics of infectious diseases. Tree 12, 395–399.

Guimerá, R., Mossa, S., Turtschi, A., Amaral, L.A.N., 2005. The worldwide air transportation network: anomalous centrality, community structure, and cities' global roles. Proc. Natl Acad. Sci. USA 102, 7794–7799.

Hanski, I., Gilpin, M.E., 1997. Metapopulation Biology: Ecology, Genetics, and Evolution. Academic Press, San Diego.

Hanski, I., Gaggiotti, O.E., 2004. Ecology Genetics and Evolution of Metapopulations. Elsevier, Academic Press, Amsterdam, New York.

Harris, T.E., 1989. The theory of branching processes. Dover Publications.

Hethcote, H.W., 1978. An immunization model for a heterogeneous population. Theor. Popul. Biol. 14, 338–349.

Hollingsworth, T.D., Ferguson, N.M., Anderson, R.M., 2006. Will travel restrictions control the international spread of pandemic influenza? Nat. Med 12, 497–499.

Hufnagel, L., Brockmann, D., Geisel, T., 2004. Forecast and control of epidemics in a globalized world. Proc. Natl Acad. Sci. USA 101, 15124–15129.

Keeling, M.J., 2000. Metapopulation moments: coupling, stochasticity and persistence. J. Anim. Ecol. 69, 725–736.

Keeling, M.J., Rohani, P., 2002. Estinating spatial coupling in epidemiological systems: a mechanistic approach. Ecol. Lett. 5, 20–29.

Liljeros, F., Edling, C.R., Amaral, L.A.N., Stanley, H.E., Aberg, Y., 2001. The web of human sexual contacts. Nature 411, 907–908.

Levins, R., 1969. Some demographic and genetic consequences of environmental heterogeneity for biological control. Bull. Entomol. Soc. Am. 15, 237–240.

Levins, R., 1970. Extinction. Lecture Notes in Mathematics 2, 75–107.

Lloyd, A.L., May, R.M., 1996. Spatial heterogeneity in epidemic models. J. Theor. Biol. 179, 1–11.

Lloyd, A.L., May, R.M., 2001. How viruses spread among computers and people. Science 292, 1316–1317.

Lloyd-Smith, J.O., Schreiber, S.J., Kopp, P.E., Getz, W.M., 2005. Superspreading and the effect of individual variation on disease emergence. Nature 438, 355–359.

Longini, I.M., 1988. A mathematical model for predicting the geographic spread of new infectious agents. Math. Biosci. 90, 367–383.

Marro, J., Dickman, R., 1999. Nonequilibrium Phase Transitions in Lattice Models. Cambridge University Press, Cambridge.

May, R.M., Anderson, R.M., 1979. Population biology of infectious diseases part II. Nature 280, 455–461.

May, R.M., Anderson, R.M., 1984. Spatial heterogeneity and the design of immunization programs. Math Biosci. 72, 83–111.

Meyers, L.A., Pourbohloul, B., Newman, M.E.J., Skowronski, D.M., Brunham, R.C., 2005. Network theory and SARS: predicting outbreak diversity. J. Theor. Biol. 232, 71–81.

Molloy, M., Reed, B., 1995. A critical point for random graphs with a given degree sequence. Random Struct. Algorithms 6, 161–179.

Moreno, Y., Pastor-Satorras, R., Vespignani, A., 2002. Epidemic outbreaks in complex heterogeneous networks. Eur. Phys. J. B 26, 521–529.

Murray, J.D., 2005. Mathematical Biology. third ed., Springer, Berlin.

Newman, M.E.J., 2003. The structure and function of complex networks. SIAM Rev. 45, 167.

Park, A.W., Gubbins, S., Gilligan, C.A., 2002. Extinction times for closed epidemics: the effects of host spatial structure. Ecol. Lett. 5, 747–755.

Pastor-Satorras, R., Vespignani, A., 2001a. Epidemic spreading in scale-free networks. Phys. Rev. Lett. 86, 3200–3203.

Pastor-Satorras, R., Vespignani, A., 2001b. Epidemic dynamics and endemic states in complex networks. Phys. Rev. E 63, 066117.

Pastor-Satorras, R., Vespignani, A., 2004. Evolution and Structure of the Internet: A Statistical Physics Approach. Cambridge University Press, Cambridge.

Pastor-Satorras, R., Vázquez, A., Vespignani, A., 2001. Dynamical and correlation properties of the internet. Phys. Rev. Lett. 87, 258701.

Riley, S., 2007. Large-scale transmission models of ingectious disease. Science 316, 1298–1301.

Rohani, P., Earn, D.J.D., Grenfell, B.T., 1999. Opposite patterns of synchrony in sympatric disease metapopulations. Science 286, 968–971.

Ruan, S., Wang, W., Levin, S., 2006. The effect of global travel on the spread of SARS. Math. Biosci. Eng. 3, 205–218.

Rvachev, L.A., Longini, I.M., 1985. A mathematical model for the global spread of influenza. Math. Biosci. 75, 3–22.

Sattenspiel, L., Dietz, K., 1995. A structured epidemic model incorporating geographic mobility among regions. Math. Biosci. 128, 71–91.

Schneeberger, A., Mercer, C.H., Gregson, S.A.J., Ferguson, N.M., Nyamukapa, C.A., Anderson, R.M., Johnson, A.M., Garnett, G.P., 2004. Scale-free networks and sexually transmitted diseases. Sexually Transmitted Dis. 31, 380–387.

Tilman, D., Kareiva, P., 1997. Spatial Ecology. Princeton University Press, Princeton.

van Kampen, N.G., 1981. Stochastic Processes in Chemistry and Physics. North-Holland, Amstredam, 1981.

Vázquez, A., 2006. Polynomial growth in age-dependent branching processes with diverging reproductive number. Phys. Rev. Lett. 96, 038702.

Vázquez, A., 2007. Epidemic outbreaks on structured populations. J. Theor. Biol. 245, 125–129.

Viboud, C., Bjørnstad, O.N., Smith, D.L., Simonsen, L., Miller, M.A., Grenfell, B.T., 2006. Synchrony, waves, and spatial hierarchies in the spread of influenza. Science 312, 447–451.

Watts, D., Muhamad, R., Medina, D.C., Dodds, P.S., 2005. Multiscale resurgent epidemics in a hierarchical metapopulation model. Proc. Natl Acad. Sci. USA 102, 11157–11162.