

## Critical behavior in interdependent spatial spreading processes with distinct characteristic time scales

Piergiorgio Castioni <sup>1,2</sup>, Riccardo Gallotti <sup>2</sup> & Manlio De Domenico<sup>2</sup>

The spread of an infectious disease is well approximated by metapopulation networks connected by human mobility flow and upon which an epidemiological model is defined. In order to account for travel restrictions or cancellation we introduce a model with a parameter that explicitly indicates the ratio between the time scales of the intervening processes. We study the critical properties of the epidemic process and its dependence on such a parameter. We find that the critical threshold separating the absorbing state from the active state depends on the scale parameter and exhibits a critical behavior itself: a metacritical point – a critical value in the curve of critical points – reflected in the behavior of the attack rate measured for a wide range of empirical metapopulation systems. Our results have potential policy implications, since they establish a non-trivial critical behavior between temporal scales of reaction (epidemic spread) and diffusion (human mobility) processes.

<sup>1</sup> Department of Physics, University of Trento, Povo, TN, Italy. <sup>2</sup> CoMuNe Lab, Fondazione Bruno Kessler, Povo, TN, Italy. email: [piergiorgio.castioni@alumni.unitn.it](mailto:piergiorgio.castioni@alumni.unitn.it)

The spatio-temporal spread of infectious diseases is inherently interdependent with human behavior<sup>1–4</sup>. On the one hand, individuals move between spatial patches, effectively creating human flows through which pathogens travel from one place to another. On the other hand, social dynamics can favor or hinder potential transmission channels: a typical example for the latter relies on risk perception and spread of awareness<sup>3,5–7</sup>, determining the choice to take individual actions—such as social distancing and wearing masks—or mid-scale non-pharmaceutical interventions—such as school or workplace closure and household quarantine—and have the direct effect of reducing the probability that a pathogen is transmitted from one infectious individual to susceptible ones in their social neighborhood. Convincing evidence for the role of human behavior in the spreading of several infectious diseases has been reported in the last decades for Foot-and-mouth disease<sup>8</sup>, Influenza and its strains<sup>9–13</sup>, SARS<sup>14,15</sup>, Ebola<sup>16</sup>, Dengue<sup>17,18</sup>, Zika<sup>19</sup> and, very recently, for COVID-19<sup>20–24</sup>.

Consequently, a huge effort has been devoted to understand the fundamental dynamical mechanisms where it is possible to take action through policy in order to drive a system subjected to an ongoing epidemic toward a controlled state, rather than an uncontrolled one. A standard approach is to coarse-grain the system under investigation in terms of spatial patches, named metapopulations, which account for groups of individuals localized at a given geographic scale. Metapopulations can be thought as nodes of a complex network of spatial patches, where links encode human flows from one place to another and are responsible for between-patch transmission<sup>25</sup>. In fact, a metapopulation can be one neighborhood within a city<sup>26,27</sup>, a city<sup>28,29</sup>, a sub-national region or higher scales<sup>30–34</sup>. Models for human mobility can range from Markov chains encoding standard diffusion<sup>35</sup> to more sophisticated dynamics encoding higher-order memory<sup>36,37</sup>: they are often encoded into origin-destination (OD) matrices which are later encoded into the essential diffusive part in the dynamical equations that describe the spatio-temporal evolution of epidemics in terms of reaction-diffusion processes.

However, quite often those models assume that an epidemic unfolds slower than the average time people need to move around<sup>38</sup>. This assumption, known as separation of scale approximation, is verified in practice, since the incubation period and the average duration of many diseases is of the order of days, which is much longer than the typical time scale of human movements, ranging from about 30 min for urban trips to 8 h for interurban ones<sup>39,40</sup>. Nevertheless, there are specific scenarios which can alter the typical time scales of the intervening dynamical processes, e.g., non-pharmaceutical interventions employed for disease containment can dramatically change the time scales of social dynamics. Specifically, changes in human mobility can be due to extraordinary travel restrictions, cancellations or to the effects of other restrictions such as school and workplace closure or delayed access to points of interest, from restaurants and gyms to shopping malls and museums, as happened during the COVID-19 pandemic<sup>41–44</sup>.

In this work, we propose a model able to cope with this scenario by explicitly accounting for the existence of two distinct time scales, each one characterizing the intervening reaction and diffusion processes. The time-scale separation is therefore encoded into a parameter  $\epsilon$ , which is defined in terms of the ratio between the diffusion and reaction time scales. This parameter allows us to tune the behavior of a system between two extremal scenarios—such as reaction faster than diffusion ( $\epsilon > 1$ , Fig. 1a) and diffusion faster than reaction ( $\epsilon < 1$ , Fig. 1b)—and all the intermediate ones. For simplicity, in the following we will refer to this model as the  $\epsilon$ -model. To allow for an analytical treatment of the

problem, we focus our attention on the emblematic Susceptible-Infected-Recovered (SIR) epidemic model for a metapopulation network, where particles flow at each time step according to transition rules encoded by empirical human flows. Using mean-field approximations and perturbative analysis, we find an analytic expression for the basic reproduction number  $R_0(\epsilon)$ , which indicates if the epidemics will reach a finite fraction of the population ( $R_0(\epsilon) > 1$ ) or to die out ( $R_0(\epsilon) < 1$ ) in the stationary state, effectively separating the active state from the absorbing state, respectively. We then find that the co-evolution of reaction and diffusion on different time scales reveals the existence of a phase transition characterizing the behavior of the critical threshold  $R_0(\epsilon)$  as a function of the parameter  $\epsilon$ : the value  $\epsilon^*$  separates the regime where  $R_0(\epsilon)$  is constant ( $\epsilon < \epsilon^*$ ) from the regime where it grows monotonically ( $\epsilon > \epsilon^*$ ). Note that for this behavior to be observed it is required that transmission rates are distributed among nodes in a nonuniform way. We identify a quantity  $R_0^{\min}$  that parametrizes such nonuniform distributions so that the point  $(R_0^{\min}, \epsilon) = (1, \epsilon^*)$  is a metacritical point, which separates all epidemic outbreaks between certain or possible based on the separation of time scales, the infectiousness distribution and the topology of the network.

## Results and discussion

**Modeling of epidemics through multiscale reaction-diffusion processes.** Let us consider the scenario of  $M$  geographical patches where individuals move from patch  $i$  to patch  $j$  ( $i, j = 1, 2, \dots, M$ ) with probability rate  $\mathcal{P}_{ij}$ , encoding the entries of the corresponding mobility matrix  $\mathcal{P}$ . Let us also consider a standard SIR model for the spreading of an infectious disease, subjected to the conditions  $S_i(t) + I_i(t) + R_i(t) = N_i(t)$ , where  $S_i(t)$ ,  $I_i(t)$ , and  $R_i(t)$  are the number of susceptible, infected and recovered agents in the  $i$ -th patch at time  $t$ , while  $N_i(t)$  is the overall population of the  $i$ -th patch at time  $t$ , with  $\sum_i N_i = N$  the total population, assumed to be constant over time. The variation of the number of infected individuals in each patch due to the epidemic spreading is given by

$$\Delta I_i^{\text{epi}}(t) = \left[ \beta_i S_i(t) \frac{I_i(t)}{N_i(t)} - \mu I_i(t) \right] \Delta t_{\text{epi}} \quad (1)$$

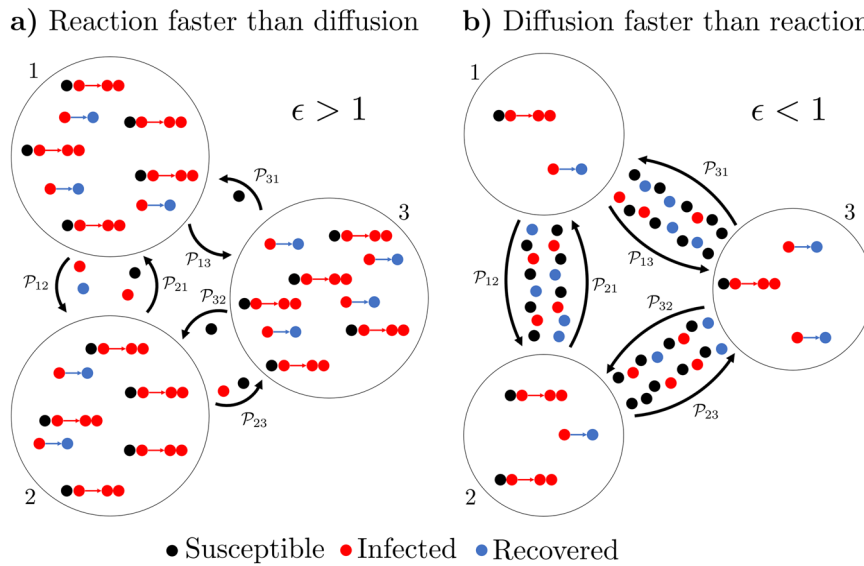
where  $\beta_i$  encodes the transmission rate in patch  $i$  for a communicable disease with empirical recovery rate given by  $\mu$ . The possibility of having patch-dependent transmission rates is plausible for a metapopulation network where each spatial patch is characterized by a population with socioeconomic or clinical differences, such as average age or income. During the COVID-19 pandemic, for instance, such differences were reflected in the variance of infection fatality rates<sup>45</sup>.

If the population is allowed to move among patches according to the mobility matrix, the variation due to diffusion is given by

$$\Delta I_i^{\text{mob}}(t) = \left[ \sum_{j=1}^M \mathcal{P}_{ji} I_j(t) - \sum_{j=1}^M \mathcal{P}_{ij} I_i(t) \right] \Delta t_{\text{mob}} \quad (2)$$

for infected individuals. Since we are working under the assumption that the epidemiological state of an agent does not affect their diffusive behavior, the equations for susceptible and recovered individuals have the exact same form. The time scales  $\Delta t_{\text{epi}}$  and  $\Delta t_{\text{mob}}$  should be interpreted as the minimum time necessary to have a certain small variation in the number of infected individuals, because of the epidemic process or the mobility one, respectively.

Usually, those two distinct dynamics happen at the same time but, in principle, the corresponding time scales of epidemics spreading and mobility dynamics are not necessarily the same.



**Fig. 1 Schematic illustration of reaction-diffusion processes characterized by different time scales as captured by our model.** The agents are represented with dots whose colors indicate their state with respect to the disease. The big circles represent different spatial patches and the probability of going from patch  $j$  to  $i$  is represented with the black arrows and encoded in the matrix element  $\mathcal{P}_{ij}$ . The agents can either move around in the network, and in that case they are represented along the arrows outside the spatial patches, or stay where they are and undergo one of two epidemiological “reactions”: infection or recovery. The infections are represented with the transformations with red arrows, while for the recoveries we used blue arrows. Finally according to the time scale separation parameter  $\epsilon$  we can distinguish two regimes of this model: in (a)  $\epsilon > 1$ , so reaction events are more frequent than diffusive ones while in (b)  $\epsilon < 1$ , so the opposite is true.

This richness of dynamic scenarios can be encoded into a tunable parameter  $\epsilon$  that slows down or speeds up diffusion with respect to epidemic spreading. We define it as the ratio  $\epsilon = \Delta t_{\text{mob}}/\Delta t_{\text{epi}}$  between the two time scales, so that the differential equations for the model become

$$\frac{dI_i(t)}{dt} = \beta_i S_i(t) \frac{I_i(t)}{N_i(t)} - \mu I_i(t) + \frac{1}{\epsilon} \left[ \sum_{j=1}^M \mathcal{P}_{ji} I_j(t) - \sum_{j=1}^M \mathcal{P}_{ij} I_i(t) \right] \quad (3)$$

see Methods for a more explicit derivation of this equation. In the remaining of this paper, we will refer to this model as the  $\epsilon$ -model, since the presence of the  $\epsilon$  parameter, characterizing the interplay between diffusive and reactive time scales, is its most distinctive feature. At this point, it is worth remarking that we can redefine both  $\mathcal{P}$  and  $\epsilon$  without loss of generality such that  $\sum_{j=1}^M \mathcal{P}_{ij} = 1$  and all the information on the time scale of the mobility lies in  $1/\epsilon$ . The reason for doing this will become clear in the next section.

Note that having a parameter that multiplies the mobility term in a metapopulation model is not uncommon<sup>29,35</sup>. In these cases such a parameter indicates the fraction of people that move between patches in a given time step. This is exactly equivalent to our interpretation of  $\epsilon$  as the ratio between the time scales of the two processes of reaction and diffusion.

**Calculation of the  $R_0$ .** In a simple SIR model without spatial structure, where every agent interacts with every other agent and there is only one transmission rate  $\beta$  for all spatial patches, it can be shown that the basic reproduction number is simply  $R_0 = \beta/\mu$ . However, in our framework we assume that there is a  $\beta_i$  for each patch  $i$  and we expect to have some dependence also on the mobility matrix and on  $\epsilon$ . To this aim, we use the next-generation

matrix formalism proposed by van den Driessche et al.<sup>46</sup> to calculate the critical threshold  $R_0(\epsilon)$ .

First, it is convenient to rewrite Eq. (3) in terms of the concentration of infectious and susceptible individuals, indicated respectively as  $\rho_i(t) = I_i(t)/N_i(t)$  and  $\sigma_i(t) = S_i(t)/N_i(t)$ . For simplicity, it is also useful to write the differential equation in matrix form as

$$\dot{\rho}(t) = \mathcal{B}\sigma(t) \cdot \rho(t) - \mu\rho(t) - \frac{1}{\epsilon} \mathcal{L}^\top \rho(t), \quad (4)$$

where  $\mathcal{L}^\top = \mathbf{I} - \mathcal{P}^\top$  and  $\mathcal{B} = \text{diag}(\beta_1, \dots, \beta_n)$ . Let us consider the infection-free steady state, i.e., the case right above the critical point, meaning that there is a negligible fraction of infected individuals ( $\sigma_i = 1, \forall i$ ). The resulting system of equations reduces to

$$\dot{\rho}(t) = \mathbf{J}\rho(t), \quad (5)$$

where  $\mathbf{J} = \mathcal{B} - \mu\mathbf{I} - \epsilon^{-1}\mathcal{L}^\top$  is the Jacobian matrix, whose analysis allows one to gain insights about the linear stability of the original system dynamics. By decomposing the matrix into  $\mathbf{J} = \frac{1}{\epsilon}(\mathbf{F} - \mathbf{V})$ , it is possible to demonstrate that the highest eigenvalue of the so-called next-generation matrix  $\mathbf{K} = \mathbf{F}\mathbf{V}^{-1}$  is in fact the basic reproduction number  $R_0$ <sup>46,47</sup>. In our case  $\mathbf{F} = \epsilon\mathcal{B}$  and  $\mathbf{V} = \epsilon\mu\mathbf{I} + \mathcal{L}^\top$ , so the matrix  $\mathbf{K}$  is

$$\mathbf{K} = \epsilon\mathcal{B}(\epsilon\mu\mathbf{I} + \mathcal{L}^\top)^{-1}. \quad (6)$$

However, solving the eigenvalue problem for this  $M \times M$  matrix is, in general, a very complicated task and tells you very little about the general dependences of  $\mathbf{K}$  on the parameters of the problem. To gain some insights, we look for approximated solutions in the extreme cases in which mobility is either much slower or much faster than the epidemic, in order to characterize the behavior of the system from an analytical perspective.

**Diffusion slower than reaction.** In this first case we have that  $\epsilon \gg 1$ , so we can use perturbation theory (PT) using  $\epsilon^{-1}$  as our parameter and writing the next-generation as  $\mathbf{K} = \mathbf{K}_0 + \mathbf{K}'/\epsilon$ , where

$$\mathbf{K}_0 = \frac{\mathbf{B}}{\mu} \quad \text{and} \quad \mathbf{K}' = -\frac{\mathbf{B}\mathcal{L}^\top}{\mu} \left( \mathbf{I} + \frac{\mathcal{L}^\top}{\epsilon\mu} \right)^{-1}. \quad (7)$$

The eigenvalues and eigenvectors of the unperturbed part are respectively  $\beta_i/\mu$  and the canonical basis  $\{\mathbf{e}_i\}$ . The resulting spectral radius, quantifying  $R_0$ , is given by

$$R_0 = \frac{\epsilon\beta_{\max}}{\mu\epsilon + (1 - \mathcal{P}_{\max,\max})} + \frac{\mu\epsilon^2}{(\mu\epsilon + 1)^4} \sum_{j \neq \max} \frac{\beta_{\max}\beta_j}{\beta_{\max} - \beta_j} \times \left( \frac{\mathcal{P}}{\mathbf{I} + \mathcal{P}/(1 + \mu\epsilon)} \right)_{\max,j} \left( \frac{\mathcal{P}}{\mathbf{I} + \mathcal{P}/(1 + \mu\epsilon)} \right)_{j,\max} + O(\epsilon^{-3}), \quad (8)$$

where the index *max* indicates the patch with the highest transmission rate (see Methods). Equation (8) highlights that when there is no mobility at all ( $\epsilon \rightarrow \infty$ ) one obtains  $R_0 = \beta_{\max}/\mu$ : i.e., the condition for the epidemic to have a finite-size outbreak is for that to happen at least in one spatial patch, as expected. It should be noted that in this case we have a strong dependence of the systems on the initial distribution of infected individuals, since the epidemic cannot grow unless the patches with  $\beta_i > \mu$  have at least one infected individual at time  $t = 0$ . Furthermore, notice how the mobility network  $\mathcal{P}$  only appears in the second order term and therefore its effects become more and more negligible as  $\epsilon$  increases.

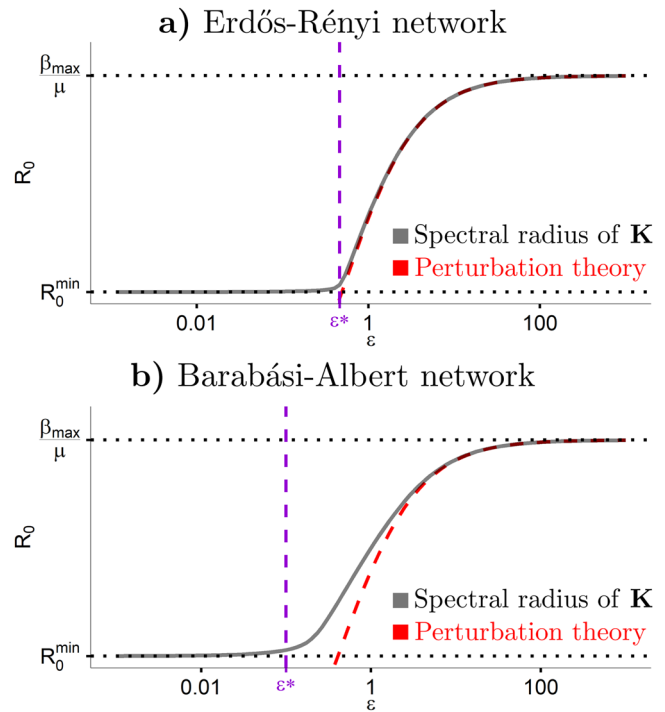
**Diffusion faster than reaction.** When the mobility is much faster than the epidemic (meaning that  $\epsilon \ll 1$ ) we cannot use the same approach, since for  $\epsilon \rightarrow 0$  the next-generation matrix becomes ill-defined as it is defined in terms of the inverse of the Laplacian matrix  $\mathcal{L}$ , which is singular. Nevertheless, one can solve this problem by using a corollary to Perron–Frobenius theorem which states that if one has a non-negative irreducible matrix, such as  $\mathbf{K}$ , then its highest eigenvalue corresponds to the spectral radius  $\rho(\mathbf{K})$  and the following inequalities hold:

$$\min_i \sum_j K_{ij} \leq \rho(\mathbf{K}) \leq \max_i \sum_j K_{ij}. \quad (9)$$

It is possible to show (see Methods) that in the limit  $\epsilon \rightarrow 0$  the next-generation matrix can be found exactly, leading to

$$R_0^{\min} = \sum_j \frac{\beta_j}{\mu} \frac{s_j}{\sum_k s_k}. \quad (10)$$

A closer inspection of the above results reveals that in this scenario  $R_0$  is the weighted average of the local basic reproduction numbers—defined by  $\beta_i/\mu$  in the case of isolated spatial patches—where the  $i$ -th weight is given by the relative strength  $s_i$  of node  $i$ , i.e., the sum of all the flows involving that node, with respect to the overall human flow. It is straightforward to verify that the above relation reduce to the expected  $R_0 = \beta/\mu$  when  $\beta_i = \beta$  for all spatial patches. It is also worth noticing that Eq. (10) tells us that  $R_0^{\min}$  does not depend on degree correlations or on topology in general and, as a result, as  $\epsilon$  vanishes, every network with the same strength sequence leads exactly to the same result. That also means that, in the case of unweighted metapopulation networks, the configuration model constitutes a good approximation with respect to  $R_0$  when  $\epsilon \rightarrow 0$ . The notation  $R_0^{\min}$  encodes the fact that, since  $R_0(\epsilon)$  is a monotonically increasing function, the basic reproduction number for  $\epsilon \rightarrow 0$  corresponds to the lower bound (see Fig. 2).



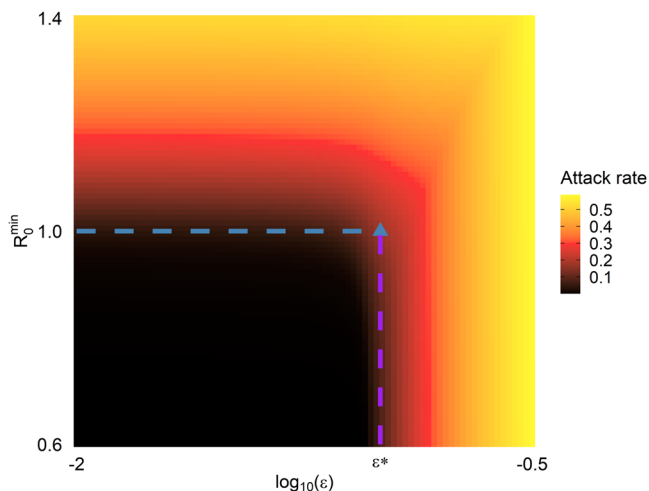
**Fig. 2 Basic reproduction number  $R_0$  of the whole metapopulation network as a function of the time scale ratio  $\epsilon$ .** The two panels show the behavior of a (a) Erdős-Rényi and a (b) Barabási-Albert metapopulation network, models for homogeneous and scale-free connectivity of spatial patches, respectively. The solid gray curve indicates the values of  $R_0$  found numerically by solving the eigenvalue problem for the next-generation matrix  $\mathbf{K}$ , whereas the red line corresponds to the value analytically derived in Eq. (8) by using perturbation theory. The approximation is excellent for the homogeneous connectivity, whereas in the case of the heterogenous one it works only for  $\epsilon > 1$ . The vertical dashed violet line indicates the position of the metacritical point  $\epsilon^*$  obtained analytically, that separates the regime in which  $R_0$  stays constant and equal to its minimum value  $R_0^{\min}$  (see Eq. (10)) from the regime in which it start growing. Finally the horizontal dotted black lines correspond to the lower bound on  $R_0$ , which in this case is  $R_0^{\min} = 1$ , and the upper bound  $\beta_{\max}/\mu = 2$ , where  $\mu = 2$  is the recovery rate and  $\beta_{\max} = 4$  is the maximum transmission rate among those in the network.

**Diffusion comparable with reaction.** It is plausible to wonder about what happens in the intermediate regime  $\epsilon \simeq 1$ . Figure 2 shows that  $R_0(\epsilon)$  can be written as the piecewise-definite function

$$f(\epsilon) = \begin{cases} R_0^{\min} & \text{for } \epsilon < \epsilon^* \\ \frac{(\beta_{\max} - b) \cdot \epsilon}{\mu} + \frac{b}{\mu} & \text{for } \epsilon > \epsilon^* \end{cases} \quad (11)$$

where  $R_0^{\min}$  is given by Eq. (10) and  $a$ ,  $b$  and  $\epsilon^*$  are three parameters, one of which is fixed by the continuity condition of  $f(\epsilon)$  in  $\epsilon = \epsilon^*$ . We can then safely assess that  $R_0(\epsilon)$  displays itself a phase transition, and as such it is possible to distinguish between the ordered and disordered phase just by looking at the initial condition in the number of infectious individuals in each patch. For instance, for  $\epsilon < \epsilon^*$  the diffusion is so fast compared to the reaction that no matter what the initial conditions are in the long-time limit the number of infected individuals will be uniformly spread among all the patches; therefore this phase is ordered or, equivalently, invariant under the choice of initial conditions. Conversely, when  $\epsilon > \epsilon^*$  the long-time limit will strongly depend





**Fig. 3 Phase diagram representing the attack rate in simulations parameterized by  $\epsilon$  and  $R_0^{\min}$ .** We recall that  $\epsilon$  is the parameter that indicates the ratio between the time scale of the mobility and that of the epidemic, while  $R_0^{\min}$  is the weighted average of the local basic reproduction numbers—defined in the case of isolated spatial patches—using the patches' strengths as weights (Eq. (10)). The dashed lines correspond to  $R_0^{\min} = 1$  (blue) and the correspondent critical value of the parameter  $\epsilon = \epsilon^*$  (violet), while the blue triangle indicates the position of the metacritical point. An Erdős–Rényi network with 500 nodes and wiring probability  $p = 0.04$  is considered.

on initial conditions, therefore the aforementioned invariance is broken and the phase is disordered. This behavior may be seen as analogous to the so called “global invasion threshold”<sup>48</sup>, which indicates the condition under which the epidemic goes from being localized only in a few geographical patches to invade the whole network.

To estimate the location of the critical point in the parameter space we can either obtain it from fitting Eq. (11)—as we have done for drawing the dotted lines in Fig. 2—or look for the interception between the sigmoid in Eq. (8) and the horizontal line  $R_0 = R_0^{\min}$ . In particular for  $R_0^{\min} = 1$  and its corresponding  $\epsilon^*$  we get a metacritical point: i.e., a value of the critical threshold that divides absorbing states from active ones (see Fig. 3). In this case we have  $\epsilon^* = 1/(\beta_{\max} - \mu)$ , obtained by intersecting the line  $R_0 = 1$  with the first term in Eq. (8) while assuming that there are no self-loops. We obtain a really good estimate for synthetic mobility networks characterized by homogeneous connectivity and our approximation still works as an upper bound for the ones characterized by heterogeneous connectivity. This is true because, as shown in Fig. 2, in heterogeneous networks the second order term in PT starts being relevant, and since this term can only be positive the net effect is to lowering the metacritical point. Therefore, we can conclude that the maximum possible value for  $R_0$  should be the one obtained only with the first order in PT, that is very close to that of homogeneous networks (see again Fig. 2). Finally, note that if Eq. (11) is correct (as it is for homogeneous networks) then the critical exponent  $\nu$  given by  $R_0 \sim (\epsilon - \epsilon^*)^\nu$  for  $\epsilon \rightarrow \epsilon^*$ , is equal to 1.

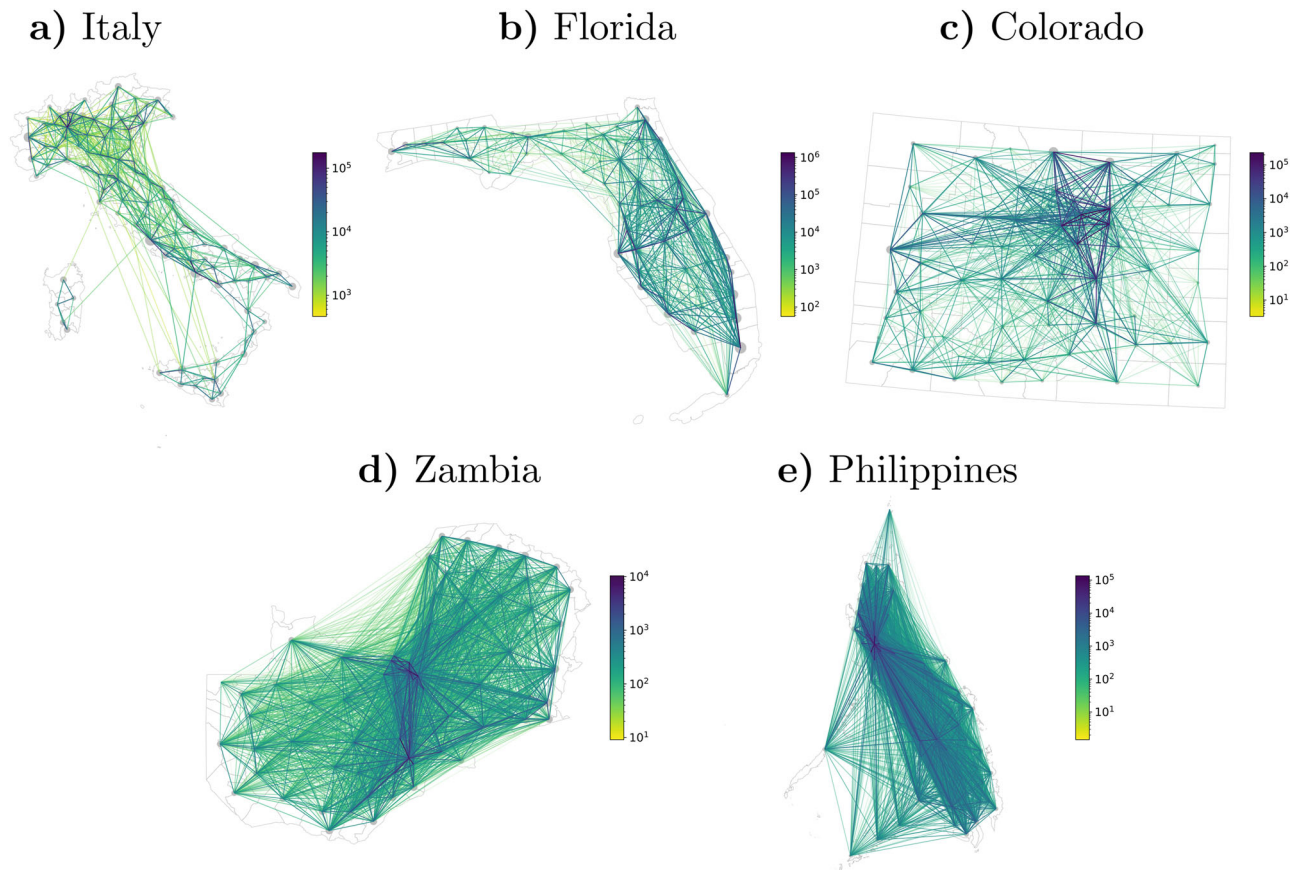
**Metacritical point observed in real human flows.** We now apply our theoretical framework to analyze empirical metapopulation networks, where nodes are either counties of a U.S. state or other types of administrative regions in other countries, and where links are observed human mobility flows.

We use a series of datasets describing the anonymized mobility flows in different areas of the world, aggregated at administrative levels. All sources originated by the usage of mobile phones: this kind of passively collected data has been successfully validated for epidemics modeling<sup>49</sup> and became standard in recent years. The flows for France (FRA), Spain (ESP) and Portugal (PRT)<sup>49</sup>, Ivory Coast (CIT)<sup>7</sup> and Turkey (TUR)<sup>33</sup> are reconstructed from the Call Detail Records, the billing dataset held by mobile phone companies. The flows for USA (Florida and Colorado) are estimated on the base of data kindly provided by Cuebiqu, a location intelligence and measurement platform that collects privacy enhanced GPS trajectories from mobile apps users who have opted-in to provide access to their aggregated location data anonymously. Also those of Italy (ITA) are derived from Cuebiqu data, and are obtained from a published dataset<sup>50</sup>. These flows have been used for assessing change in mobility during the COVID-19 pandemics in 2020<sup>51,52</sup>, but for our purposes we analyze here only a static snapshot of the mobility in the area. See Fig. 4(a), (b) and (c) for an illustration of the mobility flows in Italy, Florida, and Colorado. For these datasets, the population of each node is obtained from census data.

The flows for Senegal (SEN), Nigeria (NGA), Zambia (ZMB), China (CHN), India (IND), Mexico (MEX), Brazil (BRA) and the Philippines (PHL) are publicly available<sup>53</sup> and have been reconstructed using an extended gravity model including several geographical and socioeconomic factors. See Fig. 4(d) and (e) for an illustration of these reconstructed flows for Zambia and The Philippines. In this case, the population at each site is assumed to be proportional to the total out-bound flow, upscaled so that the sum of the node populations matches the country population.

In order to verify the validity of our theoretical framework when dealing with real-world mobility flows, we assign to each node a transmission rate  $\beta_i$  which is proportional to its population. We choose to do this because it is reasonable to think that more populated areas are linked to a larger number of interaction and therefore to a higher risk of transmission. However we cannot expect this mechanism to hold for arbitrarily large populations, since the number of contact each person can have has a limit that does not depend on the size of the population itself (we got the idea behind this from<sup>54,55</sup>). For this reason we set the maximum value to  $\beta_{\max} = 2\mu$ . In addition we set  $R_0^{\min} = 0.5$  to be sure that we are in the regime where a transition happens for some value of  $\epsilon$  (see Fig. 3). Each simulation was run until the variation in the number of recovered individuals after 20 days was 0.01% of the total population (note that for China and India this fraction was lowered to 0.0001% due to the large number of people living inside each node). For each of simulation, initial conditions were set to ten infected individuals per node. Finally the heterogeneity of every weighted network was calculated using the Gini coefficient  $G$  of the strength distribution of the network.

The results are shown in Fig. 5, highlighting the behavior of the attack rate at different stages of the epidemic, for various real-world mobility networks. Here with epidemic stages we simply mean the status reached by the epidemic after different times, measured in multiples of the time necessary to reach the maximum number of infected (or infectious peak). We chose to do this so that, even if the absolute duration of each simulation had varied, we would have been at the same point (or “stage”) of the evolution of the epidemics. The results display, qualitatively, a similar behavior: for low values of  $\epsilon$  the attack rate is approximately zero, meaning that there is no epidemic outbreak at all, while for larger  $\epsilon$  this number grows rapidly, confirming the existence of a phase transition in the curve of critical thresholds  $R_0(\epsilon)$ . Notice that such a transition point  $\epsilon^*$  always remains



**Fig. 4 Real human mobility flows used in this study.** Metapopulation networks consisting of empirical populations (nodes) and empirical human flows (links) are shown, embedded in space (coordinates are given by longitude and latitude), in the case of **(a)** Italy; **(b)** Florida (USA); **(c)** Colorado (USA); **(d)** Zambia; **(e)** The Philippines. For Italy, Florida and Colorado flows are movements between provinces or counties, estimated using anonymized GPS trajectories of mobile apps users and upscaled to the total population. For Zambia and the Philippines, flows are migrations between administrative units reconstructed through gravity modeling<sup>53</sup>. The intensity of edge colors is proportional to the observed flows, while node sizes are proportional to the local population.

approximately below the upper bound  $1/(\beta_{\max} - \mu)$ , which is equal to 10 in this case. Finally, when  $\epsilon$  is very large there is a drop in the attack rate, which happens because inter-node travels are so limited that the epidemic can only spread inside the nodes where  $\beta_i > \mu$ , quickly reaching the stationary state.

Note that here, when taking the long-time limit or stationary state, we are actually referring to the stage in which the epidemic has slow down so much that variations are extremely small. Curves in Fig. 5 weakly depend on such a condition, even though their qualitative shape would remain the same: an abrupt growth for low values of  $\epsilon$  followed by a sharp drop for high values of  $\epsilon$ .

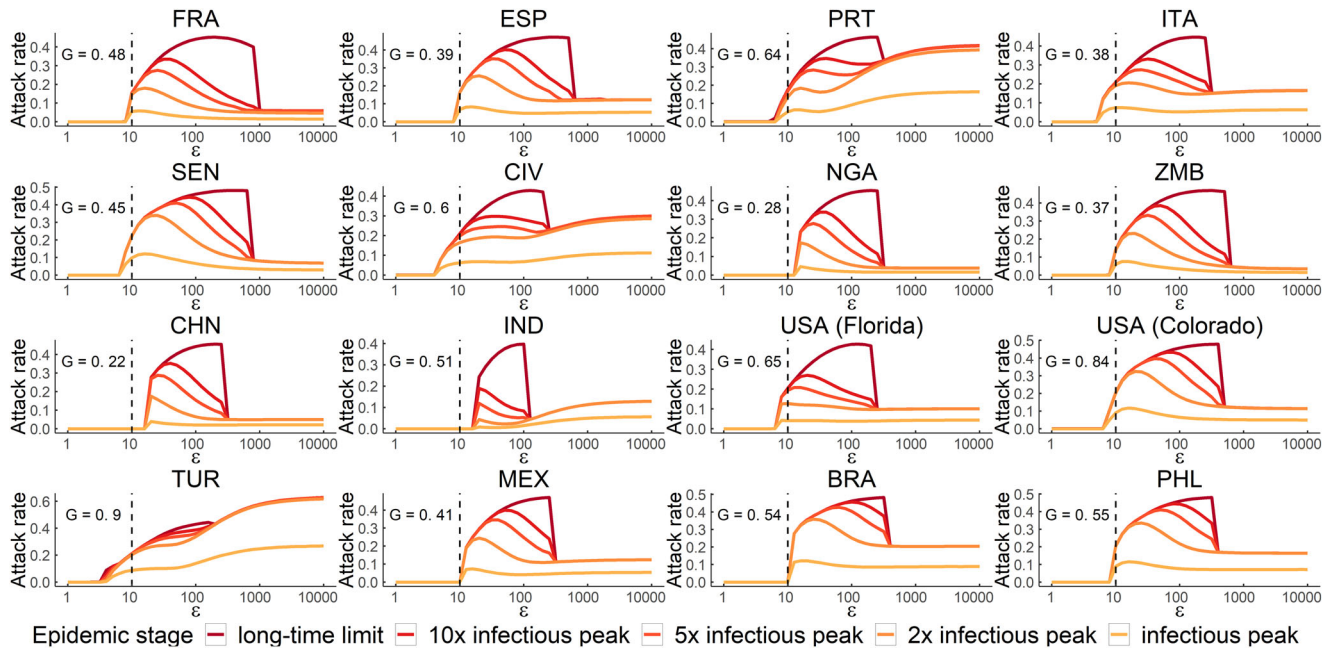
Finally, from a quantitative perspective, the different behavior of the attack rate across countries depends mainly on the population distribution at their sites and on the degree distribution of the corresponding metapopulation networks. In particular, countries with a more homogeneous degree distribution (such as Zambia or Philippines) tend to have a transition point  $\epsilon^*$  closer to the upper bound  $1/(\beta_{\max} - \mu)$ , while for heterogeneous networks (such as Portugal or Turkey) this value is lower, as confirmed by Fig. 2. Then, as  $\epsilon$  increases, the population distribution becomes more and more important, since in this limit people cannot move so much and they are bound to stay in their original patch. In this case, the outbreak will spread only in those sites where the local  $R_0$  is larger than one. Since we have defined transition rates to be proportional to the local population,

the nodes with the highest  $\beta_i$  are also the most populated ones. Therefore, if we have a network with a very heavy-tailed population distributions (such as Turkey) the attack rate in the limit  $\epsilon \rightarrow \infty$  will be large while, conversely, in the case of a network with a more homogeneous population distribution (such as Nigeria) this limit is low.

## Conclusion

In this work, we introduced a model for epidemic spreading on metapopulation network with arbitrary topology. It is structured as a reaction-diffusion process where the reaction part is that of a standard SIR model with site-dependent transition rates, while the diffusive part corresponds to a random walk whose dynamics is scaled by a parameter which allows us to tune the speed of mobility with respect to the epidemic spread. This framework is more convenient than classical single-scale reaction-diffusion models when conditions such as non-pharmaceutical interventions and behavioral changes become non-standard, altering the characteristic temporal scales of human mobility.

We provided analytic predictions for the basic reproduction number  $R_0$  in two extreme cases, namely the “diffusion much faster than reaction” regime, thanks to a perturbative approach, and the “diffusion much slower than reaction” regime, where we make use of the Perron–Frobenius theorem. In the latter case we



**Fig. 5 Criticality in empirical networks characterized by measured human flows and population distribution.** The panels show the dependence of the attack rate on the scale parameter  $\epsilon$  for different metapopulation networks, each one identified with the ISO 3166-1 alpha-3 code of the country it describes. Different colors encode distinct stages of the epidemic as shown in the legend. With epidemic stages we simply mean the stages of development of the epidemic measured using as time unit the time necessary for an epidemic to reach the maximum number of infected (or infectious peak). In all cases, the critical point  $\epsilon^*$  is clearly visible, since it corresponds to the value of  $\epsilon$  such that the attack rate is larger than zero. It is worth noticing that such a critical point is approximately always lower than our analytically found upper bound  $1/(\beta_{\max} - \mu)$  represented with a vertical dashed black line (with  $\beta_{\max} = 0.2$  and  $\mu = 0.1$  in this case). The  $G$  in each sub-figure indicates the Gini coefficient of the strength distribution of the corresponding network and it is used to quantify the heterogeneity of each network. Finally, we notice a drop in the attack rate for large values of  $\epsilon$ : an effect due to the slowing down of the mobility.

discovered that the global  $R_0$  can be well approximated by the weighted average of the local  $R_0$  with the patches' degrees as their weights.

In the intermediate regime, the most interesting one, we discovered the existence of a metacritical point on synthetic networks and on real human mobility flows. Such a point separates two distinct regimes observed in the behavior of the critical threshold, identifying the region of the parameter space where the conditions for having a non-vanishing attack rate change abruptly, providing useful information on how the different interventions to control or mitigate the epidemic in each site work in combination with travel restrictions.

To the best of our knowledge this type critical behavior in the curve of critical points has been observed in the presence of intertwined dynamical processes such as awareness and epidemics<sup>6</sup>—where the spreading of awareness about a disease is able to inhibit the epidemic spreading—and in competing epidemics spreading<sup>56</sup>—where the spreading of a disease is able to facilitate or inhibit infection by another disease—on multilayer networks<sup>57</sup>, but it is the first time that it is observed for multiscale processes in classical networked systems.

**Methods**

**Derivation of the  $\epsilon$ -model.** Here we explain how to go from Eq. (1) and (2) to Eq. (3). First of all we recall that  $\Delta t_{\text{epi}}$  and  $\Delta t_{\text{mob}}$  are the time intervals that we have to wait in order to get a variation in the number of infected individuals equal to some arbitrarily fixed value. This value however should be small, because the transition rates in both Eq. (1) and (2) depend on time, so waiting too long would mean to change the value of the transition rates. For instance we could decide that  $\Delta t_{\text{epi}}$  and  $\Delta t_{\text{mob}}$  are those that you need to have  $\Delta I_i^{\text{epi}}$  and  $\Delta I_i^{\text{mob}}$  equal to 1, respectively.

Now we ask ourselves: at each time step how much of the the variation in infected individuals is due to epidemics and how much to mobility? In order to

answer that we compute the joint variation in infected individuals after an arbitrary time  $\Delta t$

$$\begin{aligned} \Delta I_i(t) &= \left( \beta_i S_i(t) \frac{I_i(t)}{N_i(t)} - \mu I_i(t) \right) \Delta t \\ &\quad + \left( \sum_{j=1}^M \mathcal{P}_{ji} I_j(t) - \sum_{j=1}^M \mathcal{P}_{ij} I_i(t) \right) \Delta t \\ &= \left( \beta_i S_i(t) \frac{I_i(t)}{N_i(t)} - \mu I_i(t) \right) \Delta t_{\text{epi}} \frac{\Delta t}{\Delta t_{\text{epi}}} \\ &\quad + \left( \sum_{j=1}^M \mathcal{P}_{ji} I_j(t) - \sum_{j=1}^M \mathcal{P}_{ij} I_i(t) \right) \Delta t_{\text{mob}} \frac{\Delta t_{\text{epi}}}{\Delta t_{\text{mob}}} \frac{\Delta t}{\Delta t_{\text{epi}}} \\ &= \frac{1}{\Delta t_{\text{epi}}} \left( \Delta I_i^{\text{epi}}(t) \Delta t + \frac{1}{\epsilon} \Delta I_i^{\text{mob}}(t) \Delta t \right) \end{aligned} \tag{12}$$

From this equation one can see that the contribution to  $\Delta I_i^{\text{epi}}$  due to the epidemics is always a factor  $1/\epsilon$  larger (or smaller) than the contribution  $\Delta I_i^{\text{mob}}$  due to the mobility, which justify the presence of the  $1/\epsilon$  factor in Eq. (3). Finally if we incorporate the  $1/\Delta t_{\text{epi}}$  factor in the definitions of the rates and then send  $\Delta t$  to zero we end up with the differential equation of the  $\epsilon$ -model.

**Perturbation theory.** First of all we recall that the general form of the next-generation matrix for the  $\epsilon$ -model is  $\mathbf{K} = \epsilon \mathcal{B}(\epsilon \mu \mathbf{I} + \mathcal{L}^T)^{-1}$  and that we can write it us a sum  $\mathbf{K} = \mathbf{K}_0 + \epsilon^{-1} \mathbf{K}'$  where the terms are defined in Eq. (7). We also recall that the aim of PT is that of writing the eigenvalues as a power series in the parameter  $\epsilon^{-1}$ . In our case it will be sufficient to consider only the expansion up to second order, that is

$$\lambda_i = \lambda_i^{(0)} + \frac{1}{\epsilon} \lambda_i^{(1)} + \frac{1}{\epsilon^2} \lambda_i^{(2)} + O\left(\frac{1}{\epsilon^3}\right) \tag{13}$$

The unperturbed eigenvalues are simply  $\lambda_i^{(0)} = \beta_i/\mu$ , which are the “local” basic reproduction numbers, meaning the values that  $R_0$  would assume inside each patch if they were all isolated, while the unperturbed eigenvectors are those of the



canonical basis  $\{e_i\}$ . On the other hand the first order contribution is

$$\begin{aligned} \lambda_i^{(1)} &= e_i^\top K' e_i \\ &= -e_i^\top \frac{\mathcal{B} \mathcal{L}^\top}{\mu} \left( I + \frac{\mathcal{L}^\top}{\epsilon \mu} \right)^{-1} e_i \\ &= -\beta_i \left( \frac{\mathcal{L}_{ii}^\top}{\mu} - \frac{(\mathcal{L}^\top)^2_{ii}}{\epsilon \mu^2} + \frac{(\mathcal{L}^\top)^3_{ii}}{\epsilon^2 \mu^3} - \dots \right) \end{aligned} \tag{14}$$

Putting together the first two terms of the power expansion (Eq. (13)) we get

$$R_0 = \beta_{\max} \frac{\epsilon}{\mu \epsilon + (1 - \mathcal{P}_{\max, \max})} + O(\epsilon^{-2}) \tag{15}$$

where the *max* index points to the node with the highest transmission rate  $\beta_{\max}$ . For the second order in PT the key factor to compute is

$$\begin{aligned} e_i^\top K' e_j &= -e_i^\top \frac{\mathcal{B} \mathcal{L}^\top}{\mu} \left( I + \frac{\mathcal{L}^\top}{\epsilon \mu} \right)^{-1} e_j \\ &= -e_i^\top \left[ \frac{\mathcal{B} \mathcal{L}^\top}{\mu} - \frac{\mathcal{B} \mathcal{L}^\top}{\mu} \left( \frac{\mathcal{L}^\top}{\epsilon \mu} \right) + \frac{\mathcal{B} \mathcal{L}^\top}{\mu} \left( \frac{\mathcal{L}^\top}{\epsilon \mu} \right)^2 - \dots \right] e_j \\ &= \frac{\beta_i}{\mu} \left[ \mathcal{P}_{ij} \sum_{n=1}^{\infty} (-1)^n \frac{n}{(\mu \epsilon)^n} + (\mathcal{P}^2)_{ij} \sum_{n=2}^{\infty} (-1)^n \frac{\binom{n}{2}}{(\mu \epsilon)^n} + \dots \right] \\ &= \frac{\beta_i}{\mu} \left[ -\mathcal{P}_{ij} \frac{\mu \epsilon}{(\mu \epsilon + 1)^2} + (\mathcal{P}^2)_{ij} \frac{\mu \epsilon}{(\mu \epsilon + 1)^3} - \dots \right] \\ &= \beta_i \frac{\epsilon}{(\mu \epsilon + 1)^2} \left( \frac{\mathcal{P}}{I + \mathcal{P}/(1 + \mu \epsilon)} \right)_{ij} \end{aligned} \tag{16}$$

Putting this in the formula for the second order contribution and adding it to Eq. (15) we obtain Eq. (7).

**The  $\epsilon \rightarrow 0$  limit.** We now approach the problem of computing the next-generation matrix  $K = FV^{-1}$  for an arbitrary mobility network. We proceed step by step by computing one piece at the time, starting with the elements of  $V = \epsilon \mu I + \mathcal{L}^\top$ . This matrix in general is not symmetric, meaning that the right and left eigenvalues are different

$$\begin{aligned} \sum_j V_{ij} \varphi_j^{(n)} &= \nu^{(n)} \varphi_i^{(n)} \\ \sum_j \psi_j^{(n)} V_{ji} &= \nu^{(n)} \psi_i^{(n)} \end{aligned} \tag{17}$$

where  $\varphi^{(n)}$  and  $\psi^{(n)}$  are respectively the right and left eigenvectors of both  $V$  and  $\mathcal{L}^\top$ , while the eigenvalues  $\nu^{(n)}$  of  $V$  differ from the eigenvalues  $\lambda^{(n)}$  of  $\mathcal{L}^\top$  by a term  $\epsilon \mu$ . Even though we do not know neither the eigenvectors nor the eigenvalues it is useful to write the matrix in the following form

$$V_{ij} = \sum_n \nu^{(n)} \varphi_i^{(n)} \psi_j^{(n)} \tag{18}$$

Finally if the matrix  $V$  is invertible, and we know it is thanks to the term  $\epsilon \mu I$ , we can write the first building block of the next-generation matrix, that is

$$(V^{-1})_{ij} = \sum_n \frac{1}{\nu^{(n)}} \varphi_i^{(n)} \psi_j^{(n)} = \sum_n \frac{1}{\epsilon \mu + \lambda^{(n)}} \varphi_i^{(n)} \psi_j^{(n)} \tag{19}$$

The second step consists in multiplying what we have with the matrix  $\epsilon \mathcal{B}$  and then take the limit for  $\epsilon \rightarrow 0$ . This will lead to

$$\begin{aligned} K_{ij} &= \epsilon (\mathcal{B} V^{-1})_{ij} \\ &= \sum_n \frac{\epsilon \beta_i}{\epsilon \mu + \lambda^{(n)}} \varphi_i^{(n)} \psi_j^{(n)} \\ &\xrightarrow{\epsilon \rightarrow 0} \frac{\beta_i}{\mu} \varphi_i^{(1)} \psi_j^{(1)} \end{aligned} \tag{20}$$

where first we wrote the eigenvalues of  $V$  in function of those of  $\mathcal{L}^\top$  and then we noticed that the only term in the sum over  $n$  who was not multiplied by the vanishing  $\epsilon$  was the one associated to the null eigenvalue of the Laplacian matrix  $\lambda^{(1)} = 0$ , that corresponds to  $\nu^{(1)} = \epsilon \mu$ .

Fortunately for us we know that the eigenvectors corresponding to the stationary states are

$$\varphi^{(1)} = \frac{1}{2m} \begin{pmatrix} s_1 \\ \vdots \\ s_n \end{pmatrix} \quad \text{and} \quad \psi^{(1)} = \begin{pmatrix} 1 \\ \vdots \\ 1 \end{pmatrix} \tag{21}$$

where  $s_i$  is the strength of the  $i$ -th node.

Finally we can apply the corollary to Perron–Frobenius theorem mentioned in the main article, but instead of doing it on  $K$  we do it on  $K^\top$  which has the same spectrum and the property of having all the rows equal to each other, so that upper

**Table 1 Table of the fundamental properties of the networks that appear in Fig. 5.**

Country	# of nodes	# of edges	Gini
France	329	40,203	0.48
Spain	47	2125	0.39
Portugal	278	15,045	0.64
Italy	107	1454	0.38
Senegal	34	1121	0.45
Ivory Coast	237	20,117	0.6
Nigeria	38	1405	0.28
Zambia	72	5111	0.37
China	30	869	0.22
India	32	991	0.51
USA (Florida)	67	901	0.65
USA (Colorado)	63	694	0.84
Turkey	466	28,635	0.9
Mexico	32	991	0.41
Brazil	27	701	0.54
Philippines	77	5851	0.55

The bold font indicates the type of information contained in the corresponding column. Here the word “Gini” refers to the Gini coefficient of the strength distribution of the corresponding network, while the symbol # stands for the word “number”.

and lower bound coincide and we get that the spectral radius of  $K$  is exactly that written in Eq. (10).

**Mobility networks data.** Our analysis of real human flows are based on (i) mobility flow networks and (ii) local population that were obtained from a broad range of different data sources. We have in general two types of flows: estimated and empirical. In the estimated category fall the flows for Senegal (SEN), Nigeria (NGA), Zambia (ZMB), China (CHN), India (IND), Mexico (MEX), Brazil (BRA) and the Philippines (PHL). In this case the data are therefore not representing real flows but the estimated internal human migration flows in a series of countries where malaria is endemic. The reconstruction is the product of the complex data-fusion of census-based migration microdata, which are assimilated through gravity modeling. For further detail please refer to the Data Descriptor<sup>53</sup>. Being the local population resident in the nodes not part of the data shared, we used as a proxy the total out-bound flow, which we upscled so that the sum of all nodes populations was equal to the total population of the country. All other datasets represent instead empirical flows and are associated with node population estimated from census data. The FRA, ESP and PRT<sup>49</sup>, CIT<sup>7</sup> and TUR<sup>33</sup> are reconstructed from Call Details Records, which provide information on the position (closest antenna) and time for each phone call, texting, or internet access of mobile users, which are recorded by the mobile companies for billing purposes. Please refer to the cited papers for the data availability of the respective dataset. The flows for USA (Florida and Colorado) are estimated on the basis of data kindly provided by Cuebiq, a location intelligence and measurement platform that collects privacy enhanced GPS trajectories from mobile apps users who have opted-in to provide access to their aggregated location data anonymously. Cuebiq provided us with information, on a weekly basis, about: (i) the ratio of users who have been seen moving between county  $i$  and county  $j$  to the number of users seen in county  $i$   $\text{ratio}Z_{ij} = (\text{users flow})_{ij}/(\text{county users})$ , (ii) the ratio of devices that we see in county  $i$  to all the devices seen across all counties  $U_i = (\text{county users})/(\text{total users})$ . We therefore upscled to the values of flows one would have if the movements observed were to represent the entire population with the formula

$$Z_{ij} = \text{ratio}Z_{ij} \times U_i \times \text{Pop} \tag{22}$$

where *Pop* is the total population across the country considered here, the United States of America. For this analysis, we selected only the flows associated to the first week of 2020, so before the diffusion of the COVID-19 epidemics in the USA. These data are available from the authors upon request. Finally, also the flows of Italy (ITA) are derived from Cuebiq data, which in this case are again publicly available and we invite to refer to the Data Descriptor<sup>50</sup> for more details on the methodology. The italian flows shared capture the fraction of the total userbase moving between two italian provinces. We upscled to the total population by simply multiplying by the population data obtained from census. In this case, we analyzed the flows associated with the 18th January 2020, before the beginning of the COVID-19 pandemic.

The fundamental descriptors of these networks can be found in the in Table 1.



**Simulation methodology.** Here we put all the information on the Figs. 2 and 3 that we could not fit in the main text.

For instance in Fig. 2 both the networks had  $n = 500$  nodes. The Erdős-Rényi network was undirected and sampled using a wiring probability  $p = 0.04$ ; the Barabási-Albert network was also undirected and sampled adding at each time step a single edge. In both networks the weights were random integer numbers between 1 and 50, drawn from a uniform distribution. Finally in both the networks the transmission rates  $\{\beta_i\}$  were chosen in such a way that the  $R_0^{\min} \in [0.95, 1.05]$  while  $\beta_{\max} = 2\mu$  (where  $\mu$  was set equal to 2). In order to do that we proceeded as follow: first we sampled  $\{\beta_i\}$  from a normal distribution centered in  $\mu$  and with a standard deviation of 0.2. Then we set the maximum value equal to  $2\mu$ . Finally we added a cycle in which we computed the  $R_0^{\min}$  of the current distribution and if it did not belong to the mentioned interval every  $\beta_i$  except the maximum one were rescaled by the current value of  $R_0^{\min}$ . The process was repeated until  $R_0^{\min}$  was in the desired interval.

The same sampling technique for the  $\{\beta_i\}$  distribution was used in Fig. 3, but in this case we had  $\mu = 2$  and  $\beta_{\max} = 6\mu$  while  $R_0^{\min}$  was a variable. The reason why we chose to sample the transmission rates in this way is that every curve  $R_0(\epsilon)$  is uniquely identified by its maximum and minimum value, so if one's aim is to observe how the critical point depends on  $R_0^{\min}$  we have to keep everything else constant, even  $\beta_{\max}$ . If we had allowed  $\beta_{\max}$  to vary too we would have ended up comparing completely different curves and any information about the metacritical point would have been lost.

Finally for the simulations in Figs. 3 and 5 we used the R package called “deSolve” with a time step of  $\min(1, \epsilon)/10$ .

## Data availability

The data on Senegal (SEN), Nigeria (NGA), Zambia (ZMB), China (CHN), India (IND), Mexico (MEX), Brazil (BRA) and the Philippines (PHL) have been taken from an open access dataset<sup>53</sup> produced within the WorldPop project. The flows for France (FRA), Spain (ESP) and Portugal (PRT) come from<sup>49</sup>, those of Ivory Coast (CIT) from<sup>7</sup> and those of Turkey (TUR) from<sup>33</sup>. The data regarding Italy (ITA) come from Cuebiq data (see the Data Descriptor<sup>50</sup> for more details on the methodology). Finally the data on USA (Florida and Colorado) were also provided by Cuebiq and are available upon request.

## Code availability

The script we used can be find at <https://github.com/PGcastioni/Epidemics-and-Mobility>.

Received: 6 January 2021; Accepted: 14 May 2021;

Published online: 11 June 2021

## References

1. Ferguson, N. Capturing human behaviour. *Nature* **446**, 733 (2007).
2. Funk, S., Salathé, M. & Jansen, V. A. A. Modelling the influence of human behaviour on the spread of infectious diseases: a review. *J. Royal Soc. Interfac.* **7**, 1247–1256 (2010).
3. Funk, S., Gilad, E., Watkins, C. & Jansen, V. A. The spread of awareness and its impact on epidemic outbreaks. *Proc. Natl Acad. Sci. USA* **106**, 6872–6877 (2009).
4. Verelst, F., Willem, L. & Beutels, P. Behavioural change models for infectious disease transmission: a systematic review (2010–2015). *J. Royal Soc. Interfac.* **13**, 20160820 (2016).
5. Poletti, P., Ajelli, M. & Merler, S. Risk perception and effectiveness of uncoordinated behavioral responses in an emerging epidemic. *Math. Biosci.* **238**, 80–89 (2012).
6. Granell, C., Gómez, S. & Arenas, A. Dynamical interplay between awareness and epidemic spreading in multiplex networks. *Phys. Rev. Lett.* **111**, 128701 (2013).
7. Lima, A., De Domenico, M., Pejovic, V. & Musolesi, M. Disease containment strategies based on mobility and information dissemination. *Sci. Rep.* **5**, 10650 (2015).
8. Ferguson, N. M., Donnelly, C. A. & Anderson, R. M. The foot-and-mouth epidemic in Great Britain: Pattern of spread and impact of interventions. *Science* **292**, 1155–1160 (2001).
9. Ferguson, N. M. et al. Strategies for mitigating an influenza pandemic. *Nature* **442**, 448–452 (2006).
10. Cauchemez, S., Valleron, A. J., Boëlle, P. Y., Flahault, A. & Ferguson, N. M. Estimating the impact of school closure on influenza transmission from Sentinel data. *Nature* **452**, 750–754 (2008).
11. Cauchemez, S. et al. Closure of schools during an influenza pandemic. *Lancet Infect. Dis.* **9**, 473–481 (2009).
12. Balcan, D. et al. Seasonal transmission potential and activity peaks of the new influenza A(H1N1): a Monte Carlo likelihood analysis based on human mobility. *BMC Med.* **7**, 45 (2009).
13. Poletti, P., Ajelli, M. & Merler, S. The Effect of Risk Perception on the 2009 H1N1 Pandemic Influenza Dynamics. *PLoS ONE* **6**, e16460 (2011).
14. Lloyd-Smith, J. O., Schreiber, S. J., Kopp, P. E. & Getz, W. M. Superspreading and the effect of individual variation on disease emergence. *Nature* **438**, 355–359 (2005).
15. Smith, R. D. Responding to global infectious disease outbreaks: Lessons from SARS on the role of risk perception, communication and management. *Soc. Sci. Med.* **63**, 3113–3123 (2006).
16. Merler, S. et al. Spatiotemporal spread of the 2014 outbreak of Ebola virus disease in Liberia and the effectiveness of non-pharmaceutical interventions: a computational modelling analysis. *Lancet Infect. Dis.* **15**, 204–211 (2015).
17. Wesolowski, A. et al. Impact of human mobility on the emergence of dengue epidemics in Pakistan. *Proc. Natl Acad. Sci. USA* (2015). URL <http://www.pnas.org/content/early/2015/09/02/1504964112.abstract>
18. Scarpino, S. V., Allard, A. & Hébert-Dufresne, L. The effect of a prudent adaptive behaviour on disease transmission. *Nat. Phys.* **12**, 1042–1046 (2016).
19. Carey, J. M., Chi, V., Flynn, D. J., Nyhan, B. & Zeitzoff, T. The effects of corrective information about disease epidemics and outbreaks: Evidence from Zika and yellow fever in Brazil. *Sci. Adv.* **6**, eaaw7449 (2020).
20. Lai, S. et al. Effect of non-pharmaceutical interventions to contain COVID-19 in China. *Nature* **585**, 410–413 (2020).
21. Davies, N. G. et al. Effects of non-pharmaceutical interventions on COVID-19 cases, deaths, and demand for hospital services in the UK: a modelling study. *Lancet Public Health* **5**, e375–e385 (2020).
22. Cowling, B. J. et al. Impact assessment of non-pharmaceutical interventions against coronavirus disease 2019 and influenza in Hong Kong: an observational study. *Lancet Public Health* **5**, e279–e288 (2020).
23. Chinazzi, M. et al. The effect of travel restrictions on the spread of the 2019 novel coronavirus (COVID-19) outbreak. *Science* **368**, 395–400 (2020).
24. Zhang, J. et al. Changes in contact patterns shape the dynamics of the COVID-19 outbreak in China. *Science* **368**, 1481–1486 (2020).
25. Hagenaars, T. J., Donnelly, C. A. & Ferguson, N. M. Spatial heterogeneity and the persistence of infectious diseases. *J. Theor. Biol.* **229**, 349–359 (2004).
26. Gómez-Gardeñes, J., Soriano-Paños, D. & Arenas, A. Critical regimes driven by recurrent mobility patterns of reaction-diffusion processes in networks. *Nat. Phys.* **14**, 391–395 (2018).
27. Soriano-Paños, D., Lotero, L., Arenas, A. & Gómez-Gardeñes, J. Spreading Processes in Multiplex Metapopulations Containing Different Mobility Networks. *Phys. Rev. X* **8**, 031039 (2018).
28. Colizza, V., Barrat, A., Barthélemy, M. & Vespignani, A. The role of the airline transportation network in the prediction and predictability of global epidemics. *Proc. Natl Acad. Sci.* **103**, 2015–2020 (2006).
29. Brockmann, D. & Helbing, D. The hidden geometry of complex, network-driven contagion phenomena. *Science* **342**, 1337–1342 (2013).
30. Balcan, D. et al. Multiscale mobility networks and the spatial spreading of infectious diseases. *Proc. Natl Acad. Sci. USA* **106**, 21484–9 (2009).
31. Della Rossa, F. et al. A network model of Italy shows that intermittent regional strategies can alleviate the COVID-19 epidemic. *Nat. Commun.* **11**, 1–9 (2020).
32. Gatto, M. et al. Spread and dynamics of the COVID-19 epidemic in Italy: Effects of emergency containment measures. *Proc. Natl Acad. Sci. USA* **117**, 10484–10491 (2020).
33. Bosetti, P. et al. Heterogeneity in social and epidemiological factors determines the risk of measles outbreaks. *Proc. Natl Acad. Sci.* **117**, 30118–30125 (2020).
34. Bertuzzo, E. et al. The geography of COVID-19 spread in Italy and implications for the relaxation of confinement measures. *Nat. Commun.* **11**, 1–11 (2020).
35. Belik, V., Geisel, T. & Brockmann, D. Natural human mobility patterns and spatial spread of infectious diseases. *Phys. Rev. X* **1**, 011001 (2011).
36. Rosvall, M., Esquivel, A. V., Lancichinetti, A., West, J. D. & Lambiotte, R. Memory in network flows and its effects on spreading dynamics and community detection. *Nat. Commun.* **5**, 1–13 (2014).
37. Matamalas, J. T., Domenico, M. D. & Arenas, A. Assessing reliable human mobility patterns from higher order memory in mobile communications. *J. Royal Soc. Interfac.* **13**, 20160203 (2016).
38. Balcan, D. et al. Modeling the spatial spread of infectious diseases: the GLOBE Epidemic and Mobility computational model. *J. Comput. Sci.* **1**, 132–145 (2010).
39. Gallotti, R., Bazzani, A. & Rambaldi, S. Towards a statistical physics of human mobility. *Int. J. Modern Phys. C* **23**, 1250061 (2012).
40. Barbosa, H. et al. Human mobility: Models and applications. *Phys. Rep.* **734**, 1–74 (2018).
41. Report 9: Impact of non-pharmaceutical interventions (NPIs) to reduce COVID19 mortality and healthcare demand. <https://spiral.imperial.ac.uk/8443/handle/10044/1/77482>

42. Maier, B. F. & Brockmann, D. Effective containment explains subexponential growth in recent confirmed COVID-19 cases in China. *Science* **368**, 742–746 (2020).
43. Li, Y. et al. The temporal association of introducing and lifting non-pharmaceutical interventions with the time-varying reproduction number (R) of SARS-CoV-2: a modelling study across 131 countries. *Lancet Infect. Dis.* **0** (2020). [www.thelancet.com/infection](http://www.thelancet.com/infection)
44. Dehning, J. et al. Inferring change points in the COVID-19 spreading reveals the effectiveness of interventions. *Science* **369** (2020). <https://arxiv.org/abs/2004.01105><https://doi.org/10.1126/science.abb9789>
45. Levin, A. T. et al. Assessing the age specificity of infection fatality rates for covid-19: Systematic review, meta-analysis, and public policy implications. *medRxiv* (2020). <https://www.medrxiv.org/content/early/2020/10/31/2020.07.23.20160895>
46. van den Driessche, P. & Watmough, J. Reproduction numbers and sub-threshold endemic equilibria for compartmental models of disease transmission. *Math. Biosci.* **180**, 29–48 (2002).
47. Diekmann, O., Heesterbeek, J. A. P. & Metz, J. A. J. On the definition and the computation of the basic reproduction ratio  $R_0$  in models for infectious diseases in heterogeneous populations. *J. Math. Biol.* **28**, 365–382 (1990).
48. Colizza, V. & Vespignani, A. Epidemic modeling in metapopulation systems with heterogeneous coupling pattern: theory and simulations. *J. Theor. Biol.* **251**, 450–467 (2008).
49. Tizzoni, M. et al. On the use of human mobility proxies for modeling epidemics. *PLoS Comput. Biol.* **10**, e1003716 (2014).
50. Pepe, E. et al. Covid-19 outbreak response, a dataset to assess mobility changes in italy following national lockdown. *Sci. Data* **7**, 1–7 (2020).
51. Klein, B. et al. Assessing changes in commuting and individual mobility in major metropolitan areas in the united states during the covid-19 outbreak. <https://www.networkscienceinstitute.org/publications/assessing-changes-in-commuting-and-individual-mobility-in-major-metropolitan-areas-in-the-united-states-during-the-covid-19-outbreak> (2020).
52. Gauvin, L. et al. Socioeconomic determinants of mobility responses during the first wave of covid-19 in italy: from provinces to neighbourhoods. *medRxiv* <https://doi.org/10.1101/2020.11.16.20232413> (2020).
53. Soricetta, A. et al. Mapping internal connectivity through human migration in malaria endemic countries. *Sci. Data* **3**, 1–16 (2016).
54. Cattuto, C. et al. Dynamics of Person-to-Person Interactions from Distributed RFID Sensor Networks. *PLOS ONE* **5**, e11596 (2010).
55. Leung, K., Jit, M., Lau, E. H. Y. & Wu, J. T. Social contact patterns relevant to the spread of respiratory infectious diseases in Hong Kong. *Sci. Rep.* **7**, 7974 (2017).
56. Sanz, J., Xia, C.-Y., Meloni, S. & Moreno, Y. Dynamics of interacting diseases. *Phys. Rev. X* **4**, 041005 (2014).
57. Domenico, M. D., Granell, C., Porter, M. A. & Arenas, A. The physics of spreading processes in multilayer networks. *Nat. Phys.* **12**, 901–906 (2016).

## Acknowledgements

We thank Filippo Privitera and Brennan Lake at Cuebq for providing us free access to the USA Flow Network Data through their Data for Good programme. M.D.D. thanks Stefano Merler and Dirk Brockmann for useful discussions about an early version of this work.

## Author contributions

All authors made a substantial contribution to the writing of this paper. The project was originally conceived by M.D.D. who provided the general guidelines and step by step support. The methods and numerical simulations were developed by P.C. as part of his master thesis. R.G. provided and analyzed the data on human mobility flow.

## Competing interests

The authors declare no competing interests.

## Additional information

Correspondence and requests for materials should be addressed to P.C.

**Peer review information** *Communications Physics* thanks the anonymous reviewers for their contribution to the peer review of this work.

**Reprints and permission information** is available at <http://www.nature.com/reprints>

**Publisher's note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



**Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this license, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2021