Epidemic spreading induced by diversity of agents' mobility

Jie Zhou and Ning Ning Chung
Temasek Laboratories, National University of Singapore, Singapore 117411

Lock Yue Chew
Division of Physics and Applied Physics, School of Physical and Mathematical Sciences, Nanyang Technological University, 21 Nanyang Links, Singapore 637371

Choy Heng Lai
Temasek Laboratories, National University of Singapore, Singapore 117411,
Beijing–Hong Kong–Singapore Joint Centre for Nonlinear and Complex Systems (Singapore), National University of Singapore, Kent Ridge, Singapore 119260, and
Department of Physics, National University of Singapore, Singapore 117542

(Received 18 January 2012; revised manuscript received 29 July 2012; published 27 August 2012)

In this paper, we study the impact of the preference of an individual for public transport on the spread of infectious disease, through a quantity known as the public mobility. Our theoretical and numerical results based on a constructed model reveal that if the average public mobility of the agents is fixed, an increase in the diversity of the agents' public mobility reduces the epidemic threshold, beyond which an enhancement in the rate of infection is observed. Our findings provide an approach to improve the resistance of a society against infectious disease, while preserving the utilization rate of the public transportation system.

DOI: 10.1103/PhysRevE.86.026115 PACS number(s): 89.75.Hc, 05.45.Xt, 89.75.Fb

I. INTRODUCTION

The advent of modern transportation systems has enhanced the mobility of mankind and has increased their range of travel. At the same time, it has intensified the contact between human beings because of the higher human density within transportation systems resulting from a confluence of people within limited physical spaces. The close proximity between travelers provides an opportunity for diseases to spread and it is well known that infectious disease is the main cause of death, disability, as well as social and economic disruption that affects millions of people [1–4]. To stop the proliferation of infectious diseases and their spread, researchers have searched for ways to hinder their diffusion [5–9]. A typical strategy is to adjust the level of human contact through the temporary closure of companies and educational institutes. This strategy, however, comes at a very high price for both society and economy.

In this paper, we have focused our research on epidemic spreading in public transportation systems, with the aim of understanding how the disease spreads within such a system so that mitigating strategies can be determined to reduce its social and economic impact. Human mobility relates to the activity of moving from one point in space to another and can be measured by the frequency and distance of travel [10–14]. Such human movements have been enhanced by public transportation networks which are an indispensable component of the major metropolitan areas of a country. For example, the Mass Rapid Transit (MRT) system in Singapore has a daily load of around 700 000 passengers (i.e., 15% of the total population) [15]. About 90% of Hong Kong citizens rely on public transport facilities for commuting with the main concern being exposure to airborne pollutants within the public vehicles [16]. The large flux of commuters in public transportation systems has typically led to extreme overcrowding, especially during peak hours. The resulting high rate of human contact implies a high rate of transmissibility of infectious diseases. For example, the risk of contracting pulmonary tuberculosis in Peru is higher by a factor of 4.09 for those commuting by minibus compared to those traveling by private transportation [17]. In consequence, commuters tend to avoid public transportation during an epidemic outbreak. They choose either to stay at home or to commute by private transport. The outcome is undesirable: a severe shortage of manpower in the workplace and the possible occurrence of major traffic congestion.

In the past decades, there has been a lot of interest in studying the spread of epidemics within complex networks, which includes (i) the influence of network structure on epidemic spreading [18–22], (ii) the development of immunization strategies [5–7,23–25], and (iii) epidemic spreading in community networks [26–28], in dynamic networks [29–32], and in adaptive networks [33–36]. This has motivated us to construct a model on epidemic spreading that relates to public transportation systems. It is known that commuters have diverse preferences in choosing their mode of travel [37] based on their socioeconomic status. As a result, the frequency of different agents using the public transportation system may differ. In this paper, we shall denote the frequency of utilizing public transport for travel as "public mobility." Individuals with high public mobility use public transports very often, while individuals with low public mobility hardly use the public transportation system. This dichotomy in the usage of public transport prompts the following question: How does the diversity of public mobilities affect the speed of epidemic spreading? The purpose of this work is to give a definite answer to this question.

The structure of our paper is as follows. The details of the model are discussed in Sec. II of this paper. In Sec. III of the paper, we provide a theoretical analysis that enables us to

determine the lower bound of the epidemic threshold. Then, in Sec. IV, we present simulation results which are found to support our theoretical analysis. Finally, we end our paper with a discussion and conclusion in Sec. V.

II. MODEL

In order to gain a better understanding on our model, it is useful to first study a simple model, which can be regarded as a null model that serves the purpose of a benchmark and validity check. In this simple model, a square with length L satisfying the periodic boundary condition is used to represent a society. There are N agents in the square. The positions of the agents are randomly assigned with a uniform distribution and for the sake of simplicity are assumed to be fixed over time. We also assume that an agent only interacts with agents who are located at a distance that is less than r away. In other words, links appear between all pairs of agents whose distance from each other is smaller than r. Since the agents are fixed in their position, the links that are established in this way do not change with time. This null model is simple, and its similar forms have also been adopted in other works [28,29]. In this model, the average degree of the network of agents in the square, $\langle k\rangle$, which is defined as the average number of links that an agent has, is approximately given by $N\pi r^2/L^2$, and the degree distribution of this model satisfies the binomial distribution $p(k) = C_N^k q_L^k (1-q_L)^{N-k}$ with $q_L = N\pi r^2/L^2$. As the ratio of the standard deviation of the distribution to its mean value tends to zero when $N \to \infty$, we expect the degree of the network connection between agents to be homogenous.

In this paper, we use the SIS model to describe the epidemiological process, which is widely adopted to describe infectious diseases [38–41]. In this model, agents can be in either of two distinct states: susceptible or infected. A susceptible agent may become infected if there are infected agents within the interaction radius. Suppose a susceptible agent has k neighbors within its interaction region, of which $k_{\rm inf}$ are infected, and the probability of being infected by each infected neighbor is p; then the probability that an agent becomes infected is $[1-(1-p)^{k_{\rm inf}}]$. At the same time, each infected agent can recover from the disease and becomes susceptible. We assume that this occurs at a rate of μ. When the ratio p/μ is fixed, different pairs of p and μ only affect the definition of the time scale of the disease propagation [41]. Therefore, we can set p and μ to be sufficiently small so as to use the approximation $[1-(1-p)^{k_{\rm inf}}] \to p\,k_{\rm inf}$. In other words, we can maintain the same results (except for the time scale) as long as the ratio p/μ is fixed. This approximation has been widely adopted in the literature (see Refs. [35,42]).

In this paper, we have fixed $r = 0.02$, $p = 0.1$, $\mu = 0.2$. In the simulations, all the averaged results and their standard deviation (which is indicated by the error bars) are obtained from 1000 different realizations, if not otherwise specified.

The epidemic threshold of the null model is determined by the basic reproductive number R_0 with [3,43]

$$R_0 = p\langle k\rangle/\mu. \tag{1}$$

When $R_0 < 1$ the infection dies out in the long run, and when $R_0 > 1$ the infection may spread over the population. This condition leads to a critical average degree $\langle k\rangle^{\rm th} = \frac{\mu}{p}$ and correspondingly the critical number $N^{\rm th}$ for a given set of r, p, μ, and L with $N^{\rm th} = \frac{\mu}{p}\frac{L^2}{\pi r^2}$.

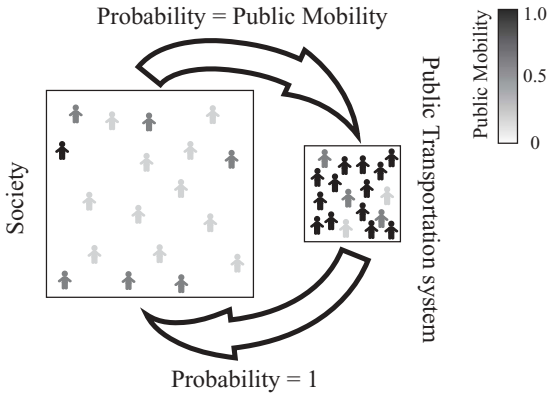

FIG. 1. Schematic illustration of the model. The smaller square represents the public transportation system and the larger square represents the rest of the society. At the beginning of each time step, each agent may transit to the public transportation system with a probability that equals his public mobility. At the end of each time step, all the agents in the public transportation system return to the society. Note that an agent in darker gray possesses a higher public mobility.

Now we are ready to introduce our model which focuses primarily on the public transportation system. In order to study epidemic spreading in the public transportation system, we have separated the society into two parts: the public transportation system (A) and the rest of the society (B) (see Fig. 1). Each part is represented by a square which satisfies periodic boundary condition. The length of square A is L_A and the length of square B is L_B. Since human contact within the public transportation system is typically denser, we have set $L_A \ll L_B$. There are a total of N agents in the society, each (labeled by the index e) with a public mobility (denoted as "PM" in the following) of m_e that ranges between 0 and 1. Our model begins by assigning a random position for each of the N agents in square B. At the beginning of each time step, agent e either transits to square A with probability m_e and then chooses a random position there to stay, or remains in square B at the originally assigned position. At the end of each time step, all the agents that transit to square A return to their original position in square B and the whole system prepares for the next time step.

Similar to the null model, each agent has a contact radius of r and a link between two agents is formed whenever they are within this radius. The average degree of agents in square A (B) is $\langle k\rangle_A = N_A\pi r^2/L_A^2$ ($\langle k\rangle_B = N_B\pi r^2/L_B^2$), where N_A (N_B) is the number of agents in the square A (B).

III. THEORETICAL ANALYSIS

In this section, we provide a theoretical analysis on the epidemic threshold of our model. Suppose $\rho(m)$ is the fraction of agents with PM m such that $\int \rho(m)dm = 1$. Then, the average PM $\overline{m} = \int \rho(m)m dm$ and the second moment $D = \int \rho(m)m^2 dm$. Let us denote $i(m,t)$ as the fraction of infected nodes with PM m at time step t. Since $i(m,t)$ represents the fraction of infected agents in both squares A and B at time t,

we expect the evolution of $i(m,t)$ to consist of two parts:

$$i(m,t+1) = \Pi_A(m,t+1) + \Pi_B(m,t+1). \tag{2}$$

Here, $\Pi_A(m,t+1)$ denotes the fraction of infected agents with PM m at time $t+1$ as a result of having visited square A at time t. It can be expressed as follows:

$$\Pi_A(m,t+1) = i(m,t)\,m + [\rho(m) - i(m,t)]\,m\,p\bar{k}_{\rm inf}^{(A)}(t) - \mu i(m,t)\,m, \tag{3}$$

where $\bar{k}_{\rm inf}^{(A)}(t)$ is the average number of infected neighbors of an agent in square A at time t and is defined by

$$\bar{k}_{\rm inf}^{(A)}(t) = \langle k\rangle_A \frac{\int i(m',t)m'dm'}{\int \rho(m')m'dm'}. \tag{4}$$

The first term on the right-hand side of Eq. (3) denotes the fraction of infected agents with PM m that have moved to square A at time t. The second term on the right-hand side denotes the fraction of susceptible agents with PM m who are infected at time t due to their transit to square A. The third term on the right-hand side represents the fraction of infected agents with PM m who have recovered from the infected state and are now in the susceptible state, as they traverse to square A at time step t. On the other hand, $\Pi_B(m,t+1)$ gives the fraction of infected agents with PM m at time $t+1$ in lieu of having remained in square B at time t. Similar to $\Pi_A(m,t+1)$, $\Pi_B(m,t+1)$ takes the following form:

$$\Pi_B(m,t+1) = i(m,t)\,(1-m) + [\rho(m) - i(m,t)]\,(1-m)\,p\bar{k}_{\rm inf}^{(B)}(t) - \mu i(m,t)\,(1-m), \tag{5}$$

where $\bar{k}_{\rm inf}^{(B)}(t)$ is the average number of infected neighbors in contact with an agent in square B at time t and is defined by

$$\bar{k}_{\rm inf}^{(B)}(t) = \langle k\rangle_B \frac{\int i(m',t)(1-m')dm'}{1-\int \rho(m')m'dm'}. \tag{6}$$

In the steady state, we expect $i(m,t+1) = i(m,t) = i^*(m)$. Thus, we have

$$\frac{\mu}{p} i^*(m) = [\rho(m) - i^*(m)]\,m\,\bar{k}_{\rm inf}^{(A)}(t) + [\rho(m) - i^*(m)]\,\bar{k}_{\rm inf}^{(B)}(t) - [\rho(m) - i^*(m)]\,m\,\bar{k}_{\rm inf}^{(B)}(t). \tag{7}$$

Multiplying both sides of the equation by m and then integrating throughout with respect to m, we obtain

$$\Omega\left[\frac{\langle k\rangle_A}{\overline{m}} + \frac{\langle k\rangle_B}{1-\overline{m}}(\lambda-1)(\eta-1)\right] = D\left[\frac{\langle k\rangle_A}{\overline{m}} - \frac{\langle k\rangle_B}{1-\overline{m}}(\lambda-1)\right] + \frac{\langle k\rangle_B}{1-\overline{m}}\overline{m}(\lambda-1) - \frac{\mu}{p}, \tag{8}$$

where $\lambda = i^*/\Theta$ and $\eta = \Theta/\Omega$ for $i^* = \int i^*(m)dm$, $\Theta = \int i^*(m)mdm$, and $\Omega = \int i^*(m)m^2dm$. It is easy to see that $\lambda - 1 \geqslant 0$ and $\eta - 1 \geqslant 0$ when $m \in [0,1]$. Thus, the terms in the square bracket on the left-hand side of Eq. (8) are positive. In order for the fraction of infected agents to be nonzero, i.e., $i^* > 0$ and $\Omega > 0$, the right-hand side of Eq. (8) has to be positive. Hence,

$$D\left[\frac{\langle k\rangle_A}{\overline{m}} - \frac{\langle k\rangle_B}{1-\overline{m}}(\lambda-1)\right] + \frac{\langle k\rangle_B}{1-\overline{m}}\overline{m}(\lambda-1) > \frac{\mu}{p}. \tag{9}$$

Equation (9) indicates the presence of a lower bound for D, which is

$$\underline{D} = \frac{\frac{\mu}{p} - \frac{\langle k\rangle_B}{1-\overline{m}}\overline{m}(\lambda-1)}{\frac{\langle k\rangle_A}{\overline{m}} - \frac{\langle k\rangle_B}{1-\overline{m}}(\lambda-1)}. \tag{10}$$

When $\langle k\rangle_B \ll \langle k\rangle_A$ and $\langle k\rangle_B$ is small, Eq. (10) can be approximated by

$$\underline{D} = \frac{\mu}{p}\frac{\overline{m}}{\langle k\rangle_A}. \tag{11}$$

Since $\langle k\rangle_A = \frac{N_A\pi r^2}{L_A^2} = \frac{\overline{m}N\pi r^2}{L_A^2}$, $\underline{D}$ can also be expressed in terms of N and L_A. This expression has the implication that the value of $\underline{D}$ remains unchanged as we scale the variables N and L_A^2 by the same factor.

Let us next consider the case where $\langle k\rangle_B$ is not neglected in Eq. (10). Since human contact in the public transportation system is typically denser, we anticipate square A to dominate the infection process. Therefore, agents that transit frequently to square A have a larger probability of being infected. Hence, we expect $i^*(m) \sim \rho(m)m$, and therefore $\lambda = i^*/\Theta = \int \rho(m)mdm/\int \rho(m)m^2dm = \overline{m}/D$. Thus, Eq. (10) becomes

$$\underline{D} = \frac{\frac{\mu}{p} - \frac{\langle k\rangle_B}{1-\overline{m}}\overline{m}\left(\frac{\overline{m}}{\underline{D}} - 1\right)}{\frac{\langle k\rangle_A}{\overline{m}} - \frac{\langle k\rangle_B}{1-\overline{m}}\left(\frac{\overline{m}}{\underline{D}} - 1\right)}. \tag{12}$$

Solving Eq. (12) for $\underline{D}$, we obtain

$$\underline{D} = \frac{\sqrt{\left(\frac{\mu}{p} + 2\overline{m}\frac{\langle k\rangle_B}{1-\overline{m}}\right)^2 - 4\overline{m}^2\frac{\langle k\rangle_B}{1-\overline{m}}\left(\frac{\langle k\rangle_A}{\overline{m}} + \frac{\langle k\rangle_B}{1-\overline{m}}\right)}}{2\left(\frac{\langle k\rangle_A}{\overline{m}} + \frac{\langle k\rangle_B}{1-\overline{m}}\right)} + \frac{\left(\frac{\mu}{p} + 2\overline{m}\frac{\langle k\rangle_B}{1-\overline{m}}\right)}{2\left(\frac{\langle k\rangle_A}{\overline{m}} + \frac{\langle k\rangle_B}{1-\overline{m}}\right)}. \tag{13}$$

By using $\langle k\rangle_A = \frac{\overline{m}N\pi r^2}{L_A^2}$ and $\langle k\rangle_B = \frac{(1-\overline{m})N\pi r^2}{L_B^2}$, we can also express $\underline{D}$ via N, L_A, and L_B.

Similar to Eq. (11), $\underline{D}$ is found here to remain invariant when N, L_A^2, and L_B^2 are varied by the same scaling factor if $\overline{m}$ is fixed. In fact, Eqs. (11) and (13) allow us to obtain the lower bound of the variance of PM:

$$\underline{\sigma^2} = \underline{D} - \overline{m}^2. \tag{14}$$

IV. SIMULATION RESULTS

Before we show the influence of the diversity of PM on epidemic spreading, let us first study the simple case when all the agents have the same PM value m. We suppose the system contains N agents. When the PM of all the agents is m, the number of agents in square A in each time step is $N_A = mN$, while that in square B is $N_B = (1-m)N$. Therefore, the long-time average of the degree of all the agents is $\langle k\rangle = m\,\frac{mN\pi r^2}{L_A^2} + (1-m)\times\frac{(1-m)N\pi r^2}{L_B^2}$, which is approximately $m^2\frac{N\pi r^2}{L_A^2}$ when

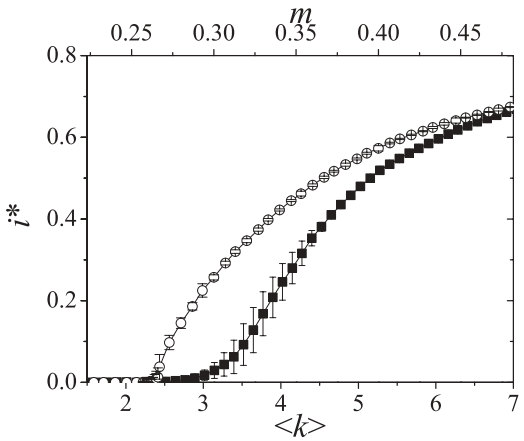

FIG. 2. The open circle symbols shows the fraction of infected agents i^* at the steady state versus the average degree $\langle k\rangle$ (the lower abscissa) and the average PM m (the upper abscissa) for the case that the PM of all the agents are equal to m, with $N = 1500$, $L_A = 0.25$, and $L_B = 10$. The solid square symbols illustrate the case for the null model for the sake of comparison, with $L = 1$.

$L_A \ll L_B$. By adopting Eq. (1), we can obtain the threshold m^{th} as follows:

$$m^{\text{th}} = \sqrt{\frac{\mu}{p}\frac{L_A^2}{N\pi r^2}}. \tag{15}$$

Note that above this threshold, the system may become endemic.

Figure 2 shows the fraction of infected agents i^* at the steady state versus the values m and $\langle k\rangle$ for the case when the PM of all the agents is m (circle symbols) and the null model (square symbols), respectively. By using Eqs. (1) and (15), we have $\langle k\rangle^{\text{th}} = 2$ and $m^{\text{th}} = 0.26$ (detailed parameters are indicated in the caption of the figure). It shows that our theoretical estimates on $\langle k\rangle^{\text{th}}$ and m^{th} are in accord with the simulation results. Moreover, it shows that when $\langle k\rangle > \langle k\rangle^{\text{th}}$, the value of the circle symbols may be much larger than that of the square symbols, which indicates that the extent of the epidemic prevalence is strongly enhanced by the public transportation system. In the following, we shall show that a modification to this threshold behavior can occur when we take the diversity of the agents' PM into consideration.

Let us begin by considering the case in which all the agents belong to either of two groups, G_1 or G_2, with all the agents in each group having the same PM. The PM of the agents in G_1 (G_2) is m_1 (m_2), and the size of the group is N_1 (N_2). Thus, the size of the system $N = N_1 + N_2$, the average PM $\overline{m} = (N_1 m_1 + N_2 m_2)/N$, and the variance of the PM $\sigma^2 = (N_1 m_1^2 + N_2 m_2^2)/N - \overline{m}^2$. At each time step, the expected number of agents in G_1 (G_2) transiting to square A is equal to $m_1 N_1 (m_2 N_2)$. Hence, on average, square A contains $N_A = m_1 N_1 + m_2 N_2 = \overline{m} N$ agents, and square B contains $N_B = (1 - \overline{m})N$ agents. Therefore, $\langle k\rangle_A$ and $\langle k\rangle_B$ can be expressed as $\langle k\rangle_A = \overline{m} N \pi r^2 / L_A^2$ and $\langle k\rangle_B = (1 - \overline{m}) N \pi r^2 / L_B^2$, respectively. Moreover, after the process of time averaging, the degree of agent e with PM m_e is $k_e = m_e \langle k\rangle_A + (1 - m_e)\langle k\rangle_B$.

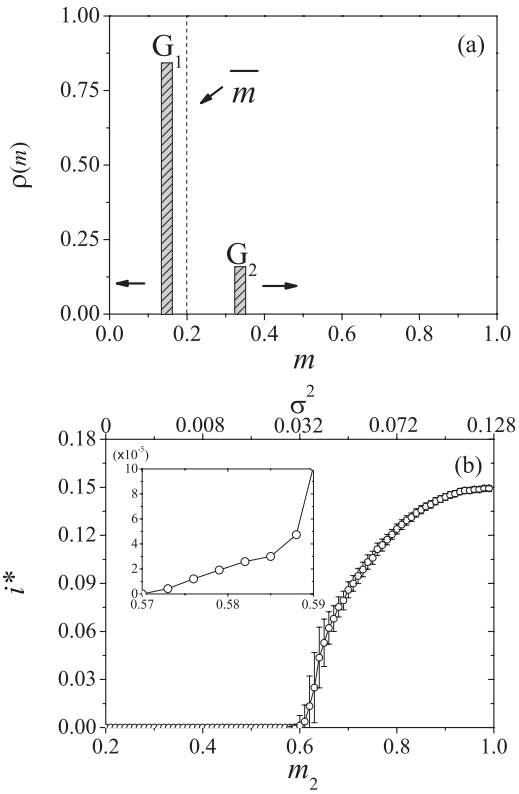

FIG. 3. (a) The fraction of agents ρ against the PM m. The two bars in the figure indicate the presence of only two groups, with each group having different PM. This figure illustrates how the standard deviation of the PM is tuned without changing the average PM, which is $\overline{m} = 0.2$, as indicated by the dashed line. (b) The fraction of infected agents i^* at the steady state against the PM m_2 and σ^2. The inset shows the detailed behavior of i^* near the zone of transition. In this case, $N_1 = 1250$ and $N_2 = 250$. Note that all the other parameters take the same value as those employed in Fig. 2.

In order to demonstrate the effect of diversity in PM on epidemic spreading, we first set $m_1 = m_2$. After that, we decrease m_1 and increase m_2 such that $\overline{m}$ remains unchanged. This operation increases σ^2 from 0 without changing $\overline{m}$. Figure 3(a) illustrates the manner in which the PMs of the two groups are tuned. Figure 3(b) shows the fraction of infected agents i^* at the steady state as a function of m_2. When $m_1 = m_2 = 0.2$ (i.e., $\sigma^2 = 0$), the system is in a disease-free state. When m_2 exceeds the threshold, $m_2^{\text{th}} \sim 0.568$ which is determined from Eq. (11) (note that the rest of the parameters are indicated in the caption of the figure), the system becomes endemic. Compared with the results shown in Fig. 2 and Eq. (15), this example shows that the diversity of the PM can induce epidemic spreading, even when $\overline{m}$ is smaller than the m^{th} in Eq. (15). Since in this case $L_A \ll L_B$ which makes $\langle k\rangle_A \gg \langle k\rangle_B$, we can use Eq. (11) to calculate the lower bound of the variance $\underline{\sigma^2}$. Our calculation gives $\underline{D} = 0.067$, and correspondingly $\underline{\sigma^2} = \underline{D} - \overline{m}^2 = 0.027$. This figure shows that our simulation results are in accord with the theoretical estimate on the epidemic threshold, above which a finite fraction of the infected agents is found to exist. Thus, for a fixed $\overline{m}$, there exists a threshold for the variance of the PM, exceeding which the epidemic spreads and the system becomes

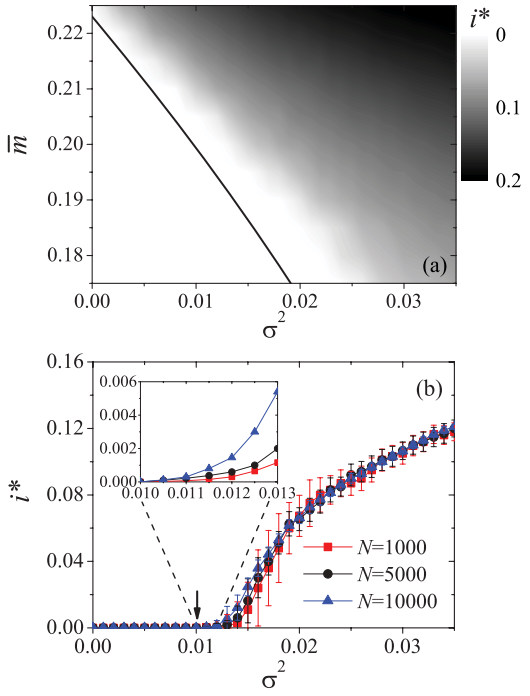

FIG. 4. (Color online) (a) A gray-scale plot on the fraction of infected agents i^* at the steady state in the $\overline{m}$-σ^2 plane. The sizes of the error bars are small and hence are not shown here. The parameters employed are $\alpha = 3$, $N = 2000$, $G = 20$, $L_A = 0.25$, and $L_B = 5$. (b) The fraction of infected agents i^* at the steady state as a function of σ^2 when $\overline{m} = 0.2$ for $N = 1000$, 5000, and 10 000. In all the three cases, N/L_A^2 and N/L_B^2 are maintained at the same value. That is (i) $N = 1000$, $L_A \simeq 0.177$, $L_B \simeq 3.54$; (ii) $N = 5000$, $L_A \simeq 0.4$, $L_B \simeq 7.9$; (iii) $N = 10\,000$, $L_A \simeq 0.56$, $L_B \simeq 11.2$. As a result, $\langle k \rangle_A \simeq 8$ and $\langle k \rangle_B \simeq 0.08$ for all three cases. The inset shows the zoom in results for the range $\sigma^2 \in [0.010, 0.013]$. The results in this panel are obtained from 5000 different realizations. Note that the values of the other parameters are the same as those used in (a).

endemic. Beyond the threshold, the fraction of infected agents increases as σ^2 increases.

In a more general setting, the PM of the population may follow an arbitrary distribution. In order to study our model in this more general situation, we first need to develop an approach to assign PMs to agents following a given distribution, with $\overline{m}$ and σ^2 tunable in this method. The details of this method are presented in the Appendix.

With the observation that the commuters' behavior [10] and their traveling properties such as distance and time interval between journeys are found to be characterized by a power-law distribution [11–14], we here assume that PM follows the distribution $\rho(m) \sim m^{-\alpha}$. [Note that our conclusions do not rely on any particular form of $\rho(m)$]. By utilizing the method introduced in the Appendix, we can tune $\overline{m}$ and σ^2 without changing the form of $\rho(m)$. The results on the fraction of infected agents i^* at the steady state in the $\overline{m}$-σ^2 plane are shown in Fig. 4(a), where the exponent of the power-law distribution is $\alpha = 3$. In this figure darker gray levels indicate a larger fraction of infected agents. We observe that for a given σ^2, i^* increases with an increase in $\overline{m}$, which can be understood as follows. A larger $\overline{m}$ means the transit of a larger number of agents to square A. This implies a higher average degree for all the agents since $\langle k \rangle_A$ is greater than $\langle k \rangle_B$. The outcome is an increase in the contacts between humans within the society. Thus, epidemic threshold reduces and disease spreads more easily. In particular, when the threshold for σ^2 drops to zero, the system can still be endemic as long as $\overline{m}$ is large enough. The presence of a threshold here for σ^2 is similar to the two-mobility group model that we have discussed earlier. Just like the two-mobility group model, we observe that as σ^2 increases beyond a threshold, the fraction of infected agents increases as the variance increases. Note that the solid line in the figure is obtained from Eqs. (13) and (14) which denotes our theoretical estimate of the epidemic threshold. We can see that the theoretical estimation conforms with our numerical simulation results. It is important to note that the theoretical analysis above indicates that $D \sim \Theta$ and through Eq. (4) reveals that D is proportional to the probability that a link is infected. In other words, for a fixed $\overline{m}$, a larger σ^2 [see Eq. (14)] implies a larger probability that a link is infected. This explains the observation as duly shown in Figs. 3 and 4 that increasing σ^2, i.e., the diversity of PM, invariably increases the fraction of infected agents. It thus clarifies our inference that diversity in PM has the effect of reducing the threshold of an epidemic outbreak. Figure 4(b) shows the fraction of infected agents i^* at the steady state as a function of σ^2 when $\overline{m} = 0.2$ for $N = 1000$, 5000, and 10 000. In all three cases, we have kept the value of N/L_A^2 and N/L_B^2 constant. We observe that $\underline{\sigma^2} \simeq 0.01$ for all three cases as indicated by the black arrow, which is consistent with the analytical results obtained from Eqs. (13) and (14). Moreover, when $\sigma^2 > \underline{\sigma^2}$, i^* is observed to have very similar monotonically increasing behavior for the different number of agents N.

V. DISCUSSION AND CONCLUSION

In summary, we have studied the effects of diversity in public mobility on epidemic spreading by proposing a model which separates a society into two parts: the public transportation system and the rest of the society. In our model, we have defined public mobility as the probability of an agent in the society who opts to take public transport at each time step. Our results show that a larger diversity in public mobility gives rise to a smaller epidemic threshold. Taking into account the inevitable diversity in socioeconomic status among individuals within a population, we have come to the conclusion that if we are able to control the diversity in human behavior, we would be able to enhance the resistance of a society against the onslaught of a pending epidemic. Our results show that this can be achieved without reducing the average public mobility of a society by encouraging the population to use both public and private transport with uniformity and without bias. For example, commuters with low public mobility are persuaded to take public transport more regularly while commuters having high public mobility are urged to travel in private transport with greater frequency. In this way, epidemic spreading can be slowed down without causing any traffic congestion as well as any disturbance to the proper functioning of the public transportation system.

ACKNOWLEDGMENTS

This work is supported by the Defense Science and Technology Agency of Singapore under project agreement POD0613356.

APPENDIX: A METHOD OF TUNING $\overline{m}$ AND σ^2

In this Appendix we report the details of a method of tuning $\overline{m}$ and σ^2, while maintaining the PM of the population according to an arbitrary distribution given by $\rho(m)$ as follows.

First, we distribute all the N agents evenly into G groups so that there are N/G agents in each group. Agents in the same group have the same PM, while agents in different groups may have different PMs. The PM of agents in group j is m_j, with $j = 1,\ldots,G$. The upper and lower bounds of the PM are $m_{\min} = m_0$ and $m_{\max} = m_G$, respectively. We define $F(m)$ to be the primitive function of $\rho(m)$, such that $F(m) = \int_{m_0}^{m} \rho(m')dm'$. Then, we assign PM to the agents according to the following recurrent relations:

$$F(m_j) - F(m_{j-1}) = \frac{1}{G}[F(m_G) - F(m_0)], \quad \text{for } j = 1,\ldots,G. \tag{A1}$$

Given m_0 and m_G, each m_j can be obtained by solving Eq. (A1) from $j = 1$ to $j = G$, from which $\overline{m}$ and σ^2 can be determined. Since $\rho(m)$ is positive for $m \in [0,1]$, m_j can only increase monotonically with j. The meaning of Eq. (A1) can be understood in the following way. Based on our definition, $\rho(m) = [F(m_j) - F(m_{j-1})]/\Delta m$ with $\Delta m = m_j - m_{j-1} \to 0$. This implies that $\rho(m) = C/(G\Delta m)$, where C is a constant equals to $F(m_G) - F(m_0)$, according to our construction. Then, for a well defined $\rho(m)$, we would expect $G \to \infty$ as $\Delta m \to 0$. In other words, the distribution of PM generated by our approach becomes accurate and tends towards the distribution $\rho(m)$ as $G \to \infty$.

Furthermore, Eq. (A1) can be easily extended to the more general situation of each group having a different number of agents. Suppose the size of group j is n_j; then Eq. (A1) can be generalized to

$$F(m_j) - F(m_{j-1}) = \frac{n_j}{N}[F(m_G) - F(m_0)], \quad \text{for } j = 1,\ldots,G. \tag{A2}$$

However, we shall restrict our investigation to the condition of same group size as we explore the effect of the diversity of PM.

Suppose PM follows the distribution $\rho(m) \sim m^{-\alpha}$. [Note that our formulation based on Eqs. (A1) and (A2) allows $\rho(m)$ to take any generic form.] Then, by means of Eq. (A1), we obtain the PM of agents in group j as follows:

$$m_j = \left(\frac{j}{G} m_G^{1-\alpha} + \frac{G-j}{G} m_0^{1-\alpha}\right)^{1/(1-\alpha)}. \tag{A3}$$

The average PM is given by

$$\overline{m} = \frac{1}{G}\left(\frac{m_G^{1-\alpha} - m_0^{1-\alpha}}{G}\right)^{1/(1-\alpha)} \zeta\left(\frac{1}{1-\alpha}, \frac{G m_0^{1-\alpha}}{m_G^{1-\alpha} - m_0^{1-\alpha}}, G\right), \tag{A4}$$

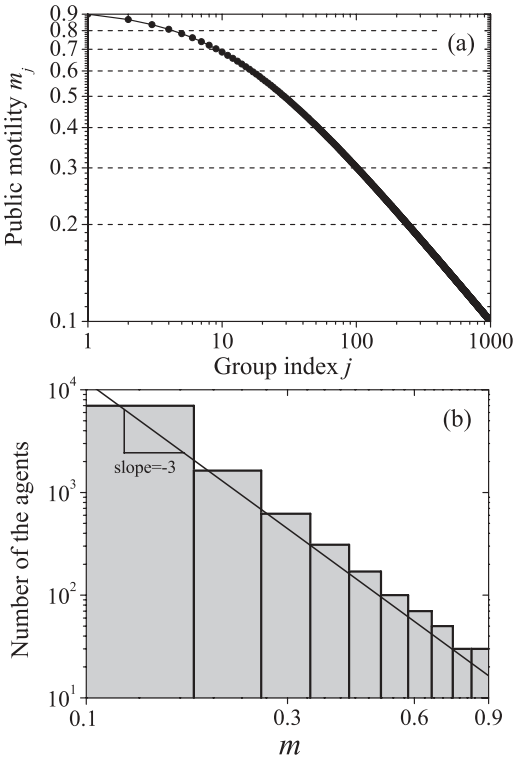

FIG. 5. (a) The relation between the group index j and the corresponding PM m_j is plotted in log-log scale, where $\alpha = 3$, $m_0 = 0.1$, $m_G = 0.9$, $G = 1000$, and $N = 10\,000$. The size of each group is $N/G = 10$. Dashed lines are plotted for reference. (b) Histogram of the number of agents with PM within the range $[m_0, m_G]$. By separating the range $[m_0, m_G]$ into 10 parts, the width of each bar is 0.8. Note that an agent with PM within the abscissa of a particular bar is counted towards the height of that bar. The plot is in log-log scale. We have plotted a straight line with a slope of -3 to serve as a guide for reference.

where $\zeta(\beta, p, N) = \sum_{j=1}^{N} (p + j)^{\beta}$ is the truncated form of the generalized ζ function [44]. Similarly, the second moment D of the distribution can be obtained from Eq. (A3) as follows:

$$D = \frac{1}{G}\left(\frac{m_G^{1-\alpha} - m_0^{1-\alpha}}{G}\right)^{2/(1-\alpha)} \zeta\left(\frac{2}{1-\alpha}, \frac{G m_0^{1-\alpha}}{m_G^{1-\alpha} - m_0^{1-\alpha}}, G\right). \tag{A5}$$

We observe that $\overline{m}$ and $\sigma^2 = D - \overline{m}^2$ are functions of m_0, m_G, and G. Hence, we can adjust the distribution of PM by varying the values of m_0 and m_G so as to obtain different $\overline{m}$ and σ^2 for a given G. Figure 5(a) shows the relation of the group index j and the corresponding PM m_j of a power-law distribution with $\alpha = 3$. In this case, we have set $N = 10\,000$, $G = 1000$, $m_0 = 0.1$, $m_G = 0.9$, $\overline{m} = 0.18$, and $\sigma^2 = 0.012$. In this figure, we observe that about 1% of the total number of groups has PM larger than 0.7, while about 90% of the groups has PM smaller than 0.3, which manifests a strong heterogeneity in the PM. Figure 5(b) shows the corresponding histogram of PM $\rho(m)$ of the generated sample. We have performed a maximum likelihood estimate of the exponent for the distribution obtained from Eq. (A3). The value of the most likely exponent is 3.08 ± 0.02 [44], which shows a good fit to the value of the target exponent, which is 3.

The slight difference between the actual and target exponent results from G being finite, as was pointed out above for a nontrivial situation; i.e., $\alpha \neq 0$. It is only when $G \to \infty$ that the actual and target exponent have a perfect match. We have plotted a straight line with a slope of -3 to serve as a guide for reference. These results demonstrate the effectiveness of the method. By tuning m_0 and m_G, we can obtain different values of $\overline{m}$ and σ^2. We note that our purpose here is not only to assign PM to the agents following a certain distribution but more importantly to find a way to adjust $\overline{m}$ and σ^2; hence here $\rho(m) \sim m^{-\alpha}$ could also serve as a simple and effective auxiliary function for adjusting $\overline{m}$ and σ^2. In consequence, as long as $\overline{m}$ and σ^2 are obtained correctly, we do not expect the slight difference between the actual and target exponent to affect our conclusions. We also note that when a power-law distribution is unbounded, the mean and variance of the distribution can be infinite if the exponent satisfies certain conditions [44]. Under these circumstances, any mean and variance obtained from a set of samples of such a distribution are not meaningful because the fluctuation of these quantities can become exceedingly large [44]. However, the PM in our model is bounded within $[0,1]$. Therefore, the mean and variance of the distribution obtained from Eq. (A3) are finite. In this situation, we expect the sample mean and sample variance to converge to the population mean and population variance of the distribution respectively, as the number of samples tends to infinity. Thus, the mean and variance of the distribution of our model obtained from Eq. (A3) are reliable.

[1] See http://www.who.int.
[2] N. Barquet and P. Domingo, Ann. Intern. Med. **127**, 635 (1997).
[3] R. M. Anderson and R. M. May, *Infectious Diseases of Humans* (Oxford University Press, New York, 1991).
[4] N. T. J. Bailey, *The Mathematical Theory of Infectious Diseases*, 2nd ed. (Berlin, Springer Verlag, 1993).
[5] Z. Dezsö and A.-L. Barabási, Phys. Rev. E **65**, 055103 (2002).
[6] Y. Chen, G. Paul, S. Havlin, F. Liljeros, and H. E. Stanley, Phys. Rev. Lett. **101**, 058701 (2008).
[7] Leah B. Shaw and Ira B. Schwartz, Phys. Rev. E **81**, 046120 (2010).
[8] See http://www.moh.gov.sg.
[9] M. Salathé and James H. Jones, PLoS **6**, e01000736 (2010).
[10] N. Lathia and L. Capra, in *Proceedings of the 13th ACM International Conference on Ubiquitous Computing* (ACM, Beijing, 2011); *Proceedings of the ACM SIGKDD 2011 Conference on Knowledge Discovery and Data Mining* (ACM, San Diego, 2011).
[11] D. Brockmann, L. Hufnagel, and T. Geisel, Nature (London) **439**, 462 (2005).
[12] M. C. González, C. A. Hidalgo, and A.-L. Barabási, Nature (London) **453**, 779 (2008).
[13] D. Balcan, V. Colizza, B. Gonçalves, H. Hu, J. J. Ramasco, and A. Vespignani, Proc. Natl. Acad. Sci. USA **106**, 21484 (2009).
[14] C. Song, T. Koren, P. Wang, and A.-L. Barabási, Nat. Phys. **6**, 818 (2010).
[15] X. Fu, S. Lim, L. Wang, G. Lee, S. Ma, L. Wong, and G. Xiao, *Proceedings of IEEE Swarm Intelligence Symposium* (IEEE, 2009).
[16] L. Y. Chan, W. L. Lau, S. C. Lee, and C. Y. Chan, Atmos. Environ. **36**, 3363 (2002).
[17] O. J. Horna-Campos, H. J. Sánchez-Pérez, I. Sánchez, A. Bedoya, and M. Martín, Emerg. Infect. Dis. **13**, 1491 (2007).
[18] M. Kuperman and G. Abramson, Phys. Rev. Lett. **86**, 2909 (2001).
[19] R. Pastor-Satorras and A. Vespignani, Phys. Rev. Lett. **86**, 3200 (2001).
[20] M. Boguna, R. Pastor-Satorras, and A. Vespignani, Phys. Rev. Lett. **90**, 028701 (2003).
[21] M. Barthelemy, A. Barrat, R. Pastor-Satorras, and A. Vespignani, Phys. Rev. Lett. **92**, 178701 (2004).
[22] G. Yan, T. Zhou, J. Wang, Z. Fu, and B. Wang, Chin. Phys. Lett. **22**, 510 (2005).
[23] D. H. Zanette and M. Kuperman, Physica A **309**, 445 (2002).
[24] R. Cohen, S. Havlin, and D. ben-Avraham, Phys. Rev. Lett. **91**, 247901 (2003).
[25] L. K. Gallos, F. Liljeros, P. Argyrakis, A. Bunde, and S. Havlin, Phys. Rev. E **75**, 045104 (2007).
[26] Z. Liu and B. Hu, Europhys. Lett. **72**, 315 (2005).
[27] Y. Zhou, Z. Liu, and J. Zhou, Chin. Phys. Lett. **24**, 581 (2007).
[28] J. Zhou and Z. Liu, Physica A **388**, 1228 (2009).
[29] M. Frasca, A. Buscarino, A. Rizzo, L. Fortuna, and S. Boccaletti, Phys. Rev. E **74**, 036110 (2006).
[30] N. H. Fefferman and K. L. Ng, Phys. Rev. E **76**, 031919 (2007).
[31] E. Volz and L. A. Meyers, Proc. R. Soc. London, Ser. B **274**, 2925 (2007).
[32] A. Buscarino, L. Fortuna, M. Frasca, and V. Latora, Europhys. Lett. **82**, 38002 (2008).
[33] T. Gross, C. J. Dommar D'Lima, and B. Blasius, Phys. Rev. Lett. **96**, 208701 (2006); T. Gross and I. G. Kevrekidis, Europhys. Lett. **82**, 38004 (2008).
[34] L. B. Shaw and I. B. Schwartz, Phys. Rev. E **77**, 066101 (2008).
[35] D. H. Zanette and S. Risau-Gusman, J. Biol. Phys. **34**, 135 (2008).
[36] J. Zhou, G. Xiao, S. A. Cheong, X. Fu, L. Wong, S. Ma, and T. H. Cheng, Phys. Rev. E **85**, 036107 (2012).
[37] E. Davidov, P. Schmidt, and S. Bamberg, Euro. Soc. Rev. **19**, 267 (2003).
[38] G. H. Weiss and M. Dishon, Math. Biosci. **11**, 261 (1971).
[39] N. T. J. Bailey, *The Mathematical Theory of Infectious Disease and its Applications*, 2nd ed. (Charles Griffn, Oxford, 1975).
[40] J. D. Murray, *Mathematical Biology* (Springer-Verlag, Berlin, 1993).
[41] O. Diekmann and J. Heesterbeek, *Mathematical Epidemiology of Infectious Diseases: Model Building, Analysis, and Interpretation* (Wiley, New York, 2000).
[42] T. Zhou, J.-G. Liu, W.-J. Bai, G. Chen, and B.-H. Wang, Phys. Rev. E **74**, 056109 (2006).
[43] C. Fraser *et al.*, Science **324**, 1557 (2009).
[44] M. E. J. Newman, Contemp. Phys. **46**, 323 (2005).