

Individual dynamics and local heterogeneity provide a microscopic view of the epidemic spreading

Youyuan Zhu(朱友源), Ruizhe Shen(沈瑞哲), Hao Dong(董昊), and Wei Wang(王炜)

Citation: Chin. Phys. B, 2024, 33 (**5**): 058301. DOI: 10.1088/1674-1056/ad1a90
Journal homepage: http://cpb.iphy.ac.cn; http://iopscience.iop.org/cpb

What follows is a list of articles you may be interested in

Source localization in signed networks with effective distance

Zhi-Wei Ma(马志伟), Lei Sun(孙蕾), Zhi-Guo Ding(丁智国), Yi-Zhen Huang(黄宜真), and Zhao-Long Hu(胡兆龙)
Chin. Phys. B, 2024, 33 (**2**): 028902. DOI: 10.1088/1674-1056/ad1482

Analysis of radiation diffusion of COVID-19 driven by social attributes

Fuzhong Nian(年福忠), Xiaochen Yang(杨晓晨), and Yayong Shi(师亚勇)
Chin. Phys. B, 2024, 33 (**1**): 018904. DOI: 10.1088/1674-1056/acd7d1

Identifying influential spreaders in social networks: A two-stage quantum-behaved particle swarm optimization with Lévy flight

Pengli Lu(卢鹏丽), Jimao Lan(揽继茂), Jianxin Tang(唐建新), Li Zhang(张莉), Shihui Song(宋仕辉), and Hongyu Zhu(朱虹羽)
Chin. Phys. B, 2024, 33 (**1**): 018901. DOI: 10.1088/1674-1056/acd3e0

Robustness of community networks against cascading failures with heterogeneous redistribution strategies

Bo Song(宋波), Hui-Ming Wu(吴惠明), Yu-Rong Song(宋玉蓉), Guo-Ping Jiang(蒋国平),Ling-Ling Xia(夏玲玲), and Xu Wang(王旭)
Chin. Phys. B, 2023, 32 (**9**): 098905. DOI: 10.1088/1674-1056/acd9c3

Identification of key recovering node for spatial networks

Zijian Yan(严子健), Yongxiang Xia(夏永祥), Lijun Guo(郭丽君), Lingzhe Zhu(祝令哲), Yuanyuan Liang(梁圆圆), and Haicheng Tu(涂海程)
Chin. Phys. B, 2023, 32 (**6**): 068901. DOI: 10.1088/1674-1056/acb75f

---

Individual dynamics and local heterogeneity provide a microscopic view of the epidemic spreading

Youyuan Zhu(朱友源)[1,2], Ruizhe Shen(沈瑞哲)[1,2], Hao Dong(董昊)[1,3,4,†], and Wei Wang(王炜)[2,3,‡]

[1] *Kuang Yaming Honors School, Nanjing University, Nanjing* 210023, *China*
[2] *Collaborative Innovation Center of Advanced Microstructures, National Laboratory of Solid State Microstructure, School of Physics, Nanjing University, Nanjing* 210093, *China*
[3] *Institute for Brain Sciences, Nanjing University, Nanjing* 210023, *China*
[4] *State Key Laboratory of Analytical Chemistry for Life Science, Nanjing University, Nanjing* 210023, *China*

(Received 1 November 2023; revised manuscript received 7 December 2023; accepted manuscript online 4 January 2024)

The COVID-19 pandemic has caused severe global disasters, highlighting the importance of understanding the details and trends of epidemic transmission in order to introduce efficient intervention measures. While the widely used deterministic compartmental models have qualitatively presented continuous "analytical" insight and captured some transmission features, their treatment usually lacks spatiotemporal variation. Here, we propose a stochastic individual dynamical (SID) model to mimic the random and heterogeneous nature of epidemic propagation. The SID model provides a unifying framework for representing the spatiotemporal variations of epidemic development by tracking the movements of each individual. Using this model, we reproduce the infection curves for COVID-19 cases in different areas globally and find the local dynamics and heterogeneity at the individual level that affect the disease outbreak. The macroscopic trend of virus spreading is clearly illustrated from the microscopic perspective, enabling a quantitative assessment of different interventions. Seemingly, this model is also applicable to studying stochastic processes at the "meter scale", e.g., human society's collective dynamics.

Keywords: Brownian motion, epidemic spreading, heterogeneity

PACS: 83.10.Mj, 89.75.–k, 87.23.Ge **DOI:** 10.1088/1674-1056/ad1a90

1. Introduction

Epidemics often occur in different forms in different places[1] and may spread globally and finally become pandemics. The pandemic coronavirus disease that erupted in late 2019 (COVID-19) led to hundreds of millions of confirmed infections and millions of deaths globally. These numbers far exceed those of the previous severe acute respiratory syndrome (SARS) and Middle East respiratory syndrome (MERS), which were also caused by coronaviruses.[2] The destructiveness of COVID-19 is tremendous, and its pandemic has caused severe disaster for humanity. However, we still lack a detailed understanding of how COVID-19 spreads between individuals in the community, which is essential for taking corresponding preventive and intervention measures.

The deterministic compartmental models,[3] including the classic susceptible–infected–recovered (SIR) model[4] and its different derivatives, are numerically studied via solving the differential equations on homogeneous populations with different labels. Due to the random nature of the transmission, a stochastic compartmental model was also proposed,[5] incorporating some random elements. Concerning COVID-19, a significant amount of research has been conducted with these models to illustrate the effects of control and prevention strategies[6–11] in the early stages of the epidemic. Human characteristics and behavior undoubtedly affect disease spreading.[12] However, when the spatiotemporal features of the community and the heterogeneity of the environment and individuals are taken into account, we need to divide the population into many sub-populations (metapopulation models) or construct a complicated agent-based model.[13] Complex models make it hard to explain epidemics from a microscopic view.

Indeed, due to the spatially variable structural features with the high complexity of connections between individuals in human society and the time evolution of the network, epidemic spreading is far more complicated than we can imagine. Statistical mechanics provides an efficient way to study the macroscopic properties of systems with a huge number of states by averaging over the microscopic entities. Seemingly, it is challenging to characterize complicated social behavior with simplified agents for the individual human. Nevertheless, the study of social dynamics with statistical physics has received substantial attention in recent years.[14] Dynamical models combined with classical compartmental models[15–24] have been constructed to explore the microscopic mechanism of epidemic transmission. For example, pedestrian dynamics was integrated with mathematical models to study disease spreading, and the study quantitatively explains the associa-

†Corresponding author. E-mail: donghao@nju.edu.cn
‡Corresponding author. E-mail: wangwei@nju.edu.cn

© 2024 Chinese Physical Society and IOP Publishing Ltd

tion between social distancing and the risk of infection.[22] In a previous study, we developed a real-scale contact-dependent dynamic (CDD) model with explicit individuals in a community and combined it with a traditional compartmental model.[24] We revealed that the distribution of individual types in a population exhibits spatial heterogeneity, influencing the pattern of epidemic spreading.

Brownian motion is a typical stochastic process of particles suspended in a medium.[25] This method can take into account the randomness and unpredictability of the movement of a large number of individuals in a crowd at the molecular level, and its mathematical treatment is relatively simple. To some extent, the Brownian features may still exist at the macroscopic scale, where the movements of individuals are random and difficult to predict. Thus, the treatment of Brownian motion can be transferred to the study of individual movement in a population.[26,27]

In this study, we propose a stochastic individual dynamical (SID) model by incorporating the individuals' stochastic dynamics and social contacts to study the transmission of epidemics. In this model, for a specific region, people are divided into three types: the susceptible, the infectious, and the removal ones (Fig. 1). Brownian dynamics (BD) governs each individual's active movement via similar molecular dynamical simulations in computational physics. If a susceptible person is within the social distancing, say 2 m, from any infector, the vulnerable one will be infected with a certain probability. At the same time, the infected one will switch to the removal subtype at a specific rate. Thus, the population typically experiences an evolution from susceptible to infected and removed status. The flowchart of the model is shown in Fig. S1.

By setting the parameters of our SID model relevant to the reported data for cities, such as Wuhan (China), Lombardy (Italy) or New York State (USA), we find that the early stages of the epidemic transmission curves of COVID-19 for these three cities can be well-reproduced. We show that the lower mobility of people leads to lower infection, which could be attributed to the shielding effect on the infectors, leading to the heterogeneity of the local environment. This demonstrates that a lockdown policy for keeping people at home or moving only locally will effectively control the spread of epidemics. We further find that the outbreak of the epidemic would be delayed for the index cases (the patients zero) aggregated somewhat locally than those dispersed in the system. These observations provide a clue for analyzing and tracing the origin of pandemics. Therefore, detailed information obtained from our SID model offers a microscopic explanation of epidemics.

2. Model and simulations

2.1. Setup of the SID model

As shown in Fig. 1, people in an area with a population of N, which is supposed to be isolated from other parts of the world, are distributed homogeneously or inhomogeneously. They fall into three groups: 1) the susceptible one P_S who is vulnerable to the disease, 2) the infector one P_I who has already been infected and could further transmit the virus to other P_S, 3) the removal one P_R who has either recovered with immunity to the disease thereafter or dead and is excluded from the disease transmission process. Their corresponding numbers are N_S, N_I and N_R, respectively. Initially, the infected one $N_I(0)$ was rare, implying that most of the individuals had not yet been infected.

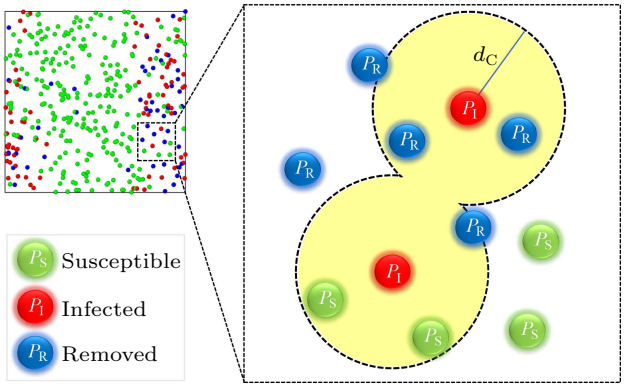


Fig. 1. Illustration of the stochastic individual dynamical (SID) model for the epidemic propagation. Shown are a snapshot of the system and an enlarged part of the local environment obtained from simulation using the SID model. Three types of individuals are defined: the susceptible one P_S (blue) with a number N_S; the infected one P_I (red) with a number N_I; the removed one P_R (green) with a number N_R. The shading in creamy yellow represents the regions within the social distancing d_C of a P_I. The numbers of three types of individuals evolve over time, with the total number $N = N_S + N_I + N_R$ remaining unchanged. The individuals are randomly distributed in a two-dimensional square space with a side length of L to mimic a region, and a periodic boundary condition (PBC) was applied to the square. $L = R_a \cdot N_a$, where R_a is a scale factor used to adjust the system's density. $N_a = \sqrt{N}$, and a broad range of N_a from 20 to 1000 has been calculated. For Wuhan in China, the population density ρ was equivalent to one person per 1000 m^2, and the R_a was then set at 32 m. Accordingly, the R_a values were 49.4 m for Lombardy in Italy, and 84 m for New York State in the USA.

Due to the human-to-human infectious feature of COVID-19,[28] close contact is the primary way to become infected. The SID model was designed to mimic the virus spreading through this kind of direct contact. Thus, a critical infectious distance $d_C \sim 2$ m can be defined, which could be considered conceptually equivalent to "social distancing" in reality,[29] though the transmission of COVID-19 via gas clouds containing pathogen-bearing droplets was found to travel greater distances.[30] In the SID model, within d_C, the infection occurs with a certain probability p_{SI}; no transmission occurs beyond d_C. At the same time, a P_I will transfer to a P_R at a removal rate p_{IR}. Both values of p_{SI} and p_{IR} are in the range of 0 to 1. The type of individuals is updated every recording time t_R.

During the evolution of time, all the individuals move stochastically with different velocities (see the following definition) in the system. Our task is to calculate the distances between each $P_{\rm I}$ and its neighbor $P_{\rm S}$ within $d_{\rm C}$ at a specific time interval. A random number is then assigned to the $P_{\rm S}$. If this number is less than $p_{\rm SI}$, the $P_{\rm S}$ is infected and is subsequently classified as an infector; otherwise, the $P_{\rm S}$ remains vulnerable until the next cycle of assessment. At the same time, another random number is assigned to each $P_{\rm I}$ in the system. If the number is less than $p_{\rm IR}$, this $P_{\rm I}$ switches to a $P_{\rm R}$; otherwise, it remains infectious until the next cycle of assessment. Random numbers follow a uniform distribution over the interval [0,1).

2.2. Parameter setting and simulations

To simulate the stochastic movements of each individual, the Langevin equation $\mathrm{d}\boldsymbol{R}^i/\mathrm{d}t = (\boldsymbol{F}^i_{\rm C} + \boldsymbol{F}^i_{\rm R})/\gamma_{\rm a}$ was used, where $i = 1,\ldots,N$. Here, the position vector $\boldsymbol{R}$ for all persons with a unit of 1 m and a time t with a unit of 0.001 d were used; $\boldsymbol{F}^i_{\rm C}$ is the force on the ith individual from all potentials and constraints, and was set as zero in the present implementation; $\boldsymbol{F}^i_{\rm R}$ is the random force with $\langle|\boldsymbol{F}^i_{\rm R}|^2\rangle = 2\mathrm{d}k_{\rm B}T\gamma_{\rm a}/\Delta t$, where the drag coefficient $\gamma_{\rm a}$ with a default value of 1 and an integration time step $\Delta t = 0.001$ day were used. T is the temperature of the system, and its value was obtained as detailed in the following paragraph. The dimension $d = 2$ was used since the system is defined in two dimensions. Clearly, the $\boldsymbol{F}^i_{\rm R}$ results from a random collision between two particles, which is the basic process of the stochastic BD.

To realize the stochastic BD simulations, we need a "temperature" T of the system, which can be reversely estimated through an average velocity $\langle\boldsymbol{v}\rangle$ of all particles. The $\langle\boldsymbol{v}\rangle$ can be calculated using $\langle\boldsymbol{v}\rangle = \lambda/\tau_{\rm c} = 1/(2\sqrt{2}\tau_{\rm c}\rho d_{\rm C})$, where the mean free path λ can be evaluated based on the population density ρ and social distancing $d_{\rm C}$. The calculated λ values are 177 m for Wuhan, 431 m for Lombardy and 1247 m for New York State. $\tau_{\rm c}$ is the average time between consecutive collisions. It was reported that each person has close contact with dozens of people every day.[7] Therefore, it is reasonable to assume $1/\tau_{\rm c} = 40$, giving a $\tau_{\rm c}$ of 0.025 day, which is an order of magnitude larger than the integration time step Δt. Thus, we obtain $\langle\boldsymbol{v}\rangle \approx 7.1$ km/day for Wuhan.

According to the Langevin equation, the displacement within one integration time step Δt is $\boldsymbol{F}_{\rm R}\Delta t/\gamma_{\rm a}$. Higher T corresponds to faster individual movement. Distribution of displacement in x-direction and y-direction $x_{\rm R}$ and $y_{\rm R}$ in recording time $t_{\rm R}$ approximates a Gaussian distribution. We assume $x_{\rm R}, y_{\rm R} \sim N\left(0,\sigma^2\right)$. Thus, the average speed $v_{\rm R}$ ($\langle\boldsymbol{v}\rangle_{\rm R} = \sqrt{2\pi}\sigma/2t_{\rm R}$) and travel distance $R_{\rm R}$

$$\left(\langle R^2_{\rm R}\rangle = \left(\frac{t_{\rm R}}{\Delta t}\right)\cdot\left(\frac{\boldsymbol{F}_{\rm R}\Delta t}{\gamma_{\rm a}}\right)^2 = \frac{2\mathrm{d}k_{\rm B}T}{\gamma_{\rm a}}t_{\rm R} = 2\sigma^2\right)$$

in recording time $t_{\rm R}$ can be estimated. Thus, we obtain the equation $k_{\rm B}T = \frac{2\gamma_{\rm a}t_{\rm R}}{\pi d}\langle\boldsymbol{v}_{\rm R}\rangle^2$. To set an initial value of temperature T, we assume $\langle\boldsymbol{v}\rangle \approx 7.1$ km/day and obtain a value of $k_{\rm B}T$ as 160000. The calculated daily travel distance of each individual falls in a range of 0–18 km (Fig. S2). With these parameters being available, the BD simulations can be carried out to generate the trajectories of movements.

A probability $p_{\rm IR}$ is required to update the state of the removal persons. This is determined by $p_{\rm IR} = \sum_t p_{\rm IR}(t)\cdot N_{\rm r,I}(t)/\sum_t N_{\rm r,I}(t)$, where $p_{\rm IR}(t) = (\mathrm{d}N_{\rm r,R}(t)/N_{\rm r,I}(t))\mathrm{d}t$ is the removal rate at time t. $N_{\rm r,I}(t)$ and $N_{\rm r,R}(t)$ denote the numbers of infectors and removers (the sum of the recovered and deceased population) in reality at time t, respectively, which is obtained from the publicly available epidemiological data.

On the other hand, the $p_{\rm SI}$ value is parameterized iteratively (Fig. S1) until the calculated data can fit the epidemic curve well. In this study, the $p_{\rm SI}$ and $p_{\rm IR}$ values are 0.0975 and 0.0001908 for Wuhan, 0.2 and 0.000379 for Lombardy, and 0.8 and 0.000202 for New York State, respectively.

It should be noted that in the SID model, the random movements of all the individuals only depend on $k_{\rm B}T$, $\gamma_{\rm a}$ and Δt, while the remaining parameters including $p_{\rm SI}$, $p_{\rm IR}$ and $d_{\rm C}$ together with the stochastic dynamic characteristics of the people determine the progress of the epidemic. The dependence of these parameters towards the prediction of the results has been thoroughly tested, illustrating that the SID model is robust enough to characterize the outbreak of diseases (Figs. S3a–S3d). In addition, the system size and the recording interval have been carefully benchmarked to demonstrate the robustness of the model (supplementary information and Fig. S4).

3. Results and discussion

3.1. Epidemic modeling at realistic scales

3.1.1. Reproduced epidemic transmission curves for the three regions

The calculated epidemic curves for Wuhan, Lombardy and New York State illustrate the cumulative prevalence of the infection and the removed ones (Figs. 2(a)–2(c)). In all three cases, the calculated curves are in line with the reported data, which show exponential growth at the early stage of the epidemic. However, the infection patterns obtained from the SID model further show that the infection takes place locally at the early stage and then gradually spreads to other areas, indicating the vital role of the local environment during the outbreak (Fig. 2(d)), as further described below. The discrepancy at the right end of the curves could be attributed to the effects of control strategies, such as actively treating patients and travel restrictions, which were not considered here with constant $p_{\rm SI}$ and $p_{\rm IR}$. Clearly, these measures effectively slow

down the spreading of the epidemic. Without these specific effects, the infection level will quadruple within half a month, from ∼ 50000 infections at 30 d to ∼ 200000 at 45 d in Wuhan. A similar situation could be observed in Lombardy and New York State.

However, the change in the infection probability in reality is elusive. Various factors, such as public health interventions, seasonal variations and viral mutations, collectively influence the propagation of epidemics. Therefore, in addition to the modeling with a fixed p_{SI}, we further studied the infection with time-dependent p_{SI}. In order to examine the impact of different patterns of p_{SI}, a constraint was applied where the average values are all identical to the constant one. As shown in Fig. S5, under the present system setup, the modeling with the monotonically decreasing p_{SI} consistently predicted an early arrival of the infection peak, though different patterns lead to slightly different curves. In contrast, the monotonically increasing p_{SI} delayed the spread of the virus. The p_{SI} with certain fluctuations, on the other hand, showed the infection curve that is basically the same as the one obtained from the constant p_{SI}.

To avoid complications caused by the pattern of infection probability, we use data from calculations under a fixed p_{SI} for further analysis.

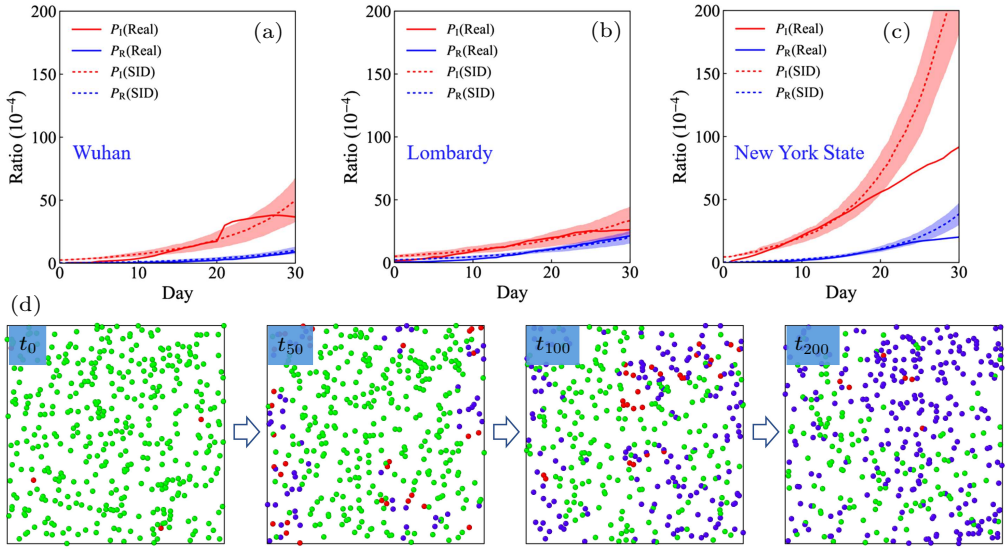


Fig. 2. The calculated epidemic curves in three regions, and a representative transmission pattern generated with the SID model. (a)–(c) The number of infections (in red) and removals (in blue) in Wuhan (a), Lombardy (b) and New York State (c) during the early stage of the COVID-19 outbreak. 50000 individuals were used for these three areas. The solid lines represent the publicly reported data, and the dashed lines are the average of the calculated results from 15 trajectories, with the deviation shown by the shading. The start time is Jan. 22nd for Wuhan, Mar. 4th for Lombardy and Mar. 17th for New York State, respectively. The y-axis in (a)–(c) represents the ratio of the population. (d) A typical pattern of the infection spreading shows the effect of local inhomogeneity on the propagation of the virus. For illustration, a model system containing 400 individuals was used.

3.1.2. Detailed nature of the spread of the epidemic

Further infections generated by a single infector, the primary case, could be identified by tracking each person's movements. R_0, an index to characterize the transmission potential of a virus, is estimated to be 1.92 by averaging the secondary cases caused by the infectors using the SID model. This value is in accordance with the reported values of 2.2 (95% confidence interval, 1.4 to 3.9),[31] 2.24 to 3.58,[32] or 2.68.[8] Accordingly, we estimate that 50% or more of the population needs to be effectively immunized (denoted as p_{immun}) to prevent the sustained spread of the infection of COVID-19, as calculated by $p_{immun} = 1 - 1/R_0$.

Superspreading is an important factor in dynamical models and has been observed in the SID model. Our results show that different individuals have different capacities to spread a virus (Fig. 3). Among them, nearly half the infected people (46.5%) were removed (more likely to recover from the disease) even without being able to transmit the virus to any others. In contrast, a few infected ones could disproportionately infect many others (Fig. 3(a)). 27.0% of infected people who excreted pathogens to at least two people caused 74.7% of infectees, which is the secondary case in Wuhan. For COVID-19, approximately 10% of the infected population causes 80% of transmission.[33] The difference between the SID model and the real world is that particles hardly gather high-density crowds by Brownian dynamics. A more realistic motion may better characterize a high-level superspreading.

In addition, the distribution of time length for switching a P_I into a P_R shows an exponential decay (Fig. 3(b)). Notably, this trend of decline is independent of the simulation time. Our calculations show that nearly half of the infected people are removed within 5 d. This data provides evidence for preventing medical run-up in the epidemic, which could bring substantial social problems. On the other hand, ∼ 22.6% of the infected people cannot be cured for more than 14 d. The latter is qualitatively consistent with the reported 20% severity of COVID-19,[28] which partially validates the p_{IR} value used in this study.

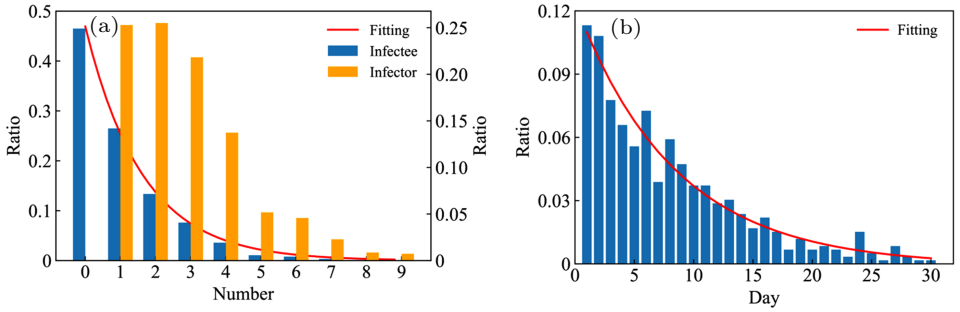


Fig. 3. The infection and recovery dynamics of the epidemic in Wuhan. (a) The distribution of the infector, the primary case, and the infectee, the secondary case. The blue bar shows the ratio of the infectors with different abilities to cause new infections to the total infectors (y-axis on the left). The fitted curve (in red) implies exponential decay of the infection ability. The yellow bar shows the ratio of infectees caused by one infector to the total number of infectees (y-axis on the right). Nearly half the infectors were removed before transmitting the virus to others; 74.7% of infectees (yellow bar when $n \geq 2$) were caused by 27% of infectors. (b) The time for the infected ones to be removed also shows an exponential decay. $\sim$ 45% of them were removed within five days, but $\sim$ 23% had infections lasting longer than 14 d.

3.2. Parameter effects in epidemic modeling

3.2.1. Dynamic characteristics of epidemic transmission

Population mobility is highly associated with the spread of the virus.[34] According to the equation $k_{\mathrm{B}}T = \frac{2\gamma_{\mathrm{a}} t_{\mathrm{R}}}{\pi d}\langle v_{\mathrm{R}}\rangle^2$ previously mentioned, we can adjust the mobility of particles by changing the value of drag coefficient γ_{a} or temperature T.

The SID model quantitatively measures the dynamic lagged effect caused by mobility. At relatively high mobility ($k_{\mathrm{B}}T = 160000$), the peak of the curve for P_{I} appears early, while much later emergence of the summit along with the lower peak height was observed at relatively small mobility ($k_{\mathrm{B}}T = 1600$) (Fig. 4(a)). This could partially explain the practical efforts in curbing the spread of the epidemic with the months-long lockdown strategy in China.

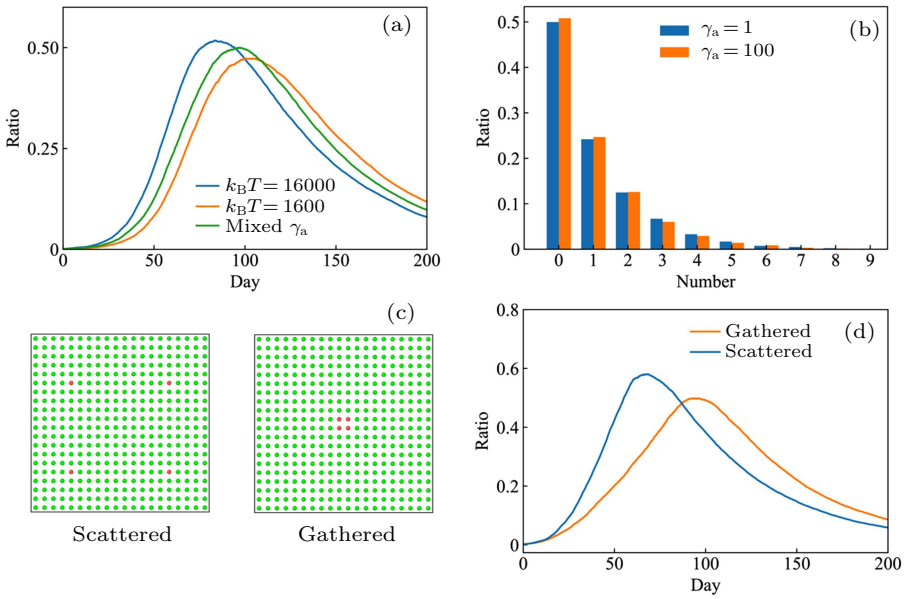


Fig. 4. The dynamic features revealed by the SID calculations. (a) The infection curves with different mobility of the community. The relatively high mobility of the population (blue line) leads to an earlier occurrence of the infection peak than the low-mobility one (orange). Mixed γ_{a} (green line) is a model system where half of the persons move fast ($\gamma_{\mathrm{a}} = 1$) and half move slowly ($\gamma_{\mathrm{a}} = 100$). (b) The number of further infections per infector in the mixed γ_{a} system. Those with higher mobilities show similar infection ability to those with lower movements. (c) and (d) The aggregated state of the index case shows a slower (orange line) spreading rate of the virus than the scattered one (blue line).

The different patterns of human mobility also affect the transmission of diseases. In a model system with mixed γ_{a}, where half of the population moves fast ($\gamma_{\mathrm{a}} = 1$) and the rest moves slowly ($\gamma_{\mathrm{a}} = 100$), an overall delay in the occurrence of a peak on the epidemic curve was observed (Fig. 4(a)). However, it is unexpected that the numbers of infections caused by different mobility groups are basically the same (Fig. 4(b)). This result probably implies that the partial lockdown policy may not efficiently interfere with transmissions in the population.

3.2.2. Infection probability p_{SI} and removal probability p_{IR}

The infection probability p_{SI} represents the spreading capacity of the virus. The removal probability p_{IR} represents the recovery rate and mortality rate. Decreasing p_{SI} or increasing p_{IR} can reduce the transmission speed in the early stages

of infection (Figs. S3a and S3b). Lower p_{SI} corresponds to the effect of public health interventions or reduced virulence of the virus. Higher p_{IR} corresponds to the improvement of treatment effect or decreased disease duration. $R_0 < 1$ and the epidemic outbreak will be prevented when p_{IR} increases fourfold.

3.2.3. Social distancing d_C

The d_C value used here represents how far the virus can travel, though it can occasionally travel much further. Decreasing d_C dramatically reduces the transmission speed in the early stages of infection (Fig. S3d). For example, the peak of the transmission curve with d_C being 2 ($\sim$ 90 d) appears almost twice as late as the one with d_C equal to 3 ($\sim$ 45 d), and much smaller values of d_C, e.g., 1, are not likely to cause large-scale infections. Therefore, our data show that wearing a mask correctly, which decreases the infectious distance, is a simple and efficient way to block viral spread and protect people during the epidemic outbreak.

3.3. Heterogeneity in epidemic modeling

3.3.1. Aggregated status of the index case

When the two systems only differ by the initial distribution state of the infected persons in the population, either gathered or scattered around, the effects of the index case (occasionally referred to as "patient zero") can be explored. As shown in Fig. 4(d), the system with initially gathered infections shows a notably reduced transmission rate and a lowered peak on the epidemic curve. The difference between the two curves depends on the overall mobility of individuals in the system, which is associated with their local environment (see below); the higher the mobility, the smaller the difference. Human movement tends to slow down during an outbreak. Therefore, it may be possible to speculate from the progression rate whether the epidemic originated locally or was imported from other places.

3.3.2. The local environment of the infection

We define the local heterogeneity of the central particle P_I as the ratio between the local and global concentrations of a certain person (either P_S or P_I). Under the low-mobility condition, this ratio is less than 1 for P_S but greater than 1 for P_I (Figs. 5(a) and 5(b)), indicating a low efficiency for the spread of diseases. In contrast, along with the increased mobility of the population, these two values gradually approach 1. Therefore, local heterogeneity makes a difference in the low-mobility condition of the community during epidemics.

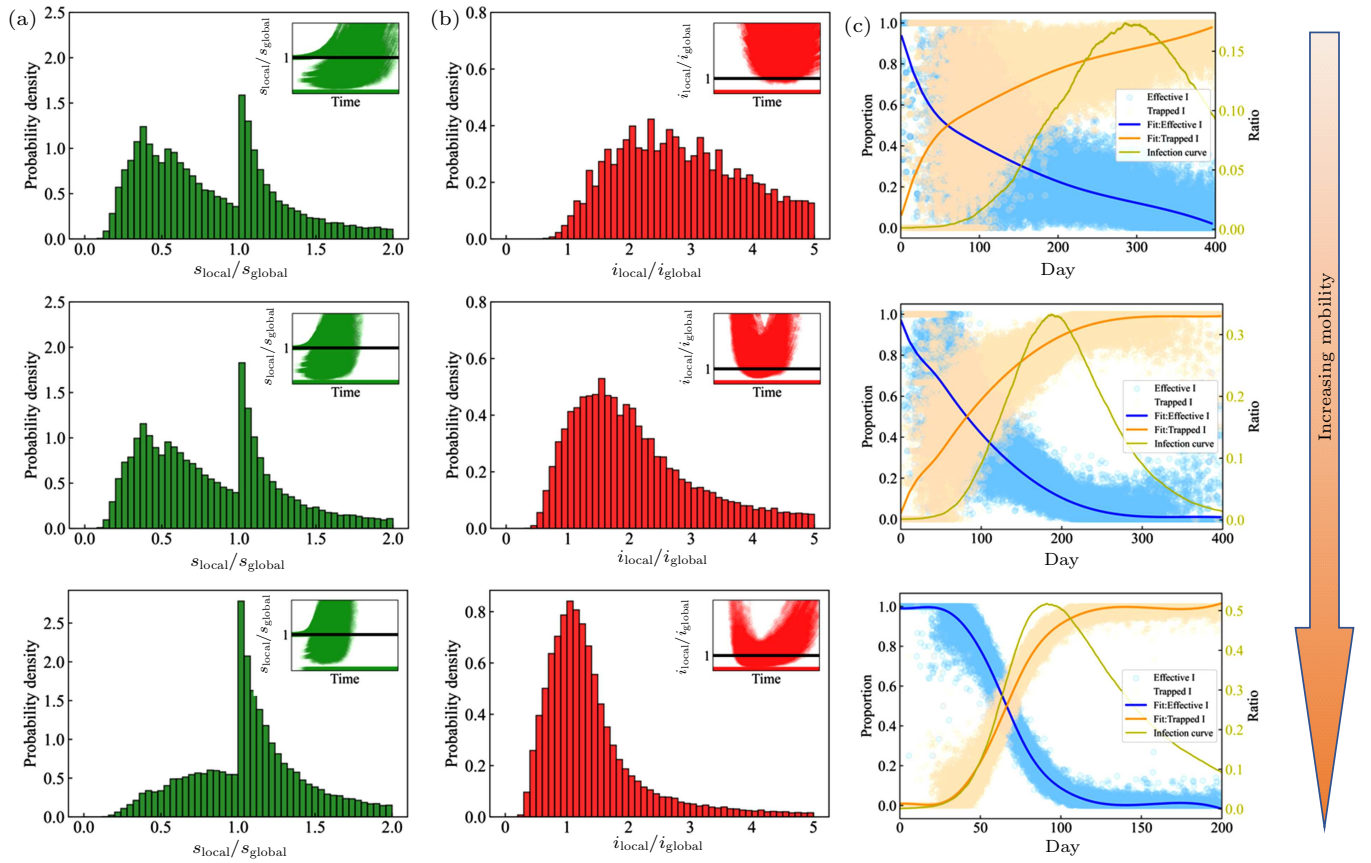


Fig. 5. The local heterogeneity significantly affects the spread of epidemics in a community with relatively low mobility. (a) and (b) Histograms of the local heterogeneity, which is defined as the ratio between the local and global concentrations of P_S (a), and P_I (b). The cases for low (top panel), medium (middle panel) and high mobility (bottom panel) are shown. The inset shows the local heterogeneity over time in each panel. (c) The time evolution of the proportion of the active (blue dots) and trapped (orange dots) infected ones in the system (y-axis on the left). The solid line shows an average of each type of particle. The infection curve is shown in the solid line in yellow (y-axis on the right). For illustration, a model system containing 1600 particles is used. All the data were obtained from 15 sets of simulations using the SID model.

To further explore the heterogeneity, $P_{\rm I}$ with neighbors within $d_{\rm C}$ was subdivided into two categories, $P_{\rm I,a}$ and $P_{\rm I,t}$. The former represents those having susceptible $P_{\rm S}$ particles within $d_{\rm C}$, while the latter has only $P_{\rm R}$ or other $P_{\rm I}$ nearby. The relative proportion of the two is associated with spreading efficiency. As shown in Fig. 5(c), $P_{\rm I,t}$ is dominant in the system with very few $P_{\rm I,a}$ when the particles' overall mobility is relatively low, mainly due to less effective contact between $P_{\rm S}$ and $P_{\rm I}$. This represents a low-infection case. In contrast, in the system with relatively high mobility, $P_{\rm I,a}$ is dominant at the early stage of the outbreak but gradually decreases to a low level due to the reduced number of $P_{\rm S}$ in the system. Indeed, the relatively local concentration of $P_{\rm I}/P_{\rm S}$ around $P_{\rm I}$ is an indicator to measure the spreading efficiency.

The inhomogeneously mixing population across space could also affect the spread of the disease. When the index cases were initially located in a high-density community, we observed an earlier outbreak of the infection (Figs. S6a and S6b). In contrast, a delayed infection pattern was found when the index cases were originally in a low-density population (Fig. S6c). A recent work discloses that population heterogeneity in age also plays an important role in affecting the spread of diseases.[35] Seemingly, the change in the transmission pattern could also be attributed to the difference in the local environment. Therefore, the shielding effect of infected ones induced by regional heterogeneity provides a microscopic interpretation of the lockdown policy.

3.4. Comparison with other epidemic models

By fitting the calculated transmission curve from the SID model, the parameters for the corresponding SIR model can be obtained. The two models show consistent results for the population with high mobility. However, we found a notable discrepancy between the two when the mobility decreases, where the SIR model underestimates the occurrence of the infection peak (Fig. 6). This could be attributed to the constant transmission rate applied in the SIR model throughout the epidemic. However, time-varying infectivity depends on each individual's local dynamics and mobility, and the heterogeneous nature of the population plays an essential role in disease transmission. In other words, simple behavior of each particle generates complicated properties at a higher level. This information is well represented in the SID model.

It should be noted that the groups used in the SID model follow the definition in the SIR model. However, each group only takes integer values in the SID model, which is different from the SIR model's continuous values. More importantly, the characterization of the spatiotemporal variation in the SID model makes the conversion from $P_{\rm S}$ to $P_{\rm I}$ and finally to $P_{\rm R}$, which is totally different to that in the SIR model.

In addition to deterministic compartmental models, such as the SIR model, stochastic dynamic models based on individual movements are one of the most important tools for studying the spread of epidemics, which can take into account the uncertainty and stochastic nature of the infection and recovery process, thus reflecting more realistically the dynamics of epidemic spread. These models have a wide range of applications and can be used to predict the spreading trend of epidemics, evaluate the effects of different interventions, etc. Our SID model is similar to the dynamical model in work[21] with regard to active Brownian motion. However, our model takes into account realistic temporal and spatial scales. Compared with dynamical models on a realistic scale,[22,24] a larger-scale system and a longer-time simulation are easier to archive in the SID model. Furthermore, the random variation of individual movement due to interactions between individuals in the SID model introduces strong stochastic features that enhance its use for describing epidemic spread.

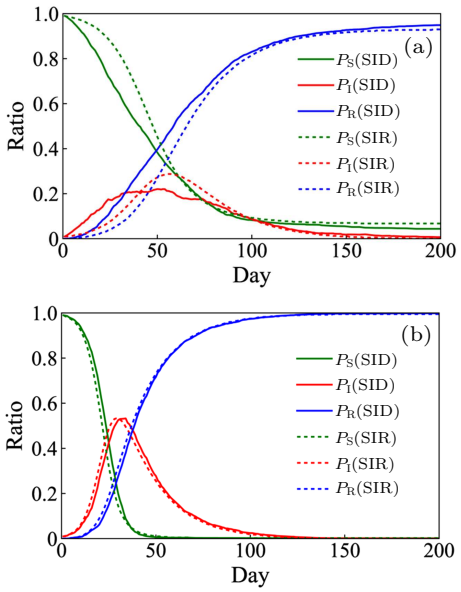


Fig. 6. A comparison between the SID and SIR models. The epidemic curves obtained by the two models show a certain difference in low-mobility condition (a), which could be attributed to the transient local heterogeneity described by the SID model. This effect cannot be well-characterized in the SIR model. However, the two give almost identical curves when the local inhomogeneity of the population cannot be maintained under the condition of relatively high mobility of everyone (b).

4. Conclusion and perspective

The COVID-19 pandemic has caused immeasurable losses to human society. Now, more than ever, we need a healthier world, and consequently, we need a better understanding of the disease. The SID model proposed in this study is robust enabling it to characterize the dynamic features of infectious disease epidemiology with randomness and heterogeneity, and quantifies the effects of control and prevention strategies. It should be noted that additional degrees of freedom, such as allowing people with incubation duration before infection, introducing isolation of the infectees and other in-

terventions resulting from the development of drugs and vaccines, could be easily introduced by properly setting the parameters as variant types (for example, time dependency) in this model.

Acknowledgements

This work was supported by the National Natural Science Foundation of China (Grant No. 22273034) and the Frontiers Science Center for Critical Earth Material Cycling of Nanjing University. Parts of the calculations were performed using computational resources on an IBM Blade cluster system from the High-Performance Computing Center (HPCC) of Nanjing University.

References

[1] Karlsson E K, Kwiatkowski D P and Sabeti P C 2014 *Nat. Rev. Genet.* **15** 379
[2] Zhu F C, Li Y H, Guan X H, *et al*. 2020 *Lancet* **395** 1845
[3] Brauer F 2008 *Mathematical Epidemiology* (Berlin, Heidelberg: Springer) pp. 19–79
[4] Kermack W O and McKendrick A G 1927 *Proc. R. Soc. London, Ser. A* **115** 700
[5] Matis J H and Hartley H O 1971 *Biometrics* **27** 77
[6] Tian H Y, Liu Y H, Li Y D, *et al*. 2020 *Science* **368** 638
[7] Yang Z F, Zeng Z Q, Wang K, *et al*. 2020 *J. Thorac. Dis.* **12** 165
[8] Wu J T, Leung K and Leung G M 2020 *Lancet* **395** 689
[9] Li R Y, Pei S, Chen B, Song Y M, Zhang T, Yang W and Shaman J 2020 *Science* **368** 489
[10] Maier B F and Brockmann D 2020 *Science* **368** 742
[11] Lai S J, Ruktanonchai N W, Zhou L C, Prosper O, Luo W, Floyd J R, Wesolowski A, Santillana M, Zhang C, Du X J, Yu H J and Tatem A J 2020 *Nature* **585** 410
[12] Funk S, Salathe M and Jansen V A A 2010 *J. R. Soc. Interface* **7** 1247
[13] Perra N 2021 *Phys. Rep.* **913** 1
[14] Castellano C, Fortunato S and Loreto V 2009 *Rev. Mod. Phys.* **81** 591
[15] Gonzalez M C and Herrmann H J 2004 *Phys. A* **340** 741
[16] Frasca M, Buscarino A, Rizzo A, Fortuna L and Boccaletti S 2006 *Phys. Rev. E* **74** 036110
[17] Peruani F and Sibona G J 2008 *Phys. Rev. Lett* **100** 168103
[18] Liu Z Z, Wang X Y and Wang M G 2010 *Chaos* **20** 023128
[19] Rodriguez J P, Ghanbarnejad F and Eguiluz V M 2019 *Sci. Rep.* **9** 6463
[20] Levis D, Diaz-Guilera A, Pagonabarraga I and Starnini M 2020 *Phys. Rev. Res.* **2** 032056
[21] Norambuena A, Valencia F J and Guzman-Lastra F 2020 *Sci. Rep.* **10** 20845
[22] Sajjadi S, Hashemi A and Ghanbarnejad F 2021 *Phys. Rev. E* **104** 014313
[23] Bestehorn M, Michelitsch T M, Collet B A, Riascos A P and Nowakowski A F 2022 *Phys. Rev. E* **105** 024205
[24] Zhu Y Y, Shen R Z, Dong H and Wang W 2023 *Plos One* **18** 0286558
[25] Einstein A 1905 *Ann. Phys.* **17** 549
[26] Bartumeus F, Catalan J, Fulco U, Lyra M and Viswanathan G 2002 *Phys. Rev. Lett.* **88** 097901
[27] Lind P G and Moreira A 2015 *Commun. Comput. Phys.* **18** 417
[28] WHO https://www.who.int/docs/default-source/coronaviruse/situation-reports/20200301-sitrep-41-covid-19.pdf?sfvrsn=6768306d_2 [2023-12]
[29] Chu D K, Akl E A, Duda S, Solo K, Yaacoub S, Schunemann H J and Review C S U 2020 *Lancet* **395** 1973
[30] Bourouiba L, Dehandschoewercker E and Bush J W M 2014 *J. Fluid Mech.* **745** 537
[31] Li Q, Guan X H, Wu P, *et al*. 2020 *N. Engl. J. Med.* **382** 1199
[32] Zhao S, Lin Q Y, Ran J J, Musa S S, Yang G P, Wang W M, Lou Y J, Gao D Z, Yang L, He D H and Wang M H 2020 *Int. J. Infect. Dis.* **92** 214
[33] Nielsen B F, Simonsen L and Sneppen K 2021 *Phys. Rev. Lett* **126** 118301
[34] Fang H M, Wang L and Yang Y 2020 *J. Public Econ.* **191** 104272
[35] Britton T, Ball F and Trapman P 2020 *Science* **369** 846