medRxiv preprint doi: https://doi.org/10.1101/2020.11.24.20238337; this version posted January 25, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC 4.0 International license .

Transition to multi-type mixing in d-dimensional spreading dynamics

Alexei Vazquez[1,*]

[1] *German Aerospace Center (DLR), Institute for the Protection of Terrestrial Infrastructures, Rathausallee 12, 53757 Sankt Augustin, Germany*

(Dated: January 24, 2021)

The spreading dynamics of infectious diseases is determined by the interplay between geography and population mixing. There is homogeneous mixing at the local level and human mobility between distant populations. Here I model spatial location as a type and the population mixing by intra- and inter-type mixing patterns. Using the theory of multi-type branching process, I calculate the expected number of new infections as a function of time. In 1-dimension the analysis is reduced to the eigenvalue problem of a tridiagonal Teoplitz matrix. In d-dimensions I take advantage of the graph cartesian product to construct the eigenvalues and eigenvectors from the eigenvalue problem in 1-dimension. Using numerical simulations I uncover a transition from linear to multi-type mixing exponential growth with increasing the population size. Given that most countries are characterized by a network of cities with more than 100,000 habitants, I conclude that the multi-type mixing approximation should be the prevailing scenario.

I. INTRODUCTION

Homogeneous mixing is a cornerstone in the mathematical analysis of infectious disease outbreaks [1, 2]. In a fully mixed population we define the basic reproductive number $R_0 = \beta/\gamma$, where β is the average rate of disease transmission from infected to susceptible individuals and γ is the recovery rate from the disease. When $R_0 < 1$ the epidemic outbreak dies out, but it grows exponentially in time when $R_0 > 1$.

Human populations are not fully mixed, calling into question the value of R_0. There are different mixing patterns according to age, immunization status and the adherence to government recommendations. Nevertheless, these heterogeneous mixing patterns can be treated under a generalized mixing approximation. If we stratified individuals into multiple types, then we can model the infectious disease spread as a multi-type spreading process [3–5]. In multi-type spreading processes R_0 is replaced by the largest eigenvalue ρ of the mixing matrix. When $\rho < 1$ the infectious disease outbreak dies out, but it grows exponentially in time when $\rho > 1$.

Geographical heterogeneity seems to challenge the mixing hypothesis [6–8]. When simulated agents are constrained to a ring with nearest neighbours transmissions, the total number of infected cases grows linearly in time [6]. Sub-exponential infectious dynamics has been reported for agent based simulations in urban settings as well [7]. There are two caveats though. First, the observation of non-exponential dynamics is not sufficient evidence to rule out the mixing hypothesis. I have shown that a combination of a high reproductive number with a truncation of the disease transmission chain yields a power law growth in the number of new infections [9]. The pre-lockdown phase of the COVID-19 outbreak is consistent with this power low growth in the number of new infections [10–12]. Second, there is no prove that geography cannot be harnessed under some mixing approximation.

FIG. 1. Geographical maps highlighting cities with more than 100,000 habitants and the road network connecting them. a) Cuba. b) Germany.

Here I demonstrate that geographical heterogeneity can be modelled as a type. Focusing on large cities, countries like Cuba look like a string of cities, with an effective 1-dimensional topology (Fig. 1a). Other countries, like Germany, are better represented by a two-dimensional mesh, with an effective 2-dimensional topology (Fig. 1b). At this level of description, we can model the network of cities as a multi-type network, where each city is represented by a type and the mixing pattern between cities accounts for the human mobility between them. In the following I introduce the multi-type mixing approximation for type networks with a d-dimensional topology, derive analytical results and test them with numerical simulations.

II. MULTI-TYPE SIR MODEL

I investigate the susceptible, infected and removed (SIR) model on the multi-type network of cities. In the SIR model each individual is in one of three states: susceptible, infected and removed. Infected individuals transmit the disease to susceptible individuals, the latter becoming infected. Infected individuals recover or die

* alexei.vazquez@dlr.de

NOTE: This preprint reports new research that has not been certified by peer review and should not be used to guide clinical practice.

medRxiv preprint doi: https://doi.org/10.1101/2020.11.24.20238337; this version posted January 25, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC 4.0 International license .

from the disease, at which point they are removed from the disease transmission chain, the removed state.

A. One-type branching process

When all individuals are of one-type we have a fully mixed population. We need to make a distinction though, between patient zero and any other case. In more detail, each individual contact other individuals at some rate ζ. Each contact results in disease transmission with probability p, resulting in the effective disease transmission rate $\beta = \zeta p$. Finally, the infected individuals are removed, due to recovery, isolation or death, at a rate γ. Patient zero is selected at random and its typical disease transmission rate is $\langle\beta\rangle$, where $\langle\cdots\rangle$ denotes the average over the variability of β across individuals. Therefore patient zero has the average reproductive number

$$R_0 = \frac{\langle\beta\rangle}{\gamma} \tag{1}$$

R_0 is called the basic reproductive number.

For infected cases other than patient zero we take into account the disease transmission bias towards individuals with a higher contact rate. Any case other than patient zero is found with a probability proportional to its contact rate: $\beta/N\langle\beta\rangle$, where N is the population size. Once infected, the individual found by contact will engage in new contacts at a rate β. Therefore, the average reproductive number for patients other than patient zero is

$$R = \frac{\langle\beta^2\rangle}{\langle\beta\rangle\gamma} \tag{2}$$

R_0 gives the average number of infectious at the first generation, those generated by patient zero. R_0R gives the average number of infections at the second generation and R_0R^{k-1} gives the average number of infections at the k generation. The actual time when an infected case at generation k becomes infected equals the sum of d generation times

$$t_d = \sum_{k=1}^{d} \Delta t_k \tag{3}$$

where Δt_k are the generation intervals, the time interval from the infection of a primary case to infection of a secondary case in the transmission chain from patient zero. If the generation intervals have the probability density function $g(t)$, then t_d has the probability density function $g^{\star k}(t)$, where the symbol $\star$ denotes convolution ($g^{\star k} = \int_0^t g^{\star(k-1)}(\tau)g(t-\tau)d\tau$). Putting the number of descendants and the timing together, we obtain the average number of new infected individuals at time t

$$\dot{I}(t) = I_0R_0\sum_{k=1}^{\infty} R^{k-1}g^{\star k}(t) \tag{4}$$

where I_0 is the of number of patient zeros.

For the standard SIR model the distribution of generating intervals is the distribution of recovery times. Given that recovery takes place at a constant rate, the distribution of generation intervals is exponential

$$g(\tau) = \gamma e^{-\gamma\tau} \tag{5}$$

In this case, equation (4) becomes

$$\dot{I} = I_0R_0\sum_{k=1}^{\infty}\frac{(R\gamma t)^{k-1}}{(k-1)!}e^{-\gamma t} \tag{6}$$

Noting that the series in (6) is the Taylor series expansion of the exponential, we obtain

$$\dot{I} = I_0R_0e^{(R-1)\gamma t} \tag{7}$$

When $R < 1$ the disease dies out, but it grows exponentially when $R > 1$. R replaces the basic reproductive number R_0 in the context of contact heterogeneity ($\langle\beta^2\rangle > \langle\beta\rangle^2$).

B. Multi-type branching process

The multi-type formalism replaces the average reproductive number, a scalar, by a matrix of reproductive numbers. Let n be the number of types, representing cities or subpopulations. To patient zero we assign the reproductive number matrix $R_{0_{n,n}}$. The matrix element $R_{0_{a,b}}$ represents the average number of cases of type b generated by a patient zero of type a. To any other case we assign the reproductive number matrix $R_{n,n}$. The matrix element $R_{a,b}$ represents the average number of cases of type b generated by a case of type a that is not a patient zero. I have previously calculated the expected number of infected individuals of epidemic outbreaks in heterogeneous populations with a multi-type structure [4]. Although the calculation is more involved it yields an expression with the same structure as the one-type case

$$\dot{I} = \gamma I_0^T R_0\sum_{k=1}^{\infty}\frac{(R\gamma t)^{k-1}}{(k-1)!}e^{-\gamma t}J_{n,1} \tag{8}$$

where I_0 is a column vector representing the number of patient zeros by type and $J_{n,1}$ is a column vector of ones. Or, making use of the exponential of a matrix,

$$\dot{I} = \gamma I_0^T R_0 e^{(R-I)\gamma t}J_{n,1} \tag{9}$$

where I is the identity matrix.

C. Diagonalizable R

When R is diagonalizable we obtain results that resemble the exponential dynamics of the single-type case.

medRxiv preprint doi: https://doi.org/10.1101/2020.11.24.20238337; this version posted January 25, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC 4.0 International license .

Let P be the transformation matrix diagonalizing R,

$$R = PDP^{-1} = P \begin{bmatrix} \lambda_1 & & 0 \\ & \ddots & \\ 0 & & \lambda_n \end{bmatrix} P^{-1} \tag{10}$$

where $\lambda_1, \ldots, \lambda_n$ are the eigenvalues of R. Note that some λ_n may be equal if some eigenvalues have multiplicity larger than 1.

I assume that R_0 has the same form of R

$$R_0 = rR \tag{11}$$

where the factor

$$r = \frac{\langle \beta \rangle^2}{\langle \beta^2 \rangle} \tag{12}$$

is the ratio between equations (1) and (2).

Finally, I will use the standard notation for functions of diagonal matrices

$$f(D) = \begin{bmatrix} f(\lambda_1) & & 0 \\ & \ddots & \\ 0 & & f(\lambda_n) \end{bmatrix} \tag{13}$$

Using (10) and (11) we can rewrite (9) as

$$\dot{I} = \gamma r I_0^T P D e^{(D-I)\gamma t} P^{-1} J_{n,1} \tag{14}$$

Whether the number of new cases grows or decays is determined by the largest eigenvalue

$$\rho = \max_n \lambda_n \tag{15}$$

When $\rho < 1$ the disease dies out, but it grows exponentially when $\rho > 1$. ρ replaces the basic reproductive number in the context of multi-type mixing.

The predictive value of ρ can be extended beyond diagonalizable R. In most realistic scenarios there is a bidirectional path between every pair of cities in the network. In the language of graph theory, the network of cities is strongly connected. In such a case R is irreducible in addition to be non-negative. Using the Perron-Frobenious theorem for irreducible matrices [13], one can demonstrate that when $\rho < 1$ the disease dies out, but it grows exponentially when $\rho > 1$ [3, 4].

The diagonalization procedure is extended to any generation interval distribution. Re-starting from equation (4), without specifying the form of $g(t)$, we obtain

$$\dot{I} = r I_0^T P D f(D,t) P^{-1} J_{n,1} \tag{16}$$

where

$$f(x,t) = \sum_{k=1}^{\infty} x^{k-1} g^{\star k}(t) \tag{17}$$

For the SIR model we have $g(t) = \gamma e^{-\gamma t}$, $f(x,t) = \gamma e^{(x-1)\gamma t}$ and we recover equation (14). I have calculated $f(x,t)$ in terms of analytical functions for the gamma distributions of generation intervals [14]. Here I restrict the analysis to the SIR model. The extensions to other generation interval distributions is obtained after plugging in the specific function $f(x,t)$.

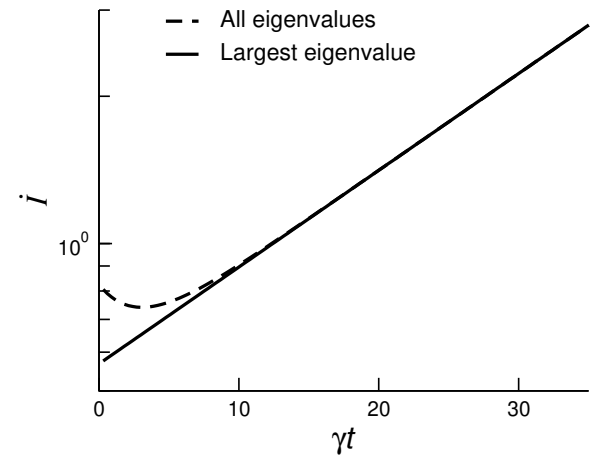

FIG. 2. Multi-type dynamics in 1-dimension for the parameter set $r = 1$, $a = 1.1/2$, $b = 1.1/4$, and $n = 6$. The lines were generated with equation (24) using the contribution of all eigenvalues (dashed) or only the largest eigenvalue (solid).

III. 1-DIMENSION

Now I analyze 1-dimensional topologies. Countries like Cuba have a 1-dimensional typology (Fig. 1a). In this case the reproductive number is given by

$$R_1 = \begin{bmatrix} R_{11} & R_{12} & & 0 \\ R_{21} & R_{22} & R_{23} & \\ & \ddots & \ddots & R_{n-1n} \\ 0 & & R_{nn-1} & R_{nn} \end{bmatrix} \tag{18}$$

where R_{ii}, R_{i-1i} and R_{ii+1} are the intra- and inter-city reproductive numbers. To obtain analytical results, I will work with homogeneous cities and left-right symmetric exchanges: $R_{ii} = a$ and $R_{i-1i} = R_{ii+1} = b$. In this case R is a symmetric tridiagonal Toeplitz matrix.

$$R_1(a,b) = \begin{bmatrix} a & b & & 0 \\ b & \ddots & \ddots & \\ & \ddots & \ddots & b \\ 0 & & b & a \end{bmatrix} \tag{19}$$

The eigenvalue problem of a tridiagonal Toeplitz matrix is solved exactly [15]. The eigenvalues are

$$\lambda_{1k} = a + 2b \cos \frac{\pi k}{n+1} \tag{20}$$

where $k = 1, \ldots, n$. From the latter equation we obtain the largest eigenvalue

$$\rho_1 = a + 2b \cos \frac{\pi}{n+1} \tag{21}$$

The intracity reproductive number a plus 2 times the intercity reproductive number b. The $\cos \frac{\pi}{n+1}$ is a boundary correction. The type at the far left has no left neighbour and the type at the far right has no right neighbour.

medRxiv preprint doi: https://doi.org/10.1101/2020.11.24.20238337; this version posted January 25, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC 4.0 International license .

Since $\cos\frac{\pi}{n+1} = 1 + \mathcal{O}(n^{-2})$ the boundary correction is irrelevant when $n \gg 1$. The value of ρ_1 can be used to determine whether the infectious disease dies out ($\rho_1 < 1$) or grows ($\rho_1 > 1$).

To account for all eigenvalues I go back to (14). The diagonalization matrix for the tridiagonal Toeplitz matrix (19) is given by [15]

$$P_{1ij} = \sqrt{\frac{2}{n+1}} \sin\frac{ij\pi}{n+1} \tag{22}$$

Since $R_1(a,b)$ in (19) is symmetric then $P^{-1} = P^T$. In this case (14) is simplified to

$$\dot{I} = rI_0^T P_1 D e^{(D-I)\gamma t} P_1^T J_{n,1} \tag{23}$$

We can work directly with this equation. We can arrive to a more explicit expression as well. Substituting the form of P_1 (22) and expanding the matrix sums we obtain

$$\dot{I} = \frac{2r}{n+1} \sum_{i,j,k=1}^{n} I_{0i} \sin\frac{ik\pi}{n+1} \sin\frac{jk\pi}{n+1} \lambda_k e^{(\lambda_k - 1)\gamma t} \tag{24}$$

Using the trigonometric identity (AD 361.1 in Ref. [16])

$$\sum_{j=1}^{n} \sin jx = \sin\frac{(n+1)x}{2} \frac{\sin\frac{nx}{2}}{\sin\frac{x}{2}} \tag{25}$$

we can calculate the sum over the j index

$$\sum_{j=1}^{n} \sin\frac{jk\pi}{n+1} = \sin\frac{k\pi}{2} \frac{\sin\frac{nk\pi}{2(n+1)}}{\sin\frac{k\pi}{2(n+1)}} \tag{26}$$

Substituting (26) into (24) and assuming a single patient zero at position i_0 ($I_{0i} = \delta_{i,i_0}$) we obtain

$$\dot{I} = \frac{2r}{n+1} \sum_{k=1}^{n} \sin\frac{k\pi}{2} \sin\frac{i_0 k\pi}{n+1} \frac{\sin\frac{nk\pi}{2(n+1)}}{\sin\frac{k\pi}{2(n+1)}} \lambda_k e^{(\lambda_k - 1)\gamma t} \tag{27}$$

Figure 2 displays the analytical solution for a specific set of a and b values, together with the contribution of the largest eigenvalue alone. For $\gamma t > 10$ the largest eigenvalue is a very good approximation to the full analytical solution. Of note, I have double checked numerically that both (23) and (27) give the same result.

IV. 2-DIMENSIONS

In 2-dimensions I take advantage of the cartesian graph product and the 1-dimensional solution to calculate the eigenvalues. First, I introduce the definition of cartesian product of weighted graphs.

A. Cartesian product of graphs

Cartesian product of graphs: Let G and F be a weighted graph with loops, adjacency-weighted matrices A and B and vertex count n and m, respectively. The cartesian product of G and F, denoted by $G\square F$, is the graph with adjacency matrix

$$A \oplus B = I_m \otimes A + B \otimes I_n \tag{28}$$

where I_n is the identity matrix of size n^2, $\oplus$ denotes the Kronecker sum and $\otimes$ denotes the Kronecker product. The Kronecker product is defined as [17]

$$A \otimes B = \begin{bmatrix} a_{11}B & a_{12}B & \cdots \\ a_{21}B & a_{22}B & \cdots \\ \vdots & \vdots & \ddots \end{bmatrix} \tag{29}$$

Example: An edge represents a complete simple graph with two nodes, denoted by K_2. K_2 has the adjacency matrix $A = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}$. The cartesian product $K_2 \square K_2$ has the adjacency matrix $A \oplus A = I_2 \otimes A + A \otimes I_2$

$$\begin{aligned} &= \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \otimes \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} + \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} \otimes \begin{bmatrix} 1 & 0 \\ 0 & 01 \end{bmatrix} \\ &= \begin{bmatrix} 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \end{bmatrix} + \begin{bmatrix} 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \end{bmatrix} = \begin{bmatrix} 0 & 1 & 1 & 0 \\ 1 & 0 & 0 & 1 \\ 1 & 0 & 0 & 1 \\ 0 & 1 & 1 & 0 \end{bmatrix} \end{aligned} \tag{30}$$

The latter is the adjacency matrix of a graph with vertices in a square. That is why the cartesian product is denoted by the symbol $\square$.

Eigenvalues: The eigenvalues of the Kronecker sum equal the sum of the eigenvalues of the summed matrices [17]. If α_k and β_l are the eigenvalues of A and B, with associated eigenvectors $u^{(k)}$ and $v^{(l)}$ and transformation matrices P and Q, then

$$\lambda_{2n(k-1)+l} = \alpha_k + \beta_l \tag{31}$$

are the eigenvalues of $A \oplus B$,

$$w^{(n(k-1)+l)} = u^{(k)} \otimes v^{(l)} \tag{32}$$

the associated eigenvectors and

$$S = P \otimes Q \tag{33}$$

the transformation matrix diagonalizing $A \oplus B$.

B. 2-dimensional grid decomposition

We can decompose a 2-dimensional grid as the cartesian graph product between a 1-dimensional graph $G_1(a,b)$ with loops and another 1-dimensional graph $G_1(0,b)$ without loops (Fig. 3),

$$G_2(a,b) = G_1(a,b) \square G_1(0,b) \tag{34}$$

medRxiv preprint doi: https://doi.org/10.1101/2020.11.24.20238337; this version posted January 25, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC 4.0 International license .

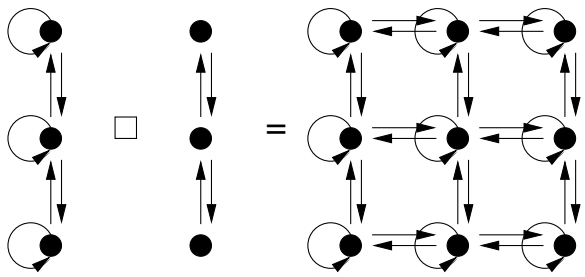

FIG. 3. The cartesian product □ between a string with loops and a string without loops yields a 2-dimensional grid with loops.

The adjacency matrix of $G_2(a,b)$ is

$$R_2(a,b) = R_1(a,b) \oplus R_1(0,b) \tag{35}$$

Substituting (20) into (31) we obtain the eigenvalues of $R_2(a,b)$,

$$\lambda_{n(k-1)+l} = a + 2b\left(\cos\frac{\pi k}{n+1} + \cos\frac{\pi l}{n+1}\right) \tag{36}$$

where $k = 1,\ldots,n$ and $l = 1,\ldots,n$. From the latter equation we obtain the largest eigenvalue

$$\rho_2 = a + 4b\cos\frac{\pi}{n+1} \tag{37}$$

Noting that $R_1(a,b)$ and $R_2(0,b)$ are both diagonalized by P_1, we substitute $P = Q = P_1$ in (33) to obtain the transformation matrix in 2-dimensions

$$P_2 = P_1 \otimes P_1 = P_1^{\otimes 2} \tag{38}$$

Substituting (38) into (14) and bearing in mind that $P_2^{-1} = P_2^T$ we obtain

$$\dot{I} = rI_0^T P_1^{\otimes 2} D e^{(D-I)\gamma t} (P_1^{\otimes 2})^T J_{n,1} \tag{39}$$

This equation is the multi-type mixing approximation for the average number of new infections in a 2-dimensional topology.

V. d-DIMENSIONS

We can iterate the cartesian graph product to generate d-dimensional grids

$$G_d(a,b) = G_1(a,b)\square[G_1(0,b)]^{\square(d-1)} \tag{40}$$

The average number of new infections in these d-dimensional grids is given by

$$\dot{I} = rI_0^T P_1^{\otimes d} D e^{(D-I)\gamma t} (P_1^{\otimes d})^T J_{n,1} \tag{41}$$

with eigenvalues

$$\lambda_{n^{d-1}(k_1-1)+n^{d-2}(k_2-1)+\cdots k_d} = a + 2b\sum_{i=1}^{d}\cos\frac{\pi k_i}{n+1} \tag{42}$$

where $k_i = 1,\ldots,n$ and $i = 1,\ldots,d$, and largest eigenvalue

$$\rho_d = a + 2db\cos\frac{\pi}{n+1} \tag{43}$$

Applications for the 3-dimensional case include the spreading of prion protein aggregates in the human brain.

VI. NUMERICAL SIMULATIONS

The basic reproductive number and its multi-type extension are defined in a contex where the number of infected individuals is much smaller than the population size. In a finite population the number of infected individuals may reach a significant percent of the population size. In that case the starting equation (4) is no longer valid. To investigate the validity of the analytical results and the population size effects, I will perform numerical simulations of the SIR model in 2-dimensional grids.

The cities are assumed of equal population size H. The number of susceptible individuals at city (i,j) is stored in the variable S_{ij}. A list L is created with elements storing the coordinates of all infected individuals. The size of L is the number of individuals in the infected state, $I = |L|$. $\gamma = 1$ and time is measured in units of $1/\gamma$.

Initial conditions: At $t = 0$ all individuals are in the susceptible state except from one individual at city $(1,1)$. L is initiated with the coordinates $(1,1)$ and $S_{ij} = H$ for all (i,j) except for $S_{1,1} = H - 1$.

Dynamics: The Gillespie algorithm is used to update the system. At a total rate $\mu = (a + 4b + \gamma)I$ a new event happens. The time interval Δt to this new even is extracted from the exponential probability density function $\mu e^{\mu\Delta t}$. A coordinate is selected at random from L and stored in (i,j). With probability γ/μ the selected coordinate is removed from L. Otherwise, the disease transmission rule is applied as specified below. The time is updated $t \to t + \Delta t$.

Disease transmission: With probability $a/(a+4b)$ we set $(i',j') = (i,j)$, otherwise we set (i',j') equal to one of the 4 neighbours $(i+1,j)$, $(i,j+1)$, $(i,j-1)$ and $(i-1,j)$ with equal probability. With probability $S_{i',j'}/H$ a susceptible individual in (i',j') gets infected, the coordinates (i',j') are added to L and the updates $S_{i',j'} \to S_{i',j'} - 1$ and $I \to I + 1$ are performed. The empty boundary condition $S_{i,-1} = S_{n+1,j} = S_{i,n+1} = S_{-1,j} = 0$ are used.

Statistics: The number of new infections is recorded in time bins of size 1 and the average is calculated over multiple realizations.

A. Subpopulation size effects

First I illustrate the transition to the multi-type branching approximation with increasing the cities population size. To this end I use the parameter set $r = 1$, $a = 2$ and $b = 1/4$. For $H = 100$ there is an evident

medRxiv preprint doi: https://doi.org/10.1101/2020.11.24.20238337; this version posted January 25, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC 4.0 International license .

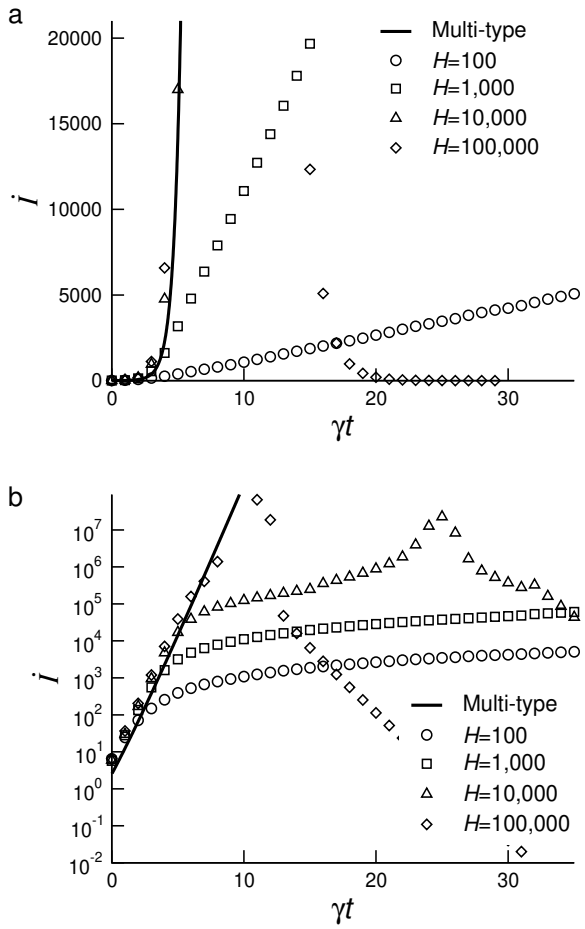

FIG. 4. Average number of new infections for the SIR model in 2-dimensional grids with 100×100 cities, each with H habitants. The model parameters are $r = 1$, $a = 2$, $b = 1/4$. The average was calculated from 100 realizations. a) Y-axis in linear scale. b) Y-axis in log scale

linear increase of $\dot{I}$ as a function of time (Fig. 4, circles). Furthermore, the numerical results are quite far from the multi-type mixing approximation (Fig. 4, line). For $H = 1,000$ the numerical solution gets closer to the multi-type mixing approximation but it follows a linear growth (Fig. 4, squares). Yet, the linear range keeps reducing for $H = 10,000$ and it is almost absent for $H = 100,000$.

As H increases we observe an increase in the region characterized by an exponential growth. Within the exponential growth regime the numerical solution coincides with the multi-type mixing approximation (Fig. 4, line).

B. Lattice size effects

Second I report a lattice size effect. Based on equation (37), for $a+4b \gtrsim 1$ we can find scenarios where the reproductive number is smaller than 1 for $L < L_c$ and larger than 1 otherwise. L_c is obtained by solving equation (37) for L with $\rho = 1$,

$$L_c = \frac{\pi}{\arccos \frac{1-a}{4b}} - 1 \tag{44}$$

For example, for $a = 1.1/2$ and $b = 1.1/4$ the multi-type calculation predicts a transition from decay for $L < 4.13$ to growth for $L > 4.13$. Figure 5 reports the average number of new infections as a function of time for different values of cities size $H = 100$, 1,000, 10,000 and 100,000 and grid linear size $L = 3$, 4, 5 and 6. For $H = 10,000$ and 100,000 there is confirmation of the multi-type mixing prediction. $\dot{I}$ decays for $L = 3, 4$ ($L < 4.13$) while $\dot{I}$ growths for $L = 5, 6$ ($L > 4.13$).

Here again there are city/subpopulation size effects. We do not observe the expected $\dot{I}$ growth for $L = 5$ when $H = 100$ or 10,00. This example teach us that some diseases may not generate outbreaks when the network of cities is small, but as they city network grows they generate outbreaks.

For the largest city sizes $H = 10,000$ and 100,000, I have plotted the analytical solution calculated from equation (41). At the early times the points from the numerical simulations fall in the multi-type mixing line. For longer times there is a deviation from the analytical results due to the finite population size.

VII. CONCLUSIONS

I have demonstrated the use of the multi-type network approach and the cartesian product of graphs to calculate the epidemic threshold of spreading dynamics in d-dimensional grids. These calculations can be deployed to estimate the epidemic threshold at the country level aggregating cities/regions of the oder of 100,000 habitants or larger. The largest eigenvalue of the reproductive number matrix will provide a good estimate of the epidemic threshold. The analytical formula for the average number of new infections will provide a good approximation to the initial dynamics of the outbreak.

The numerical simulations in 2-dimensional grids uncovered a gradual transition from linear to exponential growth with increasing the cities/subpopulation size. The same is true for epidemic spreading in fully mixed populations. Depending on the magnitude of R and the implementation of lockdowns we can expect both exponential or power law growth [9]. Today countries are characterized by a network of cities with more than 100,000 habitants. In this context the numerical simulations reported here indicate an exponential growth of the number of new infections in time. Based on this observation I find unlikely that geography is the dominant factor behind the observation of power law growth. In contrast, the power law growth reported for COVID-19 is consistent with the time scales of the lockdown implementations [12]. However, I cannot exclude that there is a combination of both the lockdown truncation of the disease transmission chain and geographical location.

medRxiv preprint doi: https://doi.org/10.1101/2020.11.24.20238337; this version posted January 25, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC 4.0 International license .

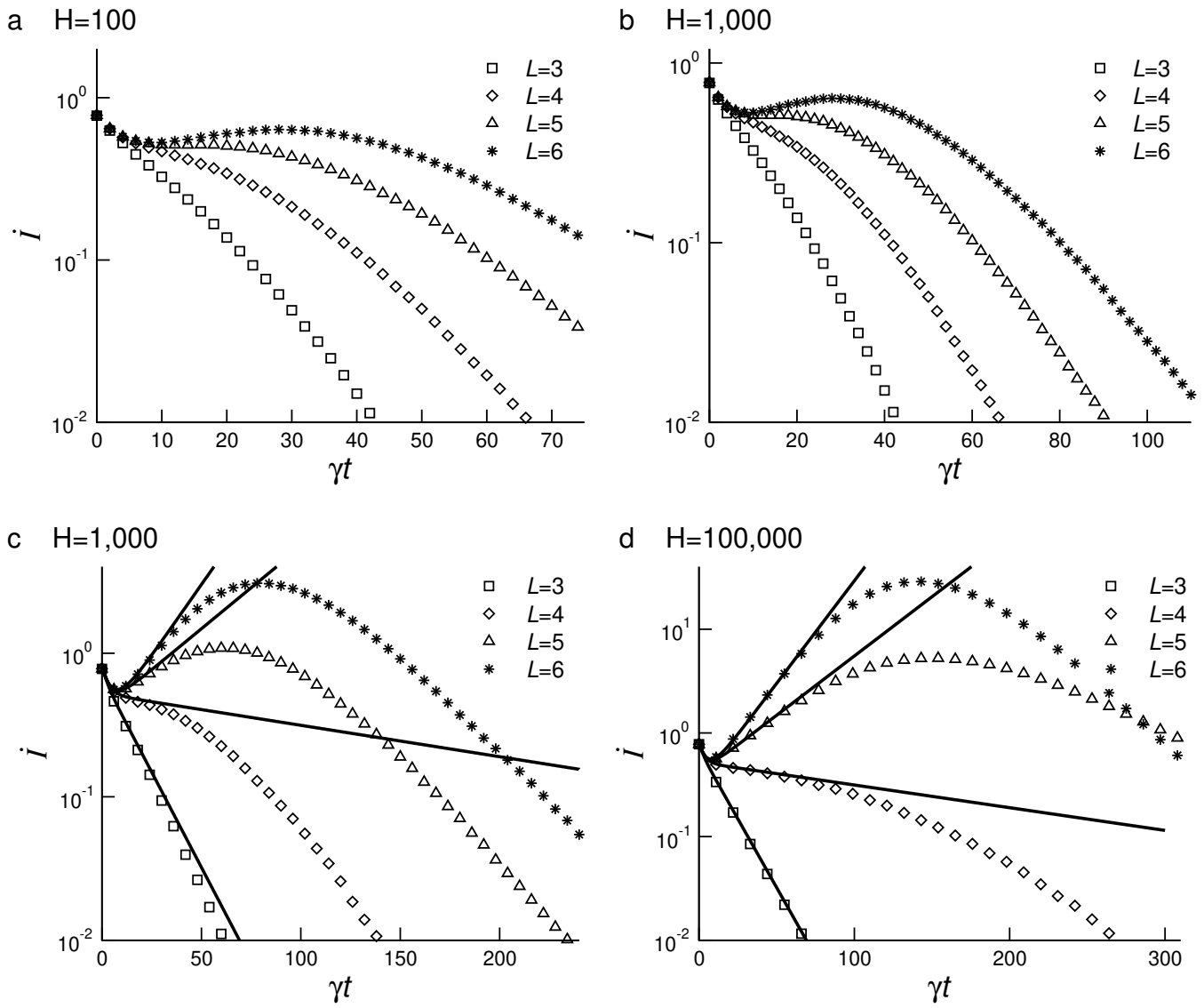

FIG. 5. Average number of new infections for the SIR model in 2-dimensional grids with $n \times n$ cities, each with H habitants. The model parameters are $r = 1$, $a = 1.1/2$, $b = 1.1/8$ and $n = L$. The average was calculated from 1,000,000 realizations.

[1] R. Anderson and R. May, *Infectious Diseases of Humans: Dynamics and Control*, Dynamics and Control (OUP Oxford, 1992).

[2] O. Diekmann and J. Heesterbeek, *Mathematical Epidemiology of Infectious Diseases: Model Building, Analysis and Interpretation*, Wiley Series in Mathematical & Computational Biology (Wiley, 2000).

[3] O. Diekmann, J. A. P. Heesterbeek, and J. A. J. Metz, Journal of Mathematical Biology **28**, 365 (1990).

[4] A. Vazquez, Phys. Rev. E **74**, 066114 (2006).

[5] A. Vazquez, medRxiv (2020), 10.1101/2020.10.09.20210252.

[6] S. Thurner, P. Klimek, and R. Hanel, Proceedings of the National Academy of Sciences **117**, 22684 (2020).

[7] D. J. Haw, R. Pung, J. M. Read, and S. Riley, Proceedings of the National Academy of Sciences **117**, 23636 (2020).

[8] N. L. Komarova, L. M. Schang, and D. Wodarz, Journal of The Royal Society Interface **17**, 20200518 (2020).

[9] A. Vazquez, Phys. Rev. Lett. **96**, 038702 (2006).

[10] A. L. Ziff and R. M. Ziff, medRxiv (2020), 10.1101/2020.02.16.20023820.

[11] H. M. Singer, Phys Biol **17**, 055001 (2020).

[12] A. Vazquez, Phys. Rev. E **102**, 040302(R) (2020).

[13] F. Gantmakher, *The Theory of Matrices*, AMS Chelsea Publishing Series No. v. 2 (AMS Chelsea, 2000).

[14] A. Vazquez, medRxiv (2020), 10.1101/2020.10.17.20214262.

[15] M. Gover, Linear Algebra and its Applications **197-198**, 63 (1994).

medRxiv preprint doi: https://doi.org/10.1101/2020.11.24.20238337; this version posted January 25, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC 4.0 International license .

[16] I. Gradshteyn and I. Ryzhik, *Table of Integrals, Series, and Products* (Elsevier Science, 2014).

[17] R. Horn, R. Horn, and C. Johnson, *Topics in Matrix Analysis* (Cambridge University Press, 1994).