MODELLING SIR-TYPE EPIDEMICS BY ODES, PDES, DIFFERENCE EQUATIONS AND CELLULAR AUTOMATA – A COMPARATIVE STUDY

Günter Schneckenreither[1], Nikolas Popper[2], Günther Zauner[2]

[1]Institute for Analysis and Scientific Computing, Vienna University of Technology,
Wiedner Hauptstraße 8-10, 1040 Vienna, Austria
[2]'Die Drahtwarenhandlung' – Simulation Services,
Neustiftgasse 57-59, 1070 Vienna, Austria

gschneck@osiris.tuwien.ac.at(Günter Schneckenreither)

Abstract

We give a comparative overview over some different approaches towards modelling spatial spread of epidemics and present methods for identifying these approaches respectively. The basis of this study is the classical Kermack-McKendrick susceptible-infected-recovered (SIR) ordinary differential equations (ODE) model. In order to introduce a spatial component in the spread of diseases, we extend the classical model in a first step by using lattice gas cellular automata (LGCA) and stochastic cellular automata (stochastic CA). These approaches are based on motion respectively distributed contacts of individuals and permit to observe simple strategies for confining epidemic outbreaks and to develop provisional methods towards an identification. We proceed with the introduction of a partial differential equations (PDE) model for simulating spatial spread. Like the stochastic CA model this approach is based on distributed contacts among individuals and uses a probability density function to describe the interaction behaviour. We can develop an universal relation between those two approaches by using the central limit theorem. Further we observe random motion as a scaling limit of Brownian motion in the LGCA model and the diffusion distribution, which we derive from the Gaussian semigroup (Brownian motion), in the PDE and stochastic CA approach on the other side. A Fast Fourier Transform (FFT) frequency analysis with susceptible-infected-recovered-susceptible (SIRS) extensions of these three model approaches shows that our identification methods deliver very good correspondence.

Keywords: epidemic spread, susceptible-infected-recovered model, lattice gas cellular automata, stochastic cellular automata, partial differential equations.

Presenting Author's Biography

Günter Schneckenreither studies mathematics at Vienna University of Technology and joined the Modelling and Simulation Group of the Institute for Analysis and Scientific Computing in 2006 where he currently works on models for SIR-type epidemics.

1   ODE model

The SIR ODE model was stated in 1927 by Kermack and McKendrick in order to describe the temporal evolution of the number of individuals in a population who are either susceptible, infected (and contagious) or recovered. It is composed of a system of three partially decoupled non-linear ordinary differential equations (1) and contains only two parameters α and β, which determine the rates of infection and recovery.

Due to its simplicity but also because of some deficiencies, this model serves as basis for many extensions, which involve incubation periods, the possibility of getting infected again after a period of immunity, demographic partitions or spatial spread. Especially for introducing the latter, there exist a variety of different approaches. Among them also cellular automata and partial differential equations. Sections 1-4 of this paper are related to an article [1], which was published in Simulation News Europe (SNE), and deal with LGCA and stochastic CA for simulating epidemic spread. Sections 5-7 present an extended PDE model and deliver advanced methods for model identifications.

$$
\begin{aligned}
\frac{\partial S(t)}{\partial t} &= -\alpha S(t)I(t) \\
\frac{\partial I(t)}{\partial t} &= \alpha S(t)I(t) - \beta I(t) \\
\frac{\partial R(t)}{\partial t} &= \beta I(t)
\end{aligned}
\tag{1}
$$

The recoveries in the classical model $\beta I(t)$ are linear dependant on the size of the infected group and assign a geometrical distribution to the duration of the infection.

The growth of the group of infected is determined by $\alpha S(t)I(t)$, which invokes that the amount of new infections depends on the interaction between susceptible and infected individuals [2]. On the one hand, if we assume that the individuals are situated on a two dimensional domain and interaction depends on geographic distances, this means that the individuals must always be absolutely uniformly distributed. And on the other hand, if a contagious and a susceptible individual come into contact, the probability of infection must be α.

2   General information on CA

For the preceding considerations some basic information on CA and consistent notations are necessary.

The observed domain (p.e. a geographical region) is discretised on a square or hexagonal structured lattice, where the resulting cells can hold different states, which (p.e. describe the stage of infection of an individual and) change only at discrete time steps. To calculate the states of all cells in the next time step, an automaton rule is applied on all cells simultaneously. This function is applied locally and depends on the current state of the cell itself and the 'neighbouring' cells. The neighbourhood can be for example the four (Van Neumann neighbourhood) respectively six nearest neighbours or any stochastically determined set of cells on the domain (stochastic CA).

A cellular automaton whose cells contain particles, that interact within the cells and move to the neighbouring cells at discrete time steps, is called lattice gas cellular automaton (LGCA). We denote LGCA with a hexagonal lattice and six particles per cell as Frisch-Hasslacher-Pomeau (FHP) automata and LGCA with a square lattice and four particles per cell as Hardy-de Pazzis-Pomeau (HPP) automata. The particles change their moving direction according to collision configurations, which are applied depending on the number of particles within the cell and their positions.

For the implementation of CA we use Matlab, which provides easy access to three dimensional matrices, array manipulation and visualisation devices.

3   LGCA model

For the SIR FHP LGCA model, which we will use as a representative for LGCA approaches, and for many other models which consider spatial spread, the population must be discretised (resolved into single individuals and placed) on the two dimensional LGCA-lattice structure depending on the desired initial conditions. Each individual is represented by a particle and can hold one of the discrete states susceptible, infected or recovered.

3.1   General considerations

We will now discuss several features of LGCA, which give a deeper insight, allow small modifications within the LGCA or need to be analysed in-depth.

We again use a parameter β, which determines the probability of recovery for an infected individual and $\tilde{\alpha}$, which determines the probability of infection when a susceptible and an infected individual come into contact (are situated in the same cell). The probability of infection for a susceptible individual depends on the number of contacts and accordingly on the number of infected individuals in the same cell (I_c) and can be written as

$$
\Psi_c = 1 - (1 - \tilde{\alpha})^{I_c}.
\tag{2}
$$

It is clear that the probability of infection is different for each cell.

Because of the hexagonal structure of the lattice, the number of contacts is limited to the five (three for the HPP LGCA) other individuals in the same cell. And by choosing a lower population density, the number of contacts becomes even smaller.

The specification of transition rules (collision configurations), allows to control the motion of individuals and thus to simulate social interaction and demographic features. Two possibilities are FHP-I rules, which conserve mass and momentum [3], or random motion, which leads to a diffusive behaviour.

Concerning the boundary conditions of the lattice, there exist two possible configurations. Periodic boundary

conditions let particles, which leave the domain at one side, reenter from the other side. Reflective boundary conditions reflect particles, which hit the border of the domain. The first type of boundary conditions can be more suitable for simulating a non geographical domain and it just seems to be a little bit easier to implement.

If the LGCA model should be adapted to an existing infection rate α and the corresponding behaviour of the ODE model, the number of infections per time step in the LGCA must necessarily be modified. The reason for this is that the number of contacts in the LGCA model is always smaller than in the continuous approach.

The expected probability of infection (we assume a uniform distribution of the individuals) in the LGCA is

$$\Psi \quad = \quad 1 - (1 - \tilde{\alpha})^{\frac{I}{N}} \qquad (3)$$

where I is the overall number of infected and N is the number of cells. A Taylor series expansion [4, 5] shows that

$$(1 - \tilde{\alpha})^{\frac{I}{N}} = 1 - \frac{\tilde{\alpha}I}{N} + \frac{\tilde{\alpha}^2 I(I - N)}{2N^2} + \ldots \qquad (4)$$

Consequently for small $\tilde{\alpha}$ the probability of infection can be approximated by $\Psi \approx \tilde{\alpha}\frac{I(t)}{N}$, which leads to $\tilde{\alpha} = \alpha N$ if we want a behaviour similar to the ODE model.

3.2 Model identification – spatial inhomogeneities

The most distinctive difference between the classical and the LGCA model is the local character of disease propagation in the latter approach. This feature favours the formation of subgroupings of infected individuals on the domain. From those 'epidemic' areas the individuals diffuse rather slowly and can not spread the disease unless they reach cells with susceptible individuals. Consequently it is not surprising that transition rules like FHP-I rules, which lead to a fast mixing of the population, deliver faster spread than transition rules like random motion, which determine diffusive motion.

Concerning the initial conditions of the LGCA with the same arguments we come to the conclusion that homogeneous initial conditions deliver faster spread of the epidemic. Now if we lower the density of the individuals in the CA (use a larger domain but with the same number of individuals) and always reconfigure the infection rate $\tilde{\alpha} = \alpha N$, simulations show that epidemic spread becomes slower (Fig. 1).

Experiments with the FHP and HPP models show that spread of epidemics is faster in the FHP model (Fig. 2). Even though the number of individuals per cell in the FHP automaton is higher, this behaviour is not expected because the infection rates have been adjusted separately for both models and should deliver the same rate of overall infections. The only possible explanations are that the lattice of the FHP automaton is smoother and that the cell-neighbourhoods are larger. Because a large neighbourhood corresponds to more dynamic, faster and wide ranged movements of the individuals,
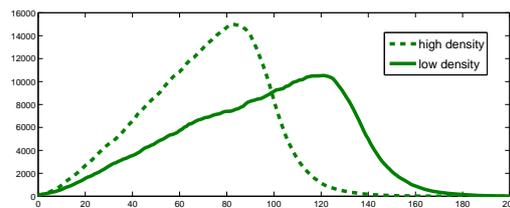

Fig. 1 Infected individuals from FHP LGCA simulations with $100\,000$ susceptibles and $240\,000$ resp. $135\,000$ cells (infection rates 0.72 resp. 0.405).

we can assume that a larger neighbourhood generally favours faster spread [1].

To understand that the LGCA model really extends the classical ODE model, we show that the time-discretisation of the ODEs is concerning the speed of spread an upper bound and concerning spatial inhomogeneities a lower bound for the FHP simulations.

In order to accelerate epidemic spread in the LGCA it is necessary to dissolve the local character of motion by repositioning the individuals randomly every time step [4]. Increasing the infection rate in order to accelerate the epidemic has no effect after a certain value has been reached, because the local character of contagion (restricted to the individuals within the same cell) limits the speed of spread. To dissolve the local character of motion very conscientiously, the lattice must be rearranged very often so that the average distribution of individuals, which then is uniform, can be applied to the lattice.

In (3) the expected probability of infection in the LGCA was calculated, which implies that a uniform distribution of the individuals was assumed. Afterwards a simplification with the help of a Taylor series expansion was presented. Therefore it is now possible to describe the evolution of the epidemic on the lattice by the difference equations system

$$
\begin{aligned}
S(t+1) &= S(t) - S(t) \cdot \tilde{\alpha}\frac{I(t)}{N} \qquad (5)\\
I(t+1) &= I(t) + S(t) \cdot \tilde{\alpha}\frac{I(t)}{N} - I(t) \cdot \beta\\
R(t+1) &= R(t) + I(t) \cdot \beta.
\end{aligned}
$$

If the individuals are not distributed uniformly, this method gets rather imprecise.

But if we assume a uniform distribution for all time steps and employ $\tilde{\alpha}\frac{I(t)}{N} = \alpha I(t)$ on this system (5), we have

$$
\begin{aligned}
S(t+1) &= S(t) - S(t) \cdot \alpha I(t) \qquad (6)\\
I(t+1) &= I(t) + S(t) \cdot \alpha I(t) - I(t) \cdot \beta\\
R(t+1) &= R(t) + I(t) \cdot \beta,
\end{aligned}
$$

which actually corresponds to the time-discretisation of the classical ODE model.

Accordingly by dissolving the local character and simultaneously reducing the step size, the LGCA con-

verges towards the ODE model (Fig. 2). But this means nothing else than that the LGCA model is a (discretised) extension of the classical model by a spatial character of disease propagation.

A more demonstrative method to achieve the same result is by applying (5) on each single cell containing $s(t) = \frac{S(t)}{N}$, $i(t) = \frac{I(t)}{N}$ and $r(t) = \frac{R(t)}{N}$ susceptible, infected and respectively recovered individuals. The resulting equations do again match (6).
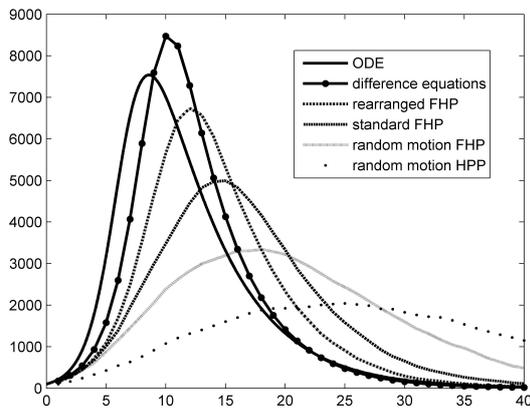

Fig. 2 [1] Infected individuals from different simulation techniques. In this case the lattice was rearranged only once every time step (rearranged FHP).

3.3   Implementation of LGCA

Generally matrices and the values $0, 1, 2, 3$ (empty position, susceptible, infected, recovered) can be used to store the positions and the states of the individuals.

Concerning the implementation of a FHP LGCA the special structure of the lattice has to be taken into account. If a three dimensional matrix with one layer for each particle-position (which might be the best possible method and also allows good optimisation with Matlab) should represent a square shaped hexagonal lattice, the even cell-rows must be shifted to the right in order to relate the lattice to the matrix. Consequently the structure of the neighbourhood and the transition rules in the matrix representation are different for even and odd rows, which actually makes things unnecessarily complicated. If the shape of the domain and the lattice is assumed to be rhomboid, all rows can be shifted in one direction in order to obtain a square shaped structure, that can be related to a matrix. In this case there exists only one distinctive structure of the neighbourhood and one kind of transition rules. This deformation of the domain does not strongly influence the behaviour of the FHP automaton if periodic boundary conditions are used. For reflective boundary conditions there are two sharp corners, which are not that easily accessible for the particles.

The implementation of random motion is rather straight forward. In order to obtain special collision configurations, the particles in the cells and accordingly along the third dimension of the matrix must be counted. Some optimised matrix calculations can deliver the empty po-

sitions and so the direction of particle-rotation can be determined for every cell. The motion of the particles between the cells can be performed by shifting the layers of the three dimensional matrix into different directions. To simulate the infections, the number of infected individuals in each cell must be counted and according to (2) a binomial distributed random number must be calculated for every susceptible individual.

3.4   Vaccination policies

One reason for the development of models for simulating epidemic spread is to find or test methods for confining outbreaks. Such investigations can not be performed with the classical ODE model and involve vaccinations and/or quarantining.

Simulating a full or partial quarantine of a group of individuals with the LGCA approach is not straight forward since quarantining involves the restriction of motion for a special set of cells on the domain, which would require a completely new implementation of the transition mechanism. What is easily accessible with the LGCA model is the observation of the behaviour of an epidemic, when a group of vaccinated (recovered) individuals is placed homogeneously or under certain considerations inhomogeneously on the domain. Such considerations can concern the surrounding of an epidemic area with vaccinations [4, 5] or to prevent a disease from spreading to certain regions of the domain or to simply slow down the speed of spread.

Simulations with the random motion FHP model show that the barrier strategy (surrounding an epidemic area with vaccinations) permits two different evolutions of an outbreak. If the barrier is very tight and large, no infected individuals from the epidemic area diffuse to the rest of the domain and the epidemic fades away. But in the other case if some infected individuals can pass through the barrier the epidemic reaches a second climax outside the primal epidemic area. In this case the spread can be even faster than for other vaccination policies like vaccinations inside the epidemic area, which do not intend to confine the outbreak, but rather target a slow down in the spread of the disease.

Especially for modelling advanced vaccination policies, which intend the use of quarantines and a chronological order for vaccinating certain subsets of the population, model approaches like agent-based CA [6] are more suitable.

3.5   Epidemic Waves

If we introduce further possible states for the particles (susceptible - infected - contagious - immune) and allow immune/recovered individuals to become susceptible again, we can observe the typical wavelike evolution of the number of individuals in each of the groups, which can also be observed for epidemics in real life.

We will use an extended model in section 7.1 in order to provide a solid basis for comparing different model approaches. And we additionally produce clearly visible waves in section 7.2.

4   Stochastic CA model

By stochastic CA we mean ordinary CA (without considering motion of particles) with a stochastically determined neighbourhood. The primary interests of the following section are a flexible definition of the interaction area/contact neighbourhood and the connections and differences between LGCA and stochastic CA. The application of stochastic CA for modelling epidemic spread can be found in [7] for example.

4.1   Stochastic neighbourhood and implementation

For simulating sociological interaction between individuals (which is also the basis for contagious contacts) it is necessary to find a neighbourhood for every cell, that represents the sociological environment of an individual. This idea applied on the FHP model would mean that a person has only contact to maximal five other persons per time step, which was also the reason for the restrictions concerning the speed of spread in the LGCA approach. On the other side this idea applied on the classical ODE model would mean that all individuals come into contact every time step (section 1).

To define a sociological environment in the stochastic CA model, that provides a gradation for the occurrence of interaction between individuals, a decaying likelihood of interaction between cells depending on the distance between them can be used. This approach delivers a radial-symmetric distribution of the contacts for each cell, what principally can be described by an arbitrary probability distribution or a similar function, which we will denote likelihood functions.

To implement this model the advantages of matrix shifting have to be abandoned and ordinary loops through the cells of the lattice must be used instead. For establishing the contacts, whose positions should follow a particular distribution, the determination of the contacts can easily be done by generating a random ankle and a random radius which is distributed with the likelihood function. If the likelihood function does not have the features of a probability distribution, a so called bounding box algorithm can be used [7]. This means that depending on the desired number of contacts per time step and the likelihood function an increasing sequence of radii must be generated so that within each of the circles defined by a radius one cell can be chosen randomly. This method for determining the contacts allows to control several new parameters, which allow a very detailed definition of the contact behaviour.

Of course it is easy to show that this approach extends the classical model by a local character of disease propagation. The ODEs are again an upper bound concerning the speed of spread and a lower bound for spatial inhomogeneities. If the stochastic CA establishes contact between each two cells (dissolution of the local character and increase of speed), a probability of infection for every individual of $\Psi = \alpha I$ is obtained and consequently the growth of the group of infected is αIS, which corresponds to the discretised ODE system.

4.2   Comparison with the LGCA model

The distinctive differences between LGCA and stochastic CA are that the stochastic CA does not consider motion of individuals but allows a more flexible definition of the neighbourhood. We try to identify these two models, by observing the spread under several special considerations, conditions and modifications and then generalise the outcomes to the original forms of the models.

First a HPP simulation with full density of the population and the rates $\alpha = 1$ and $\beta = 0$ is observed. In this special case it plays no role which transition rules are used, because if the cells contain four infected individuals in the next time step all individuals in all four surrounding cells become infected in both situations. Consequently it is possible to assign the status susceptible, infected, recovered etc. to whole cells instead of single particles. Therefore in this case the simulation with the HPP model exactly matches a simulation with a Van Neumann CA, which also uses the rates $\alpha = 1$ and $\beta = 0$. The infectious area is always a rectangular region, which grows symmetrically every time step (Van Neumann neighbourhood of radius t).

On the other side if the contact distance for the stochastic CA is always kept between $0.6999\ldots$ (0 would be possible too) and $0.7 < \sqrt{0.5^2 + 0.5^2}$ (take a look at the square lattice structure in Fig. 3 a) and the number of contacts grows infinitely, the infections are always transmitted to the whole Van Neumann neighbourhood.
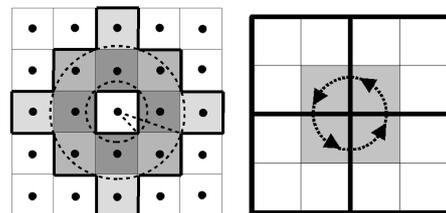

Fig. 3 [1] a) Stochastic CA: neighbourhood with radius. b) Rotational HPP LGCA: neighbourhood.

Therefore the Van Neumann CA is an upper bound concerning the speed of spread in the HPP automaton and a lower bound concerning the size of the neighbourhood for the stochastic CA – actually only when contact to all cells within the contact radius is established, which was ensured by the infinitely high number of contacts per time step.

The second approach towards an identification delivers weaker results and does not request a special infection and recovery rate, but is rather built upon modified transition rules for the HPP model what involves that not the cells, but single individuals are observed. Because particles in the LGCA do not have a fix or bounded environment of interaction as in the stochastic CA, we minimise this area by introducing rotational motion. This leads to $4 \cdot 3 = 12$ different direct contact particles for every particle, because in every discrete moment a particle has contact to three other particles (Fig. 3 b). The

numbers of direct contacts after one, two etc. time steps is $3, 6, 9, 12, 12, 12, \ldots$.

To reach the same number of contacts in the stochastic CA, we allow three contacts per time step and keep the contact distance always between $d \in [0, 1.5]$ and $1.58 < \sqrt{0.5^2 + 1.5^2}$ and thus reach that a individual can establish contact to 12 other individuals (Fig. 3 a). If lower values are used for d, the likelihood of interaction with the outer cells becomes smaller and spread slower. Simulations show that for $d \approx 1.5$ the two modified automata deliver the same behaviour.

Generally there can always be found parameters for the stochastic CA so that the behaviour of an epidemic outbreak is the same for both models.

In section 6 a more advanced method towards identifying these two models is presented.

4.3 Vaccination policies

With the stochastic CA model it is now easily possible to change the contact behaviour of regional groups of the population by defining a different number of contacts for different regions of the domain. Like with the FHP model several different vaccination policies can be examined.

A quarantining policy can be implemented by defining a lower number of contacts in and around the epidemic area. The results are the same as for the barrier strategy in the FHP model. Too weak quarantining measures can deliver a faster course of the epidemic than we would reach with a generally lower number of contacts. Combinations of vaccination and quarantining policies could for example consist of an area with a low number of contacts, surrounded by a barrier of vaccinated individuals. Of course combinations can deliver much better results for confining an outbreak.

4.4 Towards an extended model

In the section on LGCA models it was stated that the number of contacts for this type of model is limited due to the structure of the lattice. To obtain a model, which considers motion of individuals but also allows a flexible definition of the contact environment, the FHP and stochastic CA approach can be combined by allowing stochastically distributed contacts among particles in the FHP model.

5 PDE model

The partial differential equations approach we want to observe here can be found in [8]. It is based on distributed contacts like the stochastic CA approach and emerges from a Taylor series simplification that can be found in [9] for example, which also refers to [8].

5.1 Derivation

Spatial inhomogeneities are introduced by distributing the individuals on a domain and observing the densities of each group $S(t, x, y)$, $I(t, x, y)$ and $R(t, x, y)$ respectively their change in time. For every location

(x, y) on the domain $S(t, x, y) + I(t, x, y) + R(t, x, y)$ is always ≤ 1.

Further a rate or probability of infection for every location is necessary. The basis for this is an interaction coefficient $\tilde{\alpha}(\vec{x}, \vec{v})$ for each two points $\vec{x} = (x, y)$ and $\vec{v} = (v, w)$ on the domain, which allows contact between individuals, which are located at different positions. As already mentioned in the section on stochastic CA, a prototype for the likelihood of interaction is a function that depends on the distance between the two points $\|\vec{v} - \vec{x}\|$. We only observe interaction functions that satisfy $\|\tilde{\alpha}\|_{L^1(\mathbb{R}^2)} = \alpha \leq 1$ and thus have similar features like probability distributions (compare [8]).

With the assumptions from above, the probability of infection Ψ at a certain point $\vec{x}$ is

$$\int \tilde{\alpha}(\vec{x}, \vec{v}) I(\vec{v}) \, d\vec{v} = \int \tilde{\alpha}(\|\vec{v} - \vec{x}\|) I(\vec{v}) \, d\vec{v} \qquad (7)$$

and after transformation with $\vec{z} := \vec{v} - \vec{x}$ becomes

$$\int \tilde{\alpha}(\|\vec{z}\|) I(\vec{x} + \vec{z}) \, d\vec{z}. \qquad (8)$$

Now in [9] the zeroth- and second-order Taylor series expansion of $I(\vec{x} + \vec{z})$ is observed in order to simplify the resulting integral equations.

$$\begin{aligned} I(x + z) &= I(x) + I'(x)z + \frac{I''(x)}{2}z^2 + \quad (9) \\ &\quad + \frac{I'''(x)}{6}z^3 + \frac{I^{iv}(x)}{24}z^4 + \ldots \end{aligned}$$

For functions on a two dimensional space ($\vec{x} = (x, y)$ and $\vec{z} = (u, v)$) this is equal to

$$\begin{aligned} I(\vec{x} + \vec{z}) &= I + uI_x + vI_y + \qquad (10) \\ &\quad + \frac{1}{2}(u^2 I_{xx} + 2uv I_{xy} + v^2 I_{yy}) + \ldots \end{aligned}$$

After simplification of the integral equation with the zeroth-order expansion the probability of infection is the same as in the original ODE model: $\Psi_0 = \alpha I$.

Using the second-order expansion delivers

$$\Psi_2 = I \int \tilde{\alpha} dz + \frac{1}{2}I_{xx} \int u^2 \tilde{\alpha} dz + \frac{1}{2}I_{yy} \int v^2 \tilde{\alpha} dz \quad (11)$$

for the probability of infection. Because of the radial symmetry of the likelihood function $\tilde{\alpha}$ some of the terms can be eliminated

$$0 = I_x \int u\tilde{\alpha} dz = I_y \int v\tilde{\alpha} dz = I_{xy} \int uv\tilde{\alpha} dz. \quad (12)$$

This is also the reason for $\Psi_1 = \Psi_0$.

Furthermore we only observe domains with finite range and thus can say that $r = \sqrt{u^2 + v^2} \leq \rho$ and $\tilde{\alpha}(r) = 0$ for $r > \rho$. This allows to define a parameter γ with a finite number (transformation into polar coordinates)

$$\gamma = \frac{1}{2} \int u^2 \tilde{\alpha} dz = \frac{\pi}{2} \int_0^\rho r^3 \tilde{\alpha}(r) dr. \qquad (13)$$

Using this substitution the probability of infection can be written as

$$\Psi_2 = I\alpha + I_{xx}\gamma + I_{yy}\gamma = \alpha I + \gamma\Delta I. \qquad (14)$$

For a second order simplification of the probability of infection at a point $\vec{x}$, a (new) partial differential equations system, which describes the densities of the three groups can be set up

$$
\begin{aligned}
\frac{\partial S(t,\vec{x})}{\partial t} &= -S(t,\vec{x})(\alpha I(t,\vec{x}) + \gamma\Delta I(t,\vec{x})) \\
\frac{\partial I(t,\vec{x})}{\partial t} &= S(t,\vec{x})(\alpha I(t,\vec{x}) + \gamma\Delta I(t,\vec{x})) - \\
&\quad -\beta I(t,\vec{x}) \\
\frac{\partial R(t,\vec{x})}{\partial t} &= \beta I(t,\vec{x}). \qquad (15)
\end{aligned}
$$

With the Hölder Inequality and because

$$
\begin{aligned}
\alpha &= \|\tilde{\alpha}\|_{L^1(\mathbb{R}^2)} = \int \tilde{\alpha}\,dz = \qquad (16) \\
&= 2\pi\int \tilde{\alpha}(r)r\,dr = 2\pi\|r\tilde{\alpha}\|_{L^1[0,\rho]}
\end{aligned}
$$

we can calculate an upper bound for γ

$$
\begin{aligned}
\gamma &= \frac{\pi}{2}\cdot\|r^3\tilde{\alpha}(r)\|_{L^1[0,\rho]} \qquad (17) \\
&\leq \frac{\pi}{2}\cdot\|r^2\|_{L^\infty[0,\rho]}\cdot\|r\tilde{\alpha}(r)\|_{L^1[0,\rho]} \\
&= \frac{\pi}{2}\cdot\rho^2\cdot\frac{1}{2\pi}\|\tilde{\alpha}\|_{L^1(\mathbb{R}^2)} \\
&= \frac{1}{4}\rho^2\alpha.
\end{aligned}
$$

5.2 Deriving an extended version

Using the third order expansion delivers nothing new because of the symmetry of $\tilde{\alpha}$ all terms with third order derivatives vanish $\Psi_3 = \Psi_2$.

The fourth order Taylor series expansion yields to

$$
\begin{aligned}
\Psi_4 &= I\int\tilde{\alpha}\,dz + \qquad (18) \\
&+ \frac{1}{2}I_{xx}\int u^2\tilde{\alpha}\,dz + \frac{1}{2}I_{yy}\int v^2\tilde{\alpha}\,dz + \\
&+ \frac{1}{24}I_{xxxx}\int u^4\tilde{\alpha}\,dz + \frac{1}{24}I_{yyyy}\int v^4\tilde{\alpha}\,dz + \\
&+ \frac{1}{4}I_{xxyy}\int u^2 v^2\tilde{\alpha}\,dz.
\end{aligned}
$$

If $\gamma_0 := \alpha$ and $\gamma_2 := \gamma$ are the same as for the second order expansion, and we define γ_4 so that

$$
\begin{aligned}
\gamma_4 &= \frac{1}{24}\int u^4\tilde{\alpha}\,dz = \frac{\pi}{32}\int r^5\tilde{\alpha}(r)\,dr, \quad (19) \\
2\gamma_4 &= \frac{1}{4}\int u^2 v^2\tilde{\alpha}\,dz = \frac{\pi}{16}\int r^5\tilde{\alpha}(r)\,dr,
\end{aligned}
$$

we can write the probability of infection (18) as

$$
\begin{aligned}
\Psi_4 &= \gamma_0 I + \gamma_2\Delta I + \gamma_4(I_{x^4} + I_{y^4} + 2I_{x^2y^2}) = \\
&= \gamma_0 I + \gamma_2\Delta I + \gamma_4\Delta\Delta I. \qquad (20)
\end{aligned}
$$

Actually this is only possible if the surface $I(x,y)$ is smooth enough. But because we observe only discrete points we can assume an arbitrary interpolation.

As before for $\gamma = \gamma_2$ we can now find upper bounds for γ_4.

$$
\begin{aligned}
\gamma_4 &\leq \frac{1}{16}\rho^2\gamma_2. \qquad (21) \\
\gamma_4 &\leq \frac{1}{64}\rho^4\gamma_0.
\end{aligned}
$$

Generally the probability of infection can be approximated with the Taylor series expansion of any arbitrary order n, which delivers a differential equations system of the form

$$
\begin{aligned}
\frac{\partial S(t,\vec{x})}{\partial t} &= -S(t,\vec{x})\Psi_n(t,\vec{x}) \\
\frac{\partial I(t,\vec{x})}{\partial t} &= S(t,\vec{x})\Psi_n(t,\vec{x}) - \beta I(t,\vec{x}) \\
\frac{\partial R(t,\vec{x})}{\partial t} &= \beta I(t,\vec{x}), \qquad (22)
\end{aligned}
$$

where $\Psi_n(t,\vec{x})$ is the probability of infection at a certain point $\vec{x}$ at time t, which has been approximated using the n-th order Taylor series expansion.

Now we will determine and analyse Ψ_n in order to examine the behaviour for $n\to\infty$.

For $\vec{x} = (x,y)$ and $\vec{z} = (u,v)$ the n-th order Taylor series expansion of $I(\vec{x}+\vec{z})$ can be written as

$$I_n(\vec{x}+\vec{z}) = \sum_{i+j\leq n}\frac{\partial^i}{\partial x^i}\frac{\partial^j}{\partial y^j}\frac{I(x,y)}{i!j!}u^i v^j. \qquad (23)$$

Because of the radial symmetry of $\tilde{\alpha}$ all terms for which $i\cdot j$ is an odd number vanish in the integral representation of the probability of infection Ψ_n. This means that we have

$$
\begin{aligned}
\Psi_n(t,\vec{x}) &= \int\tilde{\alpha}(\|\vec{z}\|)I_n(\vec{x}+\vec{z})\,d\vec{z} = \qquad (24) \\
&= \int\tilde{\alpha}(\|\vec{z}\|)\sum_{\substack{i+j\leq n \\ i,j\in 2\mathbb{N}}}\frac{\partial^i}{\partial x^i}\frac{\partial^j}{\partial y^j}\frac{I(x,y)}{i!j!}u^i v^j\,d\vec{z} = \\
&= \sum_{\substack{k=0 \\ i+j=k}}^{\lfloor\frac{n}{2}\rfloor}\frac{\partial^{2i}}{\partial x^{2i}}\frac{\partial^{2j}}{\partial y^{2j}}\frac{I(x,y)}{(2i)!(2j)!}\int\tilde{\alpha}(\|\vec{z}\|)u^{2i}v^{2j}\,d\vec{z}.
\end{aligned}
$$

Now we claim that

$$\Psi_n(t,\vec{x}) = \sum_{k=0}^{\lfloor\frac{n}{2}\rfloor}\gamma_{2k}\Delta^k I(x,y) \qquad (25)$$

where

$$
\begin{aligned}
\gamma_{2k} &= \frac{1}{(2k)!}\int\tilde{\alpha}(\|\vec{z}\|)u^{2k}\,d\vec{z} \qquad (26) \\
&= \frac{2\pi}{(2k)!}\prod_{l=1}^{k}\frac{2l-1}{2l}\int_0^\rho\tilde{\alpha}(r)r^{2k+1}\,dr.
\end{aligned}
$$

The second equal sign follows from the Wallis product and the periodicities of the cosine function. Furthermore we can develop a relation between γ_{2k} and $\gamma_{2(k-1)}$ respectively γ_0 ($k > 0$).

$$\gamma_{2k} \leq \frac{1}{4k^2}\rho^2\gamma_{2(k-1)} \qquad (27)$$

$$\gamma_{2k} \leq \frac{1}{4^k(k!)^2}\rho^{2k}\gamma_0$$

Our proposition (25) is true because

$$\Delta^k I(x,y) = \sum_{j=0}^{k}\binom{k}{j}\frac{\partial^{2(k-j)}}{\partial x^{2(k-j)}}\frac{\partial^{2j}}{\partial y^{2j}}I(x,y) \quad (28)$$

and (γ_{2k} is involved!)

$$\frac{1}{(2(k-j))!(2j)!}\int_0^{2\pi}\cos^{2k-2j}\varphi\cdot\sin^{2j}\varphi\,d\varphi =$$
$$= \frac{\binom{k}{j}}{(2k)!}\int_0^{2\pi}\cos^{2k}\varphi\,d\varphi, \qquad (29)$$

which we can apply after rearranging (24) to a double series.

On the other hand we know that the Taylor series expansion converges locally towards the function value (actually depending on the smoothness of the function), which leads to

$$\lim_{n\to\infty}\sum_{k=0}^{\lfloor\frac{n}{2}\rfloor}\gamma_{2k}\Delta^k I(x,y) = \int\tilde{\alpha}(\|\vec{z}\|)I(\vec{x}+\vec{z})\,d\vec{z}. \quad (30)$$

5.3 Implementation

First we observe the PDE model with the second order Taylor series approximation. We need to approximate the Laplacian at every point of the lattice (and also use a shorter notation).

$$\Delta I(x,y) = I(x,y)_{xx} + I(x,y)_{yy} = \qquad (31)$$
$$= \frac{I(x+\varepsilon,y)+I(x-\varepsilon,y)-2I(x,y)}{\varepsilon^2} +$$
$$+ \frac{I(x,y+\varepsilon)+I(x,y-\varepsilon)-2I(x,y)}{\varepsilon^2} =$$
$$= \frac{I(x+\varepsilon)+I(x-\varepsilon)+I(y+\varepsilon)+I(y-\varepsilon)-4I}{\varepsilon^2}$$

Because we have a discrete lattice, the minimal value for ε is 1. This means that in the case of Ψ_2 the solution procedure for the PDE uses a Van Neumann neighbourhood of size one to calculate the growth of the density of infected in the cell.

For the fourth order approximated PDE model we additionally need to know

$$\Delta\Delta I(x,y) = \Delta(I(x,y)_{xx} + I(x,y)_{yy}), \qquad (32)$$

which we can calculate from $\{\Delta I(u,v),\,|x-u|+|y-v| \leq 2\}$. We obtain a Van Neumann neighbourhood of size two for Ψ_4.

Very interesting are the weight factors for the cells in the neighbourhood, which are determined by the γ_{2k} (Fig. 4). This also explains why, these parameters can

...			
γ_4	...		
γ_2 $-4\gamma_4$	$2\gamma_4$	...	
γ_0 $-4\gamma_2$ $+4\gamma_4$	γ_2 $-4\gamma_4$	γ_4	...

Fig. 4 Weight factors in the neighbourhood for the fourth-order simplified PDE implementation (upper-right quadrant).

not be totally arbitrary but must satisfy the conditions (21) and (27). Otherwise problems arise at the latest in the implementation, especially if all densities should be ≤ 1.

Because of (30) we know that these CA approximations converge for $n \to \infty$ towards the integral representation (8) of the model. Therefore and because we always observe discrete domains, the approximation of Ψ and consequently the calculation of the Laplacian for each location is not necessary. Instead we can calculate the probability of infection, which delivers the growth of the density of infected (and also the growth of the probability of a cell being infected as we will see in the following section), by

$$\Psi(x,y) = \sum_{\substack{(u,v)\in \\ \Omega_{(x,y)}}}\tilde{\alpha}(\|(u,v)-(x,y)\|)I(u,v) \quad (33)$$

where $\Omega_{(x,y)}$ is the set of all cells, which lie in the neighbourhood of the cell located at (x,y).

5.4 Modified solution approach and comparison

Beneath the just mentioned solution method, which we will use as a standard and which provides the densities of the three groups on the lattice (domain), we will introduce a modification, which uses the distribution of the population to create a state-discrete representation of the system.

With the help of a uniformly distributed random number for each cell, which is constant over time, a discrete state can be calculated by simply locating the random number in the segmentation of $[0,1] = [0,S] \cup (S,S+I] \cup (S+I,S+I+R] \cup (S+I+R,1]$. Now this means that Ψ describes the growth of the probability of the cell being infected.

The behaviour of this lattice representation during evaluation can be compared to one of a stochastic CA simulation. In the following we provide the theoretical link between these two approaches and therefore analyse the feature of distributed interaction in both models using a

special likelihood function $\tilde{\alpha} \in L^1$.

$$\tilde{\alpha}(r) = \begin{cases} c \cdot \exp\left(\frac{r^2}{d(r^2-\rho^2)}\right), & r < \rho \\ 0, & r \geq \rho \end{cases} \qquad (34)$$

With $c \in (0,1]$ it is possible to directly define the maximum of the function, which is always located at $r = 0$. $d \in (0,1]$ controls the decay of the function. Bigger values mean slower decay. And finally ρ is the radius of the circular neighbourhood. These parameters allow to generate likelihood functions, which satisfy $\|\tilde{\alpha}\|_{L^1} \leq 1$.

Because for both models $\tilde{\alpha}$ defines neighbourhoods of the same maximum size (ρ) and structure (c, d), the distinct difference lies in the employment of the likelihood function (compare Tab. 1).

For the PDE model $\tilde{\alpha}$ delivers a factor for each neighbourhood cell (compare γ_{2k}), which determines the amount of impact that they have on the central spot. Depending on the sum of all impacts, the densities of infected in the discrete neighbourhood and the density of susceptibles in the central cell, the growth of the (density of the) infected is performed. According to this density with the earlier mentioned constant random number, a discrete state for the lattice representation can be obtained. This also means that the probability of the central cell being infected (in the next time step) is

$$\Phi_{\text{PDE}} = I + \Psi \cdot S \qquad (35)$$

(As before I and S are the densities of infected respectively susceptibles in the central cell) minus the recoveries, which are neglected here for reasons of simplification but reintroduced later on/in the implementation without further mentioning.

In the stochastic CA on the other hand the likelihood function is used for a stochastic determination of λ contact cells within the neighbourhood in order to provide the basis for – now in contrast not only for a representation – generating the infections with a further constant parameter α, which may vary for different cells. Notice that α must not necessarily be the same like α from the previous section, to which we now refer to as γ_0. For the probability of the cell being infected (in the next time step), which we denote Φ_{CA}, this yields to

$$\Phi_{\text{CA}} = (1 - (1-\alpha)^{I_{\Omega_\lambda}}) \cdot S \qquad (36)$$

as it has been discussed earlier in this paper (I_{Ω_λ} is the number of infected contact cells and S is one if the cell is susceptible otherwise it is zero).

Like in the section on stochastic CA α represents the probability of infection when a susceptible and an infected individual come into contact and λ is the number of contacts within the neighbourhood per time step. Regard the strong connection between those two parameters, which is characterised by an indirect proportional relation. If more contacts are established, the impact of a single contact must be smaller in order to preserve

the same overall level of infectious impacts. Concerning runtime of an implementation, it is important that λ is not too large, because a high number of contacts requires a higher number of runs through the loop, which provides the random contact generation. If λ is a larger number on the other side, the interaction with the neighbourhood becomes 'smoother', which could (but actually does not – see later on) provide a behaviour of the CA closer to the pseudo continuous PDE model.

Now the remaining parameter for an identification is α.

An optimal relation between the probabilities of infection would be valid for all possible conditions concerning the neighbourhood.

$$\Phi_{\text{CA}} \approx \Phi_{\text{PDE}} \qquad (37)$$
$$S(1 - (1-\alpha)^{I_{\Omega_\lambda}}) \approx I + S \cdot \Psi$$
$$\approx I + S \sum_{u^2+v^2<\rho^2} \tilde{\alpha}(u,v)I(u,v)$$
$$\approx I + S \int \tilde{\alpha}(u,v)I(u,v)\,d\vec{z}$$

Because this might not be possible, some features of the prevailing conditions must influence the derivation of the probability of infection in the stochastic CA α. What can be assumed is that $S = 1$ and $I = 0$, because otherwise it would not be clear that the cell has not already status infected in the lattice representation. Further an estimation of the density of infected in the neighbourhood is needed $I(u,v) \approx \bar{I}$. This leads to $I_{\Omega_\lambda} \approx \bar{I}\lambda$.

After applying these assumptions on (37) and again using a Taylor series simplification (especially if α is small and λ is large) the equation can be solved for α.

$$1 - (1-\alpha)^{\bar{I}\lambda} \approx \bar{I}\|\tilde{\alpha}\|_{L^1(\mathbb{R}^2)}$$
$$\alpha\bar{I}\lambda \approx \bar{I}\|\tilde{\alpha}\|_{L^1(\mathbb{R}^2)}$$
$$\alpha \approx \frac{\|\tilde{\alpha}\|}{\lambda}. \qquad (38)$$

On the first look this method seems to be rather imprecise.

Now we present a second one, which actually works only for likelihood functions that are probability density functions and therefore discharge the idea of discrete states in the stochastic CA in the beginning and proceed with using densities.

$$\Phi_{\text{CA}} \approx \Phi_{\text{PDE}} \qquad (39)$$
$$I + S \cdot \sum_{\Omega_\lambda} \alpha I(u,v) \approx I + S \cdot \sum_{\Omega} \tilde{\alpha}(u,v)I(u,v)$$

If λ is large enough, we can assume $\Omega_\lambda = \Omega$ and we can use the same approximation as before because the contacts are distributed with $\tilde{\alpha}$.

Now on the contrary we assume that $\lambda = 1$. This means that one single contact cell with a specific density of infected I_1 is chosen. We want to have

$$\alpha I_1 \approx \int \tilde{\alpha}(u,v)I(u,v). \qquad (40)$$

Because the location of this contact cell follows the distribution $\tilde{\alpha}$, we can use the central limit theorem, which yields to

$$I_1 \;\to\; \int \tilde{\alpha}(u,v)I(u,v) \tag{41}$$

and accordingly $\alpha \to 1$ if the process of contact-establishment is repeated very often. And with the same idea we obtain for $\lambda > 1$ that $\alpha \to \frac{1}{\lambda}$, which corresponds to the first approach.

Now we must remember that we used a 'probability' of infection of the form αI. But because we know from the previously used Taylor series expansion that for small α and accordingly larger λ we can write $\alpha I_{\Omega_\lambda} \approx 1 - (1-\alpha)^{I_{\Omega_\lambda}}$, this results are also valid for the original contagion process, which was described by the latter form of the probability of infection.

With this parameter identification it is possible to generate rather good correspondence between the two different approaches as we will see later.

5.5 Remarks

As it can be seen in (37) for example, the space- (and time-) discrete implementation of the PDE model uses Riemann-Stieltjes-Sums to approximate the integral over $\tilde{\alpha}I$.

A closer look on the bounding box algorithm, which has been used in the implementation of the stochastic CA model, shows that in this case the inverse of the likelihood function applied on equidistant values delivers a special sequence of radii, which then provide by themselves the decaying likelihood of interaction. This method can be compared to the Lebesgue method of integration, which seeks step functions f that satisfy $\int f \leq \int \tilde{\alpha}$. On the other hand the process of determining an interaction area within the neighbourhood with a certain (decaying) likelihood of contact is equivalent to calculating quantiles of the probability distribution $\tilde{\alpha}$.

We recapitulate that for both models the process of propagation of a disease is composed of two components, either stochastic or deterministic (Tab. 1).

Tab. 1 Deterministic and stochastic components in the PDE and stochastic CA model. [1]This formulation is of course only valid for likelihood functions, which are probability density functions. [2]Actually only the lattice representation is generated stochastically by using the (deterministic) density and a constant random number.

	PDE	stoch. CA
construction of neighbourhood resp. contacts	deterministic by Riemann sums	stochastic by quantiles[1]
reaction on contagious contact	semi-stochastic using densities[2]	stochastic by probability of infection

Concerning the behaviour of the stochastic CA we can assume that for higher values of λ the behaviour must

lie closer to the PDE approach because the impact of contagious cells is closer to $\int \tilde{\alpha}I$. That this is not the case shows Fig. 5.

Further, if in the discretisation of the second order Taylor series simplified PDE only discrete states are allowed, the resulting CA corresponds to the Van Neumann CA, which was used in section 4.2 to establish a connection between the HPP LGCA and stochastic CA approach.

6 Diffusion – Brownian motion

We now present a method for adapting the behaviour of the PDE and stochastic CA approach to the behaviour of the LGCA model. As a basic connection we use that the diffusion of particles governed by random motion (random walk) can be compared to Brownian motion – at least in a scaling limit.

Brownian motion can be described by the diffusion or Gaussian semigroup, which is defined by [10]

$$T(t)f(\vec{x}) = \frac{1}{4\pi t} \int \exp\left(\frac{-\|\vec{z}\|^2}{4t}\right) f(\vec{x}+\vec{z})\,d\vec{z} \tag{42}$$

This invokes a likelihood function

$$\tilde{\alpha}(r) = \frac{1}{4\pi} e^{-\frac{r^2}{4}} \tag{43}$$

for the PDE and stochastic CA approach, which only defers little form the previously used likelihood functions. Now the radius of interaction is not bounded. This function also describes the mean square deviance of the particles.

We use this likelihood function in the PDE approach and adapt the parameters of the stochastic CA and LGCA to fit the PDE behaviour.

For the stochastic CA we know from the previous section how to choose α and λ.

6.1 Parameter identification for the LGCA model

We now identify positions on the domain for the PDE approach and cells for the stochastic CA with cells in the LGCA, which means that we do not observe the number of particles (I_{abs} and S_{abs}) but their densities within the cells (I and S). This allows again to calculate a discrete status for every cell by a constant random number.

The probability of a cell being infected in the FHP LGCA is given by (compare (2))

$$\begin{aligned}
\Phi_{\text{FHP}} &= \frac{I_{\text{abs}} + (1-(1-\alpha)^{I_{\text{abs}}}) \cdot S_{\text{abs}}}{6} \tag{44} \\
&= I + (1-(1-\alpha)^{6I}) \cdot S \\
&\approx I + 6\alpha \cdot I \cdot S.
\end{aligned}$$

Further the particles in the cells depend on the transition mechanism (random motion), which is described by $\tilde{\alpha}$ (random walk).

It is important to know that the variance of the diffusion distribution $\tilde{\alpha}$ describes the mean square deviation of particles.

$$
\begin{aligned}
\sigma &= \int u^2 \tilde{\alpha}(u,v)\,d\vec{z} \qquad (45)\\
&= \pi \cdot \int r^3 \tilde{\alpha}(r)\\
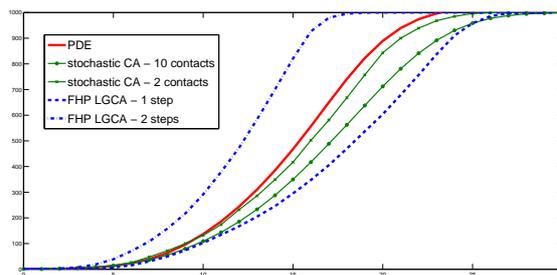&= 2\gamma_2\\
&= 2
\end{aligned}
$$

Fig. 5 Infected individuals from the PDE, stochastic CA and FHP LGCA approach. The stochastic CA was run with 10 contacts per time step and 2 contacts per time step. Accordingly the infection rates in the stochastic CA were $\frac{1}{10}$ and $\frac{1}{2}$ respectively. The FHP simulation was run with one movement per time step and two movements per time step. The infection rates for the LGCA were therefore $\frac{1}{3}$ respectively $\frac{1}{6}$.

This means that during one time unit a particle moves for two space units. Accordingly for a comparison of the LGCA with the PDE approach the particles in the LGCA must perform two discrete movements during one time step. But because this means that the number of contacts doubles, we can alternatively divide the probability of infection α by two (compare [11]) and use one discrete movement per time step in the LGCA.

Further Graham's law states that diffusion is indirect proportional to $\sqrt{\omega}$, where ω is the density of a population or a gas.

Now the probability of infection can be written as

$$
\begin{aligned}
\Psi_{\text{FHP}} &\approx 6 \cdot \frac{\alpha}{2} \cdot I \qquad (46)\\
&\approx 3\alpha \cdot \frac{1}{\sqrt{\omega}} \int \tilde{\alpha}(u,v) I(u,v)\,d\vec{z}.
\end{aligned}
$$

Comparing the probabilities of infection in both models leads to

$$
\begin{aligned}
\Psi_{\text{FHP}} &\approx \Psi_{\text{PDE}} \qquad (47)\\
\frac{3\alpha}{\sqrt{\omega}} \int \tilde{\alpha} I &\approx \int \tilde{\alpha} I\\
\alpha &\approx \frac{\sqrt{\omega}}{3}.
\end{aligned}
$$

6.2 Monte Carlo simulation

The following experiment (Fig. 5) happens on a domain of 1000 cells/individuals/positions with periodic boundary conditions. The initial condition is one (absolutely) infected cell in the centre of the lattice, all other cells are fully susceptible (full density). For reasons of simplification only infections and no recoveries are observed. The time of observation is 30 time units. The outcome is an averaged one over 20 simulation runs (except for the PDE because it is fully deterministic).

6.3 Remarks

Simulations show that also the behaviour of LGCA with FHP-I transition rules (not shown in the figures) can be compared to the PDE model with this special likelihood function.

The LGCA simulations require extremely less computation time. The CA is fast if only few susceptible cells are on the domain. PDE simulations (no Taylor series
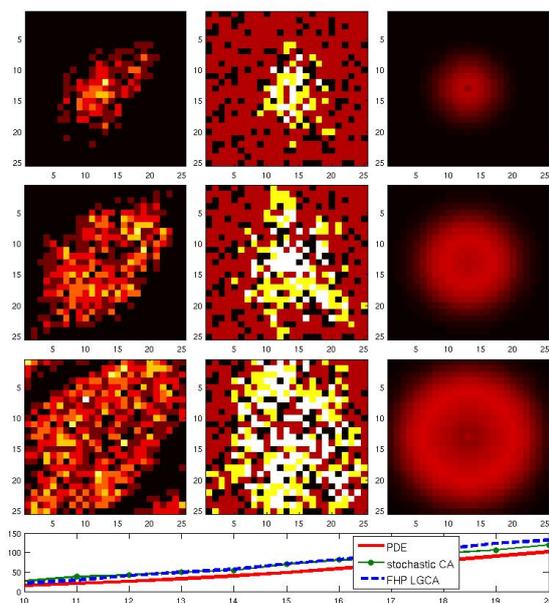

Fig. 6 The density of infected individuals from the PDE, stochastic CA and FHP approach (SIRS) on the domain after 10, 15 and 20 time steps (from left to right and the densities are increasingly represented by black-red-yellow-white except for the stochastic CA where infected cells are marked yellow). Initial condition is one infected cell in the centre of the lattice. The population consists of 500 susceptible individuals who are uniformly distributed on a domain with a density of 0.8. The rate of infection is 0.8, the recovery rate is 0.2 and the rate for loss of immunity is 0.1.

simplification) are slower – actually depending on the implementation.

It is surprising that the behaviour of the stochastic CA with two contacts per time step is closer to the PDE model than with 10 contacts, which is closer to the LGCA model. The reason for slower spread with 10 contacts could be that for a higher number of contacts more contacts are established with other susceptible individuals.

How the population density effects the behaviour of the LGCA is already known from section 3.2. Consequently the reason for why the probability of infection (46) additionally contains the population density is that otherwise a lower density would influence the overall rate of infections in the LGCA twice. In the PDE model the population density only influences the values of S, I and R. In both situations no further parameter tuning is necessary. For the stochastic CA the infection rate should be divided by the population density because the number of contacts in the calculation of the probability of infection is linked to the population density and is smaller for lower densities (compare (37) and (38)). Accordingly the reason for the influence of the population density is the same as for the LGCA approach.

For a different rate of infection the probability of infection in each of the three approaches must simply be multiplied with a parameter.

Because of the central limit theorem and the diffusion distribution from the previous sections we can say that the PDE model is an upper bound for the stochastic CA and random motion FHP LGCA concerning the symmetry and smoothness of the contact behaviour respectively motion.

7 Frequency analysis

In this section we compare the occurrence of epidemic waves in different SIRS model approaches. In section 1 it was mentioned that by using a rate (or actually probability) of recovery and a random number, which decides whether an infected individual recovers in the current time step or not, the duration of the infection is geometrically distributed. For the period of immunity (before the individuals become susceptible again) in the SIRS extension the same idea is valid if a rate for the loss of immunity is used. Principally it is also possible to assign an arbitrary distribution to the incubation, contagion or immunity period as done in section 7.2.

7.1 PDE, stochastic CA and FHP SIRS models

The simulations with the SIRS extensions with geometrically distributed durations (Fig. 7 and Fig. 8) additionally show that the parameter identifications from the previous sections deliver very accurate results.

Initial condition for the simulations in Fig. 7 and Fig. 8 is one infected cell in the centre of the lattice. The population consists of 500 susceptible individuals who are uniformly distributed on a domain with full density.

The rate of infection is 1, the recovery rate is 0.3 and the rate for loss of immunity is 0.05.
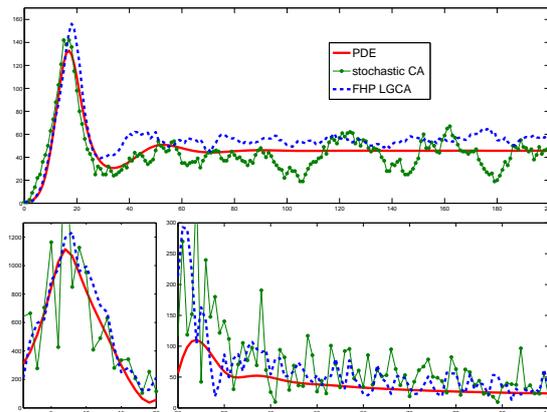

Fig. 7 above: Infected individuals from a single simulation run with the PDE, stochastic CA and FHP LGCA approach. below: FFT frequency analysis with two different scalings.
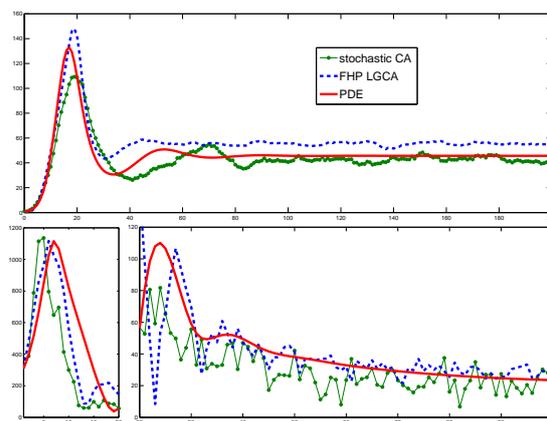

Fig. 8 above: Average number of infected individuals from 16 simulation runs with the PDE, stochastic CA and FHP LGCA approach. (Results of simulations in which the epidemic dies out were not used). below: FFT frequency analysis.

7.2 SICRS – extended model

The model, which we use for the following simulations, was mentioned earlier. It is a composition of the LGCA and stochastic CA approach and thus features distributed contacts and motion. But these are not the decisive factors for the generation of the wavelike behaviour of an epidemic, as it is the duration of the stages in which the individuals rest (susceptible, infected, contagious, immune/recovered) and the breaks between position changes.

The durations of the stages are normally distributed with the following parameters (Tab. 2).

We use FHP-I collision configurations and allow motion only every 10 time steps but permit 5 discrete

Tab. 2 Expected values and variances for the durations of the disease stages.

stage	expected duration	variance
incubation period	6	1
contagious period	14	4
resistant period	21	5

movements at once. The boundary conditions are periodic. The distribution of contacts is determined by a likelihood function of the form (34) with parameters $\rho = 5$ (radius), $d = 0.4$ (decay) and $c = 1$ (maximum). The number of contacts per time step is 39 and the probability of infection 0.2. The domain is discretised on a lattice of $15\,000$ cells and the population density is 0.3. The initial condition consists of 54 infected particles in the centre of the lattice. The outcomes of the simulation after 800 time steps are presented in Fig. 9 and a Monte Carlo simulation with 10 runs and 400 time steps can be found in Fig. 10.
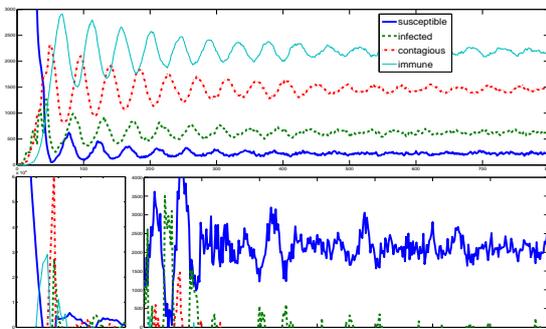

Fig. 9 above: Number of susceptible, infected, contagious and immune individuals from a single simulation run with the extended model. below: FFT frequency analysis.
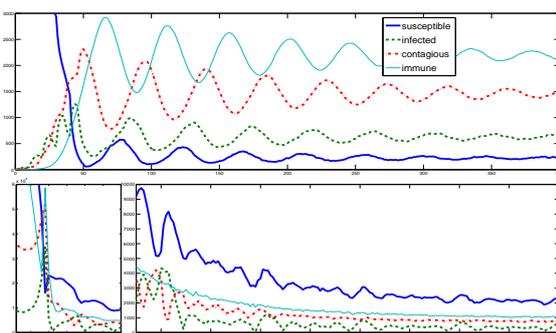

Fig. 10 above: Number of susceptible, infected, contagious and immune individuals from a Monte Carlo simulation with the extended model. below: FFT frequency analysis.

7.3   Remarks

Especially the first two enhancements in the number of infected individuals (Fig. 7) are very similar in all three

approaches. In contrast to the PDE model, the FHP LGCA and especially the stochastic CA model deliver strong fluctuations (compare Brownian motion) in the number of infected. The Monte Carlo simulations (Fig. 8) show that the expected or mean development of the number of infected can be described by the PDE model, which confirms that the PDE model is an 'upper bound concerning smoothness'.

Fig. 9 and Fig. 10 show that the extended model delivers clearly visible waves and that it ends up in a steady state. By applying Monte Carlo methods, small fluctuations can be eliminated whereas the wavelike behaviour is preserved.

8   References

[1] G. Schneckenreither, F. Breitenecker. *CA models for SIR type epidemics*. SNE 16/3, p. 27-36, December 2006.

[2] D. Smith, L. Moore. *The SIR Model for Spread of Disease*. www.math.duke.edu/education

[3] D. A. Wolf-Gladrow. *Lattice-Gas Cellular Automata and Lattice Boltzmann Models*. Springer, 2000.

[4] H. Hötzendorfer, N. Popper, F. Breitenecker. *Temporal and Spatial Evolution of a SIR-type Epidemic – ARGESIM Comparison C17 – Definition*. SNE 41/42, p. 42-45, December 2004.

[5] H. Fukś, A. Lawniczak. *Individual-based Lattice Model for Spatial Spread of Epidemics*. Discrete Dynamics in Nature and Society, vol. 6, no. 3, pp. 191-200, 2001.

[6] Š. Emrich. *Diploma Thesis – Comparison of Mathematical Models and Development of a Hybrid Approach for the Simulation and Forecast of Influenza Epidemics within Heterogeneous Populations*. Institute for Analysis and Scientific Computing, Vienna University of Technology, 2007.

[7] S. Venkatachalam, A. R. Mikler. *Towards Computational Epidemiology: Using Stochastic Cellular Automata in Modeling Spread of Diseases*. Proceedings of the 4th Annual International Conference on Statistics, Mathematics and Related Fields, Honolulu, HI, January 2005.

[8] Jones, D. S., B. D. Sleeman. *Differential Equations and Mathematical Biology*. George Allen & Unwin, London, 1983

[9] J. Callahan. *The spread of an infection through a region*. http://maven.smith.edu/~callahan/ili/sir2d.pdf

[10] M. Blümlinger, M. Kaltenbäck, H. Langer, M. Langer. *Skriptum zur Vorlesung Funktionalanalysis 2*. p. 62. Institute for Analysis and Scientific Computing, Vienna University of Technology, 2003/2004.

[11] H. Hötzendorfer, F. Breitenecker. *A Directly Programmed Implementation of ARGESIM Comparison C17 SIR-type Epidemic using Matlab*. SNE 41/42, p. 45, December 2004.