

A finite element framework for simulating residential burglary in realistic urban geometries

Baoli Hao¹, Kamrun Mily², Annalisa Quaini², and Ming Zhong²

¹ Department of Applied Mathematics, Illinois Institute of Technology, 10 West 35th St,
Chicago, IL 60616

`bhao2@hawk.illinoistech.edu`

² Department of Mathematics, University of Houston, 3551 Cullen Blvd, Houston, TX
77204, USA

`kmily@cougarnet.uh.edu`, `aquaini@central.uh.edu`, `mzhong3@central.uh.edu`

Abstract

We consider a partial differential equation (PDE) model to predict residential burglary derived from a probabilistic agent-based model through a mean-field limit operation. The PDE model is a nonlinear, coupled system of two equations in two variables (attractiveness of residential sites and density of criminals), similar to the Keller–Segel model for aggregation based on chemotaxis. Unlike previous works, which applied periodic boundary conditions, we enforce boundary conditions that arise naturally from the variational formulation of the PDE problem, i.e., the starting point for the application of a finite element method. These conditions specify the value of the normal derivatives of the system variables at the boundary. For the numerical solution of the PDE problem discretized in time and space, we propose a scheme that decouples the computation of the attractiveness from the computation of the criminal density at each time step, resulting in the solution of two linear algebraic systems per iteration. Through numerous numerical tests, we demonstrate the robustness and computational efficiency of this approach. Leveraging the flexibility allowed by the finite element method, we show results for spatially heterogeneous model parameters and a realistic geometry (city of Chicago). The paper includes a discussion of future perspectives to build multiscale, “multi-physics” models that can become a tool for the community. The robust and efficient code developed for this paper, which is shared open-source, is intended as the solid base for this broader research program.

Key words: Discrete and continuum crime models; Finite Element approximation; Partitioned methods; Pattern formation.

1 Introduction

Predictability and control of urban crime are fundamental for public safety. Despite advances in understanding the mechanisms driving crime spread and dissipation [17, 26, 24], it remains extremely hard to predict crime due to its complex nature. Prediction strategies based on mapped crime data have limited efficacy because crime distributions exhibit low regularity in both space and time and display significantly different patterns at different spatio-temporal scales [51, 36]. See Fig. 1, which shows the emergence of large burglary hotspots in the Houston area over three-month periods in 2024. These hotspots vary in lifetime and scale depending on geographic, economic, and seasonal factors. Also, different crime types exhibit distinct patterns of emergence, diffusion, and dissipation.

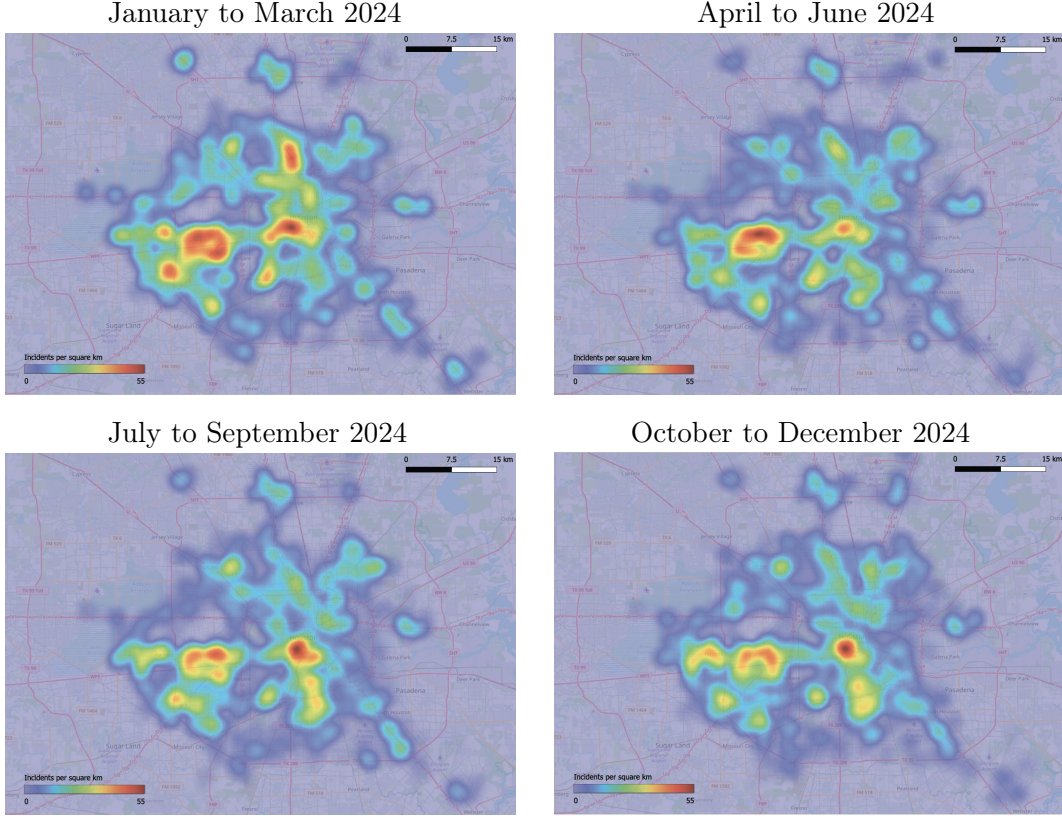


Figure 1: Number of burglaries per km^2 in the Houston area for four three-month periods in 2024. Police data are taken from a publicly available database [3] and visualized with open-source visualization software QGIS [4], using a radius of 2.5 Km to calculate the densities.

The need for better strategies to predict crime hotspots, and thus address or prevent them, has prompted mathematicians to develop crime models. In this paper, we focus on a specific kind of crime: residential burglary, which is simpler than other crimes since mobile offenders target stationary sites. This crime has been extensively studied in criminology and sociology, where researchers have observed how offenders are driven by environmental cues [52, 53] and respond to the success or failure of their activities [31, 54]. The pioneering paper that translated these studies into mathematical modeling is [47]. Therein, two types of models are proposed. The first type is a probabilistic agent-based model, where the interacting “agents” are offenders and housing units. The second type of model is derived from the first one through a mean-field limit operation and consists in a system of PDEs that describes the average behavior at large scale, i.e., for a large number of agents. Both the probabilistic model and the PDE model feature complex, possibly dynamic, parameters such as the attractiveness of residences [53] and strength of neighborhood effects [31, 52]. The use of these models for predictive policing, i.e., to suggest law enforcement where crime is most likely to occur, has attracted criticism [29, 23]. While the models can certainly be improved to reduce bias, we see them as the basis to develop a tool to assist municipalities in deciding where to direct community investments (instead of policemen) to address systemic needs and reduce crime.

A rigorous analysis of the hotspot formation in the PDE model from [47] is presented in [46]. In [45], a weakly nonlinear analysis is used to study hotspot formation, bifurcations and suppression of crime patterns in a simplified version of the PDE model from [47]. A study on the local existence and uniqueness of the solution to such simplified PDE model

is reported in [41]. The simplified PDE model contains a deterrence term for the effect of police deployment. The modeling of police intervention is improved in [55] through a dynamic optimal control problem, while the work in [53] incorporates spatial data about housing to improve the estimates on residential crime density. In [38], the Poisson process used in the agent-based model from [47] to represent the movement of criminal agents (i.e., random walk) is replaced by a truncated Lévy distribution. This leads to a different system PDEs with respect to the one in [47] and results for problems in one dimension are presented in [38]. A different agent-based model with police enforcement can be found in [32], while other chemotaxis and nonlinear diffusion models can be found in [34, 42, 15, 28]. In particular, reference [28] describes and analyzes a model for the interactions of two rivaling gangs attempting to establish or defend territories by spraying intimidating graffiti. The analysis includes construction of global, bounded classical solutions and proof of their convergence to homogeneous steady states, while the numerical results show steady-state solutions (in diffusion-dominated regime) and nontrivial asymptotic behaviors (in convection-dominated regime).

As it may be apparent from the literature review presented above, most models have been developed at the microscopic (i.e., individual-based) or macroscopic (i.e., based on continuum assumptions) scale. To the best of our knowledge, the only model developed at the mesoscopic scale, which is intermediate between the small and the large scales, is presented in [11]. This mesoscopic model incorporates tools from game theory to study how interactions between the general population, criminals, and policemen lead to patterns in criminal behavior. It is found that small, individual changes influenced by factors, such as wealth differences, can cause large, discernible differences in crime patterns. These factors are included in an additional variable, called activity. While the activity variable in [11] is constant in time for simplicity, it could evolve in time as a result of interactions and, consequently, affect the evolution of crime. We believe that more work is needed at the mesoscopic scale and, perhaps even more importantly, to integrate the scales in a multiscale vision. Indeed, only one representation scale is not sufficient to describe the overall collective dynamics of crime. See Sec. 5 for more details on these future perspectives.

More recently, data-driven methods have been developed. A deep learning technique called Spatio-Temporal Residual Networks (ST-ResNet) for real time crime forecasting is proposed in [50] and further developed in [49]. A statistics-based model is introduced in [21] and statistical study on the effect of social distancing during the COVID-19 pandemic on residential burglary and domestic violence is presented in [35]. Finally, reported gang-related violent crimes in Los Angeles from 2014 to 2017 are modeled in [39] using spatial-temporal marked Hawkes point processes with covariates.

The reader interested in more in-depth reviews of criminal behavior dynamics modeling, including data-driven approaches, is referred to [16, 25, 27].

In this paper, we consider the two original models from [47] and introduce several elements of novelty. First, we replace the periodic boundary conditions used in [47] with natural boundary conditions, i.e., conditions that specify the value of the normal derivatives of the system variables at the boundary. While periodic boundary conditions are convenient for the analysis, they assume an infinitely large domain and thus are of little practical use if one is interested in simulating crime in a real city. The boundary conditions we enforce arise naturally from the variational formulation of the PDE problem, which is the starting point for the application of a finite element method. As such, these boundary conditions are enforced automatically, i.e., without any external constraint. Of the references listed above, numerical results for the PDE model from [47] can be found only in [47, 46, 55]. These references adopt a spectral method to approximate the solution of the PDE model.

Spectral methods can be computationally less expensive and easier to implement than finite element methods. Moreover, they can be highly accurate in simple domains with smooth solutions. However, because of their global nature, they lead to linear systems with dense matrices. Thus, computational efficiency suffers as the size of the problem increases. For larger problems, finite element methods generally work better due to sparse matrices and more flexibility in handling non-trivial geometries, including better modeling of sharp bends. Because of our interest in simulating crime in real cities, we opt for a finite element-based solver for the PDE model.

The PDE model is a nonlinear, coupled system of two equations in two variables (attractiveness of residential sites and density of criminals), similar to the Keller–Segel model for aggregation based on chemotaxis [33, 20, 19]. To learn more about chemotaxis models and cross-diffusion models in complex environments (i.e., interacting with a time-evolving external system), the reader is referred to [15]. For the numerical solution of the PDE model from [47], we propose a scheme that decouples the computation of the attractiveness from the computation of the criminal density at each time step, resulting in the solution of two linear algebraic systems per iteration. Through numerous numerical tests, we demonstrate the robustness and computational efficiency of this approach. Since we cannot compare with the results from [47, 46, 55] due to the difference in boundary conditions, we validate the results obtained with the PDE model against the results provided by the agent-based model. Finally, leveraging the flexibility allowed by the finite element method, we show results for spatially heterogeneous model parameters and a realistic geometry, i.e., the city of Chicago.

The results in [47, 46, 55] assume parameter values uniform in space and time. This oversimplification does not allow to account space heterogeneity in, e.g., attractiveness of housing sites, and behavioral heterogeneity in criminals and policemen. So, we believe that the work in this paper is an important effort towards making the PDE model applicable to realistic scenarios. A way to introduce further heterogeneity is through the above-mentioned activity variable. See Sec. 5 for more details.

The rest of the paper is organized as follows. Sec. 2 discusses the agent-based model and associated PDE model from [47], together with the new boundary conditions. In Sec. 3, we present the time and space discretization of the PDE model and introduce our iterative partitioned algorithm to solve the discretized problem. Sec. 4 reports the numerical results obtained with both models. Future perspectives are discussed in Sec. 5 and conclusions are drawn in Sec. 6.

2 Problem definition

2.1 The agent-based model

This section summarizes the discrete statistical model from [47]. It is relevant to our work for two reasons. First, the PDE model for which we propose a finite element method is derived from this discrete model. Second, Sec. 4.1 compares the results obtained from the PDE model and the agent-based model.

Let us consider a smooth and bounded domain $\Omega \subset \mathbb{R}^2$ where there are two kinds of “agents”: the burglars and the houses where the burglaries can occur. The houses are located on a two-dimensional lattice. For simplicity, we will consider a rectangular domain and consider a Cartesian lattice with uniform spacing $h = \Delta x = \Delta y$. Similarly, time is discretized with a time step Δt . Houses are located at the lattice sites s and are associated with a time-dependent level of attractiveness $A_s(t)$, which measures the

burglars' perception of the attractiveness of the home at site s . In general, we will use subindex s to denote the variables in the discrete model. Variable $A_s(t)$ is written as the sum of a static (possibly spatially varying) component A_s^{st} and a dynamic component $B_s(t)$ associated with repeat and near-repeat victimization:

$$A_s(t) = A_s^{st} + B_s(t). \quad (2.1)$$

Criminals either burglarize a house or move to a neighboring house with a random walk biased toward areas of high attractiveness. The probability of each crime occurring at site s during time interval $[t, t + \Delta t]$ is given by:

$$p_s(t) = 1 - e^{-A_s(t)\Delta t}, \quad (2.2)$$

in accordance with a standard Poisson process. When a site s is burglarized, the corresponding criminal agent is removed from the lattice to represent the tendency of burglars to flee the crime location. Obviously, removed burglars can become active again, which can be seen as a regeneration of criminals. To represent this "burglar generation" in the model, a parameter Γ is introduced at each lattice site s .

The repeat victimization (i.e., a house has an elevated risk of being re-victimised shortly after a first break in), near-repeat victimization (i.e., the houses near a burglarized site are also at increased risk), and the broken-window effect (i.e., the perception of a neighborhood as crime-tolerant increases the risk of the houses in such neighborhood) are modeled through the dynamics behavior of B_s . To capture repeat victimization, the attractiveness $B_s(t)$ at location s is increased by θ for each crime committed at s in the preceding time period [47, 46]. Let parameter ω be a time scale over which repeat victimizations are most likely to occur and $E_s(t)$ the number of burglaries occurred at site s during the time interval $(t - \Delta t, t)$, which we discuss how to compute later. Thus, to account for repeat victimization the equation for $B_s(t + \Delta t)$ needs to include term:

$$B_s(t)(1 - \omega\Delta t) + \theta E_s(t)$$

To model near-repeat victimization and the broken-window effect, we need to allow $B_s(t)$ to spread spatially from a house to its neighbors. So, $B_s(t + \Delta t)$ needs to include term of this kind:

$$\eta \left[\frac{1}{4} \sum_{s' \sim s} B_{s'}(t) - B_s(t) \right] (1 - \omega\Delta t),$$

where parameter $\eta \in [0, 1]$ characterizes the strength of the near-repeat victimization and the broken-window effect and $s' \sim s$ denotes the set of houses neighboring site s . For a Cartesian lattice, the cardinality of this set is four (see Fig. 2), which explains the factor of 4 in the formula above.

Putting the two terms discussed above together, we get

$$B_s(t + \Delta t) = \left[(1 - \eta)B_s(t) + \frac{\eta}{4} \sum_{s' \sim s} B_{s'}(t) \right] (1 - \omega\Delta t) + \theta E_s(t). \quad (2.3)$$

Note that eq. (2.3) can be rewritten as

$$B_s(t + \Delta t) = \left[B_s(t) + \frac{\eta h^2}{4} \Delta_h B_s(t) \right] (1 - \omega\Delta t) + \theta E_s(t), \quad (2.4)$$

where

$$\Delta_h B_s(t) = \frac{1}{h^2} \left(\sum_{s' \sim s} B_{s'}(t) - 4B_s(t) \right) \quad (2.5)$$

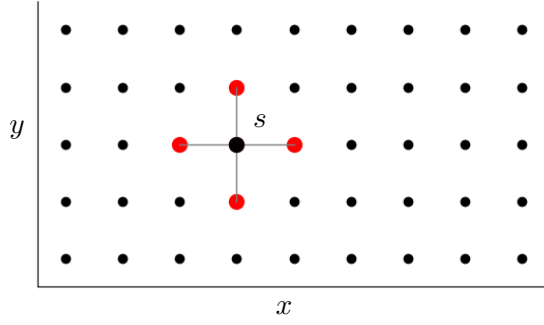


Figure 2: Visualization of a 5-point stencil on a Cartesian lattice with site s at the center to its four neighboring sites (red dots) in the horizontal and vertical directions.

is the discrete spatial Laplacian operator computed using the stencil in Fig. 2. One way to estimate $E_s(t)$ in (2.4) is given by:

$$E_s(t) = n_s(t)p_s(t), \quad (2.6)$$

where $n_s(t)$ is the number of criminals at site s at time t and $p_s(t)$ is defined in (2.2).

Next, we need to describe the evolution of $n_s(t)$. Recall that a criminal at site s at a given time either burglarizes it or moves to a neighboring site following a biased random walk, with probability of moving to site r :

$$q_{s \rightarrow r} = \frac{A_r(t)}{\sum_{s' \sim s} A_{s'}(t)}.$$

Thus, the criminals at site s at the following time step have either arrived after failing to burglarize a neighboring site or have been generated there at rate Γ . This gives us the following:

$$n_s(t + \Delta t) = A_s \sum_{s' \sim s} \frac{n_{s'}(t)[1 - p_{s'}(t)]}{T_{s'}(t)} + \Gamma \Delta t, \quad T_{s'}(t) = \sum_{s'' \sim s'} A_{s''}(t). \quad (2.7)$$

Note that, to prevent the denominator of $q_{s \rightarrow r}$ from becoming zero, we assume $A_s(t) \geq a_0 > 0$ at all times t under consideration and for all sites s in the lattice.

Problem (2.4),(2.7), with $E_s(t)$ as in (2.6) and $p_s(t)$ as in (2.2), represents the probabilistic agent-based model. As mentioned in Sec. 1, we consider a finite domain and natural boundary conditions. In all the numerical results in Sec. 4, we enforce homogeneous natural boundary conditions, meaning that at the boundary the normal derivatives of the system variables are set to zero. In the agent-based model, we enforce such boundary conditions for the level of attractiveness and the number of criminals. Let us consider that simple setting in Fig. 3 to clarify how we handle these boundary conditions: we impose that $A_s(t) = A_r(t)$ and $n_s(t) = n_r(t)$, so that the derivatives of A and n in the normal direction to the boundary $\mathbf{n}$ are zero at the boundary. Note that that boundary condition $A_s(t) = A_r(t)$, which can be rewritten as $B_s(t) + A_s^{st} = B_r(t) + A_r^{st}$, has also the effect of making site s no more attractive than site r and thus criminals would not move from s to r based on the above random walk probability $q_{s \rightarrow r}$.

It is shown in [47] that, depending on the values of the parameters, the system described by (2.4)-(2.7) can exhibit different hotspot patterns. Additionally, when all sites have the

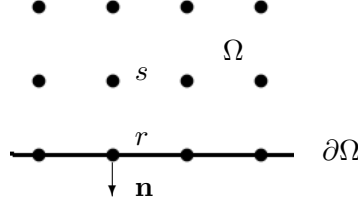


Figure 3: Part of a simplified domain Ω with boundary $\partial\Omega$ and normal to the boundary $\mathbf{n}$. Internal site s lies in the normal direction to the boundary with respect to site r , which lies on the boundary.

same level of attractiveness $\bar{A} = A^{st} + \bar{B}$ and, on average, the same number of criminals $\bar{n}$, the system has a homogeneous equilibrium solution with:

$$\bar{B} = \frac{\theta\Gamma}{\omega}, \quad \bar{n} = \frac{\Gamma\Delta t}{1 - e^{-\bar{A}\Delta t}}. \quad (2.8)$$

While it might be unsettling to see the time step, i.e., a chosen discretization quantity, in the above homogeneous equilibrium solution, we note that at the leading order the denominator of $\bar{n}$ in (2.8) is approximated by $\bar{A}\Delta t$ and so $\bar{n} \approx \Gamma/\bar{A}$.

2.2 The PDE model

Let $\rho = n_s(t)/h^2$ be the density of criminals. By manipulating eq. (2.4), with $E_s(t)$ as in (2.6), and taking its limit for $h, \Delta t \rightarrow 0$, while keeping the ratio $\frac{h^2}{\Delta t}$ fixed to D and quantity $\theta\Delta t$ fixed to ϵ , we obtain:

$$\frac{\partial B}{\partial t} = \frac{\eta D}{4} \Delta B - \omega B + \epsilon D \rho A, \quad A(\mathbf{x}, t) = A^{st}(\mathbf{x}) + B(\mathbf{x}, t), \quad (2.9)$$

where $\mathbf{x}$ gives the spatial coordinates. In (2.9), we have removed the subindex s from the variables since the model is not discrete anymore. Similarly, by manipulating eq. (2.7) and taking its limit for $h, \Delta t \rightarrow 0$, while keeping the ratio $\frac{\Gamma}{h^2}$ fixed to γ , we get:

$$\frac{\partial \rho}{\partial t} = \frac{D}{4} \nabla \cdot \left(\nabla \rho - \frac{2\nabla A}{A} \rho \right) - \rho A + \gamma. \quad (2.10)$$

Similarly to the strict positivity requirement on $A_s(t)$ stated after eq. (2.7), we assume $A(\mathbf{x}, t) \geq a_0 > 0$ for $\mathbf{x} \in \Omega$ and t in the time interval of interest.

To simplify problem (2.9)-(2.10), we introduce the following non-dimensional quantities:

$$\tilde{A} = \frac{A}{\omega} = \frac{A^{st}}{\omega} + \frac{B}{\omega}, \quad \tilde{\rho} = \epsilon L^2 \rho, \quad \tilde{\mathbf{x}} = 2 \frac{\mathbf{x}}{L}, \quad \tilde{t} = \omega t, \quad (2.11)$$

where $L = \sqrt{D/\omega}$ is a characteristic length. We note that the 2 in the third equation in (2.11) comes from $\sqrt{4}$, where 4 is the number of neighbors of a given site (see Fig. 2). For simplicity of notation, we will omit the tilde on the non-dimensional variables and from now on every variable in the continuum (or PDE) model will be non-dimensional.

The non-dimensional continuum model can be written as: find $A(\mathbf{x}, t)$ and $\rho(\mathbf{x}, t)$ such that

$$\frac{\partial A}{\partial t} - \eta \Delta A + A - \rho A = -\eta \Delta A^{st} + A^{st}, \quad (2.12)$$

$$\frac{\partial \rho}{\partial t} - \nabla \cdot \left(\nabla \rho - \frac{2\nabla A}{A} \rho \right) + \rho A = \frac{\Gamma\theta}{\omega^2}, \quad (2.13)$$

in $\Omega \times (0, T]$, where $(0, T]$ is a time interval of interest. Eq. (2.12)-(2.13) are arranged so that known forcing terms are at the right-hand side, while all the terms involving variables are at the left-hand side.

Let us denote a homogeneous equilibrium solution to (2.12)-(2.13) with $\bar{A}$ and $\bar{\rho}$. We note that there is a slight abuse of notation with respect to Sec. 2.1, since here $\bar{A}$ and $\bar{\rho}$ are nondimensionalized according to (2.11). Assuming A^{st} is homogeneous in space, solution $(\bar{A}, \bar{\rho})$ satisfies:

$$\bar{A} - \bar{\rho}\bar{A} = A^{st}, \quad (2.14)$$

$$\bar{\rho}\bar{A} = \frac{\Gamma\theta}{\omega^2}, \quad (2.15)$$

where the right-hand side of (2.15) is $\bar{B}$ is given in (2.8) and nondimensionalized according to (2.11). We continue with the same abuse of notation and denote the right-hand side of (2.15) with $\bar{B}$. Thus, (2.15) can be written as $\bar{\rho} = \bar{B}/\bar{A}$, which can be plugged into (2.14) to get the obvious $\bar{A} = A^{st} + \bar{B}$. Hence, once the value of $\bar{B} = \Gamma\theta/\omega^2$ is selected, we get automatically $\bar{A} = A^{st} + \bar{B}$, since A^{st} is given, and then we can find $\bar{\rho} = \bar{B}/\bar{A}$.

The linear stability analysis of problem (2.12)-(2.13) with periodic boundary conditions determines which combinations of parameters $(\eta, A^{st}, \bar{B})$ lead to the instability around the homogeneous steady-state that generates hotspots [47]. Such instability arises when $\bar{B} > A^{st}/2$ and the following sufficient condition is satisfied:

$$\eta < \frac{3\bar{\rho} + 1 - \sqrt{12\bar{\rho}}}{\bar{A}}. \quad (2.16)$$

This means that, provided that $A^{st} < 2\Gamma\theta/\omega^2$, hotspots form only when the neighborhood effects are weak enough. Although we do not use periodic boundary conditions, we have reported this result because in Sec. 4.1 we will use (2.16) as a heuristic indication of when to expect the above-mentioned instability, i.e., below a given threshold for η . Of course, the value of the threshold would be different in our case since we use Neumann boundary conditions.

Obviously, problem (2.12)-(2.13) needs to be supplemented with appropriate boundary and initial conditions. Given that periodic boundary conditions are not very realistic, we opt for the following natural boundary conditions

$$\nabla A \cdot \mathbf{n} = 0, \quad \text{on } \partial\Omega, \quad t > 0, \quad (2.17)$$

$$\nabla \rho \cdot \mathbf{n} = 0, \quad \text{on } \partial\Omega, \quad t > 0, \quad (2.18)$$

where $\mathbf{n}$ is the outward normal unit vector. Below, we will show how these boundary conditions arise naturally from the variational formulation of problem (2.12)-(2.13). As for the initial conditions, we set:

$$A(\mathbf{x}, 0) = A^{st}(\mathbf{x}) + B_0(\mathbf{x}) + \chi_\delta(\mathbf{x})\xi_B(\mathbf{x}) \in L^2(\Omega), \quad (2.19)$$

$$\rho(\mathbf{x}, 0) = \rho_0(\mathbf{x}) + \chi_\delta(\mathbf{x})\xi_\rho(\mathbf{x}) \in L^2(\Omega), \quad (2.20)$$

where $A^{st}(\mathbf{x}) + B_0(\mathbf{x})$ is the prescribed initial level of attractiveness and $\rho_0(\mathbf{x})$ is the initial density, while $\xi_B(\mathbf{x})$ and $\xi_\rho(\mathbf{x})$ represent noise perturbations, drawn from a Gaussian distribution $\mathcal{N}(0, \sigma^2)$. Function $\chi_\delta(\mathbf{x})$ encodes the sparsity of the noise: $\chi_\delta(\mathbf{x}) = 1$ with probability $\delta \in [0, 1]$, and $\chi_\delta(\mathbf{x}) = 0$ otherwise, independently for each $\mathbf{x}$. Note that $\delta = 1$ corresponds to spatially uniform Gaussian noise, while $\delta = 0$ corresponds to no noise. For $0 < \delta < 1$, the noise is sparse. This formulation of the initial conditions enables

a systematic investigation of the emergence of realistic patterns under both uniform and sparse noise perturbations.

For the numerical method that will be discussed in Sec. 3, we need a variational formulation of problem (2.12)-(2.13). To devise it, one multiplies (2.12) by $u \in H^1(\Omega)$ and (2.13) by $v \in H^1(\Omega)$, integrates over Ω and employs the integration by parts identity. This leads to the variational formulation: Find $A(\mathbf{x}, t) \in H^1(\Omega) \cap L^\infty(\Omega)$ and $\rho(\mathbf{x}, t) \in H^1(\Omega)$ such that

$$\begin{aligned} & \int_{\Omega} \frac{\partial A}{\partial t} u \, d\Omega + \int_{\Omega} \eta \nabla A \cdot \nabla u \, d\Omega - \int_{\partial\Omega} \eta (\nabla A \cdot \mathbf{n}) u \, d\Gamma + \int_{\Omega} A u \, d\Omega \\ & - \int_{\Omega} \rho A u \, d\Omega = \int_{\Omega} (A^{st} - \eta \Delta A^{st}) u \, d\Omega, \\ & \int_{\Omega} \frac{\partial \rho}{\partial t} v \, d\Omega + \int_{\Omega} \left(\nabla \rho - \frac{2 \nabla A}{A} \rho \right) \nabla v \, d\Omega - \int_{\partial\Omega} \left(\nabla \rho - \frac{2 \nabla A}{A} \rho \right) \cdot \mathbf{n} v \, d\Gamma \\ & + \int_{\Omega} \rho A v \, d\Omega = \int_{\Omega} \frac{\Gamma \theta}{\omega^2} v \, d\Omega, \end{aligned}$$

for all $(u, v) \in H^1(\Omega) \times H^1(\Omega)$. Boundary integrals involving $\nabla A \cdot \mathbf{n}$ and $\nabla \rho \cdot \mathbf{n}$ have naturally appeared in the above weak problem. Thanks to boundary conditions (2.17)-(2.18), such integrals vanish. Thus, we obtain the following simpler weak problem: Find $A(\mathbf{x}, t) \in H^1(\Omega) \cap L^\infty(\Omega)$ and $\rho(\mathbf{x}, t) \in H^1(\Omega)$ such that

$$\int_{\Omega} \frac{\partial A}{\partial t} u \, d\Omega + \int_{\Omega} \eta \nabla A \cdot \nabla u \, d\Omega + \int_{\Omega} A u \, d\Omega - \int_{\Omega} \rho A u \, d\Omega = \int_{\Omega} (A^{st} - \eta \Delta A^{st}) u \, d\Omega, \quad (2.21)$$

$$\int_{\Omega} \frac{\partial \rho}{\partial t} v \, d\Omega + \int_{\Omega} \left(\nabla \rho - \frac{2 \nabla A}{A} \rho \right) \nabla v \, d\Omega + \int_{\Omega} \rho A v \, d\Omega = \int_{\Omega} \frac{\Gamma \theta}{\omega^2} v \, d\Omega, \quad (2.22)$$

for all $(u, v) \in H^1(\Omega) \times H^1(\Omega)$.

In (2.21), we have not integrated by parts the Laplacian of A^{st} , which means that we require $A^{st} \in H^2(\Omega)$. That is the case for all the numerical tests discussed in Sec. 4. If one wants to require less regularity on A^{st} , i.e., $A^{st} \in H^1(\Omega)$, then eq. (2.21) becomes

$$\begin{aligned} & \int_{\Omega} \frac{\partial A}{\partial t} u \, d\Omega + \int_{\Omega} \eta \nabla A \cdot \nabla u \, d\Omega + \int_{\Omega} A u \, d\Omega - \int_{\Omega} \rho A u \, d\Omega = \int_{\Omega} A^{st} u \, d\Omega \\ & + \int_{\Omega} \eta \nabla A^{st} \cdot \nabla u \, d\Omega - \int_{\partial\Omega} \eta (\nabla A^{st} \cdot \mathbf{n}) u \, d\Gamma. \end{aligned}$$

3 Numerical method for the PDE model

PDE problem (2.21)-(2.22) poses some interesting challenges. The first main challenge is related to its two-way coupled nature: both variables ρ and A appear in both eq. (2.21) and eq. (2.22). The coupling terms create an intricate feedback mechanisms between the criminal density ρ and the attractiveness field A . The second main challenge is the strong nonlinear nature of the problem. In particular, the second term in eq. (2.22) is degenerate and nonlinear and, as such, it could complicate stability and convergence in a numerical scheme. Thus, the design of an accurate and efficient numerical method for problem (2.21)-(2.22) is non-trivial.

Let us start from the time discretization. For this, we introduce a time step $\Delta t = T/N_t$ for given N_t , set $t^n = n\Delta t$ for $n = 1, \dots, N_t$ and denote with y^n the approximation of quantity y at time t^n . For simplicity, we apply Backward Differentiation Formula of order

1 to (2.21)-(2.22) and get the following problem at the generic time t^{n+1} : for $n \geq 0$, given A^n and ρ^n , find $A^{n+1} \in H^1(\Omega) \cap L^\infty(\Omega)$ and $\rho^{n+1} \in H^1(\Omega)$ such that

$$\begin{aligned} & \frac{1}{\Delta t} \int_{\Omega} A^{n+1} u \, d\Omega + \int_{\Omega} \eta \nabla A^{n+1} \cdot \nabla u \, d\Omega + \int_{\Omega} A^{n+1} u \, d\Omega \\ & - \int_{\Omega} \rho^{n+1} A^{n+1} u \, d\Omega = \int_{\Omega} (A^{st} - \eta \Delta A^{st}) u \, d\Omega + \frac{1}{\Delta t} \int_{\Omega} A^n u \, d\Omega, \end{aligned} \quad (3.1)$$

$$\begin{aligned} & \frac{1}{\Delta t} \int_{\Omega} \rho^{n+1} v \, d\Omega + \int_{\Omega} \left(\nabla \rho^{n+1} - \frac{2 \nabla A^{n+1}}{A^{n+1}} \rho^{n+1} \right) \nabla v \, d\Omega \\ & + \int_{\Omega} \rho^{n+1} A^{n+1} v \, d\Omega = \int_{\Omega} \frac{\Gamma \theta}{\omega^2} v \, d\Omega + \frac{1}{\Delta t} \int_{\Omega} \rho^n v \, d\Omega, \end{aligned} \quad (3.2)$$

for all $(u, v) \in H^1(\Omega) \times H^1(\Omega)$. System (3.1)-(3.2) is implicit and fully coupled. Hence, if one adopts a fine mesh for space discretization, it could lead to a large nonlinear system that needs to be solved at each time step through an iterative solvers like, e.g., Newton's method. The strongly nonlinear term $\frac{2 \nabla A^{n+1}}{A^{n+1}} \rho^{n+1}$ could make the convergence of such iterative method slow, especially for large-scale simulations [40, 43].

One way to contain the computational cost is to decouple system (3.1)-(3.2), while linearizing it, so that one solves two smaller linear systems, instead of one large nonlinear system, per time step. At time t^{n+1} , given A^n and ρ^n , the decoupled, linear algorithm performs two sequential steps:

- *Step 1*: find $A^{n+1} \in H^1(\Omega)$ such that

$$\begin{aligned} & \frac{1}{\Delta t} \int_{\Omega} A^{n+1} u \, d\Omega + \int_{\Omega} \eta \nabla A^{n+1} \cdot \nabla u \, d\Omega + \int_{\Omega} A^{n+1} u \, d\Omega \\ & - \int_{\Omega} \rho^n A^{n+1} u \, d\Omega = \int_{\Omega} (A^{st} - \eta \Delta A^{st}) u \, d\Omega + \frac{1}{\Delta t} \int_{\Omega} A^n u \, d\Omega, \end{aligned} \quad (3.3)$$

for all $u \in H^1(\Omega)$.

- *Step 2*: find $\rho^{n+1} \in H^1(\Omega)$ such that

$$\begin{aligned} & \frac{1}{\Delta t} \int_{\Omega} \rho^{n+1} v \, d\Omega + \int_{\Omega} \left(\nabla \rho^{n+1} - \frac{2 \nabla A^{n+1}}{A^{n+1}} \rho^{n+1} \right) \nabla v \, d\Omega \\ & + \int_{\Omega} \rho^{n+1} A^{n+1} v \, d\Omega = \int_{\Omega} \frac{\Gamma \theta}{\omega^2} v \, d\Omega + \frac{1}{\Delta t} \int_{\Omega} \rho^n v \, d\Omega, \end{aligned} \quad (3.4)$$

for all $v \in H^1(\Omega)$.

While successfully containing the computational time, the above scheme may suffer from instability for certain combination of parameter values. Algorithms like (3.3)-(3.4) are called loosely-coupled because they do not enforce the strong coupling between the two subproblems at each time step.

One way to cure possible instability in (3.3)-(3.4) is to iterate between Step 1 and Step 2 at a given time step until a suitable convergence criterion is satisfied. This means that at every time step t^{n+1} , given A^n and ρ^n , the following steps are performed at iteration $k+1$, $k \geq 0$:

- *Step 1*: find $A^{k+1} \in H^1(\Omega)$ such that

$$\begin{aligned} & \frac{1}{\Delta t} \int_{\Omega} A^{k+1} u \, d\Omega + \int_{\Omega} \eta \nabla A^{k+1} \cdot \nabla u \, d\Omega + \int_{\Omega} A^{k+1} u \, d\Omega \\ & - \int_{\Omega} \rho^k A^{k+1} u \, d\Omega = \int_{\Omega} (A^{st} - \eta \Delta A^{st}) u \, d\Omega + \frac{1}{\Delta t} \int_{\Omega} A^n u \, d\Omega, \end{aligned} \quad (3.5)$$

for all $u \in H^1(\Omega)$. For $k = 0$, $\rho^k = \rho^n$.

- *Step 2*: find $\rho^{k+1} \in H^1(\Omega)$ such that

$$\begin{aligned} \frac{1}{\Delta t} \int_{\Omega} \rho^{k+1} v \, d\Omega + \int_{\Omega} \left(\nabla \rho^{k+1} - \frac{2 \nabla A^{k+1}}{A^{k+1}} \rho^{k+1} \right) \nabla v \, d\Omega \\ + \int_{\Omega} \rho^{k+1} A^{k+1} v \, d\Omega = \int_{\Omega} \frac{\Gamma \theta}{\omega^2} v \, d\Omega + \frac{1}{\Delta t} \int_{\Omega} \rho^n v \, d\Omega, \end{aligned} \quad (3.6)$$

for all $v \in H^1(\Omega)$.

- *Step 3*: Check the stopping criterion, e.g.

$$\frac{\|A^{k+1} - A^k\|_{L^2(\Omega)}}{\|A^k\|_{L^2(\Omega)}} < \text{tol}_1 \quad \text{and} \quad \frac{\|\rho^{k+1} - \rho^k\|_{L^2(\Omega)}}{\|\rho^k\|_{L^2(\Omega)}} < \text{tol}_2, \quad (3.7)$$

where tol_1 and tol_2 are given stopping tolerances. If not satisfied, repeat steps 1-3. If satisfied, set $A^{n+1} = A^{k+1}$ and $\rho^{n+1} = \rho^{k+1}$.

At a given time step, algorithm (3.5)-(3.7) solves two smaller linear problems, i.e., (3.5)-(3.6), as many times as needed to satisfy criterion (3.7). It can be viewed as a fixed-point scheme for nonlinear system (3.1)-(3.2) that allows to decouple the problem while linearizing it. So, algorithms like (3.5)-(3.7) are called strongly-coupled, as they enforce the coupling between two subproblems up to user-prescribed tolerances.

Next, we present the space discretization of (3.5)-(3.6). We introduce a partition $\mathcal{T}_h$ of Ω into N_e quadrilaterals Q_k such that

$$\bar{\Omega} = \bigcup_{k=1}^{N_e} \bar{Q}_k,$$

with

$$h = \max_{Q_k \in \mathcal{T}_h} h_k, \quad h_k = \text{diam}(Q_k), \quad k = 1, \dots, N_e.$$

Let $\mathbb{Q}_N(Q_k)$ be the set of algebraic polynomials, defined on Q_k , of degree less than or equal to N in each space variable, and let $V_h(\Omega)$ be the space of global continuous functions on $\bar{\Omega}$ that are polynomials of degree N on each $Q_k \in \mathcal{T}_h$:

$$V_h(\Omega) = \{v_h \in C^0(\bar{\Omega}) : v_h|_{Q_k} \in \mathbb{Q}_N, \forall Q_k \in \mathcal{T}_h\}.$$

We denote with $\{\phi_i\}_{i=1}^{N_Q}$ a basis for V_h , where N_Q the total number of degrees of freedom in Ω .

The space-discrete counterpart of (3.5) reads: find $A_h^{k+1} \in V_h(\Omega)$ such that

$$\begin{aligned} \frac{1}{\Delta t} \int_{\Omega} A_h^{k+1} u \, d\Omega + \int_{\Omega} \eta \nabla A_h^{k+1} \cdot \nabla u \, d\Omega + \int_{\Omega} A_h^{k+1} u \, d\Omega \\ - \int_{\Omega} \rho_h^k A_h^{k+1} u \, d\Omega = \int_{\Omega} (A_{0,h} - \eta \Delta A_{0,h}) u \, d\Omega + \frac{1}{\Delta t} \int_{\Omega} A_h^n u \, d\Omega, \end{aligned} \quad (3.8)$$

for all $u \in V_h(\Omega)$. Similarly, the space-discrete counterpart of (3.6) reads: find $\rho_h^{k+1} \in V_h(\Omega)$ such that

$$\begin{aligned} \frac{1}{\Delta t} \int_{\Omega} \rho_h^{k+1} v \, d\Omega + \int_{\Omega} \left(\nabla \rho_h^{k+1} - \frac{2 \nabla A_h^{k+1}}{A_h^{k+1}} \rho_h^{k+1} \right) \nabla v \, d\Omega \\ + \int_{\Omega} \rho_h^{k+1} A_h^{k+1} v \, d\Omega = \int_{\Omega} \frac{\Gamma \theta}{\omega^2} v \, d\Omega + \frac{1}{\Delta t} \int_{\Omega} \rho_h^n v \, d\Omega, \end{aligned} \quad (3.9)$$

for all $v \in V_h(\Omega)$. Obviously, stopping criterion (3.7) becomes

$$\frac{\|A_h^{k+1} - A_h^k\|_{L^2(\Omega)}}{\|A_h^k\|_{L^2(\Omega)}} < \text{tol}_1 \quad \text{and} \quad \frac{\|\rho_h^{k+1} - \rho_h^k\|_{L^2(\Omega)}}{\|\rho_h^k\|_{L^2(\Omega)}} < \text{tol}_2. \quad (3.10)$$

In order to write the matrix formulation of (3.8)-(3.9), let M be the mass matrix, K the stiffness matrix, $N(w)$ a weighted mass matrix and $D(w)$ and “weighted” divergence matrix, for suitable weight $w \in V_h(\Omega)$:

$$\begin{aligned} M_{i,j} &= \int_{\Omega} \phi_j \phi_i \, d\Omega, \quad K_{i,j} = \int_{\Omega} \nabla \phi_j \cdot \nabla \phi_i \, d\Omega, \\ N(w)_{i,j} &= \int_{\Omega} w \phi_j \phi_i \, d\Omega, \quad D(w)_{i,j} = \int_{\Omega} \frac{2\nabla w}{w} \phi_j \nabla \phi_i \, d\Omega, \end{aligned}$$

for $i, j = 1, \dots, N_Q$. Moreover, let $\mathbf{A}^{k+1}, \mathbf{r}^{k+1} \in \mathbb{R}^{N_Q}$ be the vectors of the nodal values of A_h^{k+1} and ρ_h^{k+1} .

The matrix form of (3.8) is given by: find vector $\mathbf{A}^{k+1} \in \mathbb{R}^{N_Q}$ such that

$$\frac{1 + \Delta t}{\Delta t} M \mathbf{A}^{k+1} + \eta K \mathbf{A}^{k+1} - N(\mathbf{r}^k) \mathbf{A}^{k+1} = \mathbf{b}_A^{k+1}, \quad (3.11)$$

where the vector at the right-hand side $\mathbf{b}_A^{k+1}$ has i -th component

$$\mathbf{b}_{A,i}^{k+1} = \int_{\Omega} (-\eta \Delta A_{0,h} - A_{0,h}) \phi_i \, d\Omega + \frac{1}{\Delta t} (M \mathbf{A}^k)_i,$$

for $i = 1, \dots, N_Q$. Note that the above matrix form assumes that η is constant in space. This means that step 1 can be written as: solve linear system:

$$\left[\frac{1 + \Delta t}{\Delta t} M + \eta K - N(\mathbf{r}^k) \right] \mathbf{A}^{k+1} = \mathbf{b}_A^{k+1}. \quad (3.12)$$

The matrix form of (3.9) is given by: find vector $\mathbf{r}^{k+1} \in \mathbb{R}^{N_Q}$ such that

$$\frac{1}{\Delta t} M \mathbf{r}^{k+1} + K \mathbf{r}^{k+1} - D(\mathbf{A}^{k+1}) \mathbf{r}^{k+1} + N(\mathbf{A}^{k+1}) \mathbf{r}^{k+1} = \mathbf{b}_{\rho}^{k+1}, \quad (3.13)$$

where the vector at the right-hand side $\mathbf{b}_{\rho}^{k+1}$ has i -th component

$$\mathbf{b}_{\rho,i}^{k+1} = \int_{\Omega} \frac{\Gamma \theta}{\omega^2} \phi_i \, d\Omega + \frac{1}{\Delta t} (M \mathbf{r}^k)_i,$$

for $i = 1, \dots, N_Q$. Then, step 2 can be written in the equivalent form:

$$\left[\frac{1}{\Delta t} M + K - D(\mathbf{A}^{k+1}) + N(\mathbf{A}^{k+1}) \right] \mathbf{r}^{k+1} = \mathbf{b}_{\rho}^{k+1}, \quad (3.14)$$

where $\mathbf{A}^{k+1}$ is known from solving (3.12). Systems (3.12), (3.14) clearly show that algorithm (3.5)-(3.7) is a fixed-point scheme for problem (3.1)-(3.2).

4 Numerical experiments

We assess the proposed finite element-based solver for the PDE model through several numerical experiments. We start with the validation in Sec. 4.1 and then study the robustness and computational efficiency of the solver in Sec. 4.2. For the results in Sec. 4.1 and 4.2, we use a simple square geometry and constant (in time and space) model parameters. Results for spatially heterogeneous parameters and a nontrivial geometry are presented in Sec. 4.3.

We implemented the PDE solver using an iterative partitioned scheme in FEniCSx [7, 44, 6], a popular open-source computing platform for solving PDE problems with the finite element method. For the agent-based model, we developed an in-house solver in MATLAB. Both the agent-based and the PDE solvers are available on GitHub [1]. For all the simulations based on the PDE model, we used first-order Lagrange finite elements on quadrilateral cells $\mathbb{Q}_1$.

4.1 Validation of the solver for the PDE model

As mentioned in Sec. 1, we validate the results obtained with the PDE model against the results provided by the agent-based model. For this purpose, we choose 3 representative cases. Since the PDE model is appropriate when the number of criminals is large (i.e., the criminal population can be looked at as a continuum versus separate individuals), all three cases have a large initial density. For each case, we first present the solution given by the PDE model and motivate the choice of model parameters. Then, we compare with the solution obtained with the agent-based model.

All the quantities mentioned in the following are non-dimensional. The computational domain is $\Omega = [0, 16] \times [0, 16]$ and time interval of interest is $[0, 200]$. We consider three sets of parameters and initial conditions that yield representative solutions:

- Case 1: $\eta = 0.9$, $A^{st} = 1/30$, $\Gamma\theta/\omega^2 = B_0 = 1$, and $\rho_0 = 0.8$. No noise is added to the initial data, i.e., $\chi_\delta(\mathbf{x}) = 0$ in (2.19)–(2.20).
- Case 2: $\eta = 0.3$, $A^{st} = 1/30$, $\Gamma\theta/\omega^2 = B_0 = 1$, and $\rho_0 = 0.8$. Uniform Gaussian noise is added to the initial data, with $\chi_\delta(\mathbf{x}) = 1$, $\xi_B(\mathbf{x}) \sim \mathcal{N}(0, 0.05^2)$ and $\xi_\rho(\mathbf{x}) \sim \mathcal{N}(0, 0.01^2)$ in (2.19)–(2.20), sampled independently at each spatial point $\mathbf{x}$.
- Case 3: $\eta = 0.03$, $A^{st} = 1/30$, $\Gamma\theta/\omega^2 = B_0 = 1$, and $\rho_0 = 0.8$. The same noise as in case 2 is added to the initial data.

In all three case, we have set $B_0 = \Gamma\theta/\omega^2$, which corresponds to $\bar{B}$ in (2.8) after non-dimensionalization as in (2.11). Additionally, we note that all the cases have non-dimensional $\bar{B} > A^{st}/2$. In this section, we set A^{st} constant because that is typically the case for the numerical results given by the PDE model from [47] presented in the literature [47, 46, 55]. We will consider a space-dependent $A^{st} = A^{st}(\mathbf{x})$ in Sec. 4.3. For realistic simulations, the setting of A^{st} is an open question that needs to be object of future work. See Sec. 5.

To simulate the crime dynamics in the above three cases, we consider time step to $\Delta t = 1/50$ and a mesh with size $h = 16/200$ for both the PDE solver and the agent-based solver. The PDE solver uses the iterative partitioned algorithm (3.5)–(3.7), with stopping tolerances $\text{tol}_1 = \text{tol}_2 = 10^{-9}$.

Fig. 4 shows the evolution of the level of attractiveness and density for case 1 given by the PDE solver. We see that the system quickly reaches a steady-state solution that is uniform in space, i.e., no hotspots form. The computed value of A shown in Fig. 4 is roughly 1.0333, which approximates $\bar{A} = A^{st} + \Gamma\theta/\omega^2 = 1/30 + 1$, while the computed

value of ρ is roughly 0.9677, which approximates $\bar{\rho} = \Gamma\theta/(\omega^2\bar{A}) = 1/(1/30 + 1)$. Recall that these $\bar{A}$ and $\bar{\rho}$ satisfy (2.14)-(2.15). Since we are not adding noise, initial conditions (2.19)-(2.20) become $A = \bar{A}$ and $\rho = \rho_0 \neq \bar{\rho}$. The difference between ρ_0 and $\bar{\rho}$ explains the transition to the steady state observed in Fig. 4.

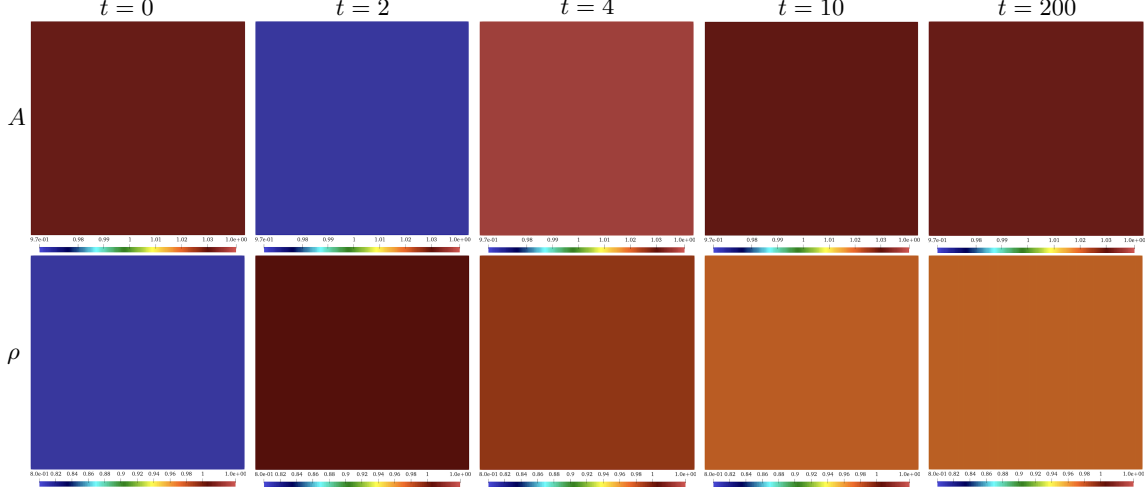


Figure 4: Case 1 - PDE model: Evolution of the level of attractiveness (top) and density (bottom).

If one adds noise to the initial data of case 1, the computed solutions look similar to the ones reported in Fig. 4, with the addition of a small noisy perturbation that we suspect would be dampened if the simulation time is long enough.

The initial conditions and the parameters used for the results in Fig. 4 are suitably dimensionalized and passed to the agent-based solver, which uses all dimensional quantities. For the agent-based solver, we set $\Gamma = 0.0077$, $\theta = 0.58$, and $\omega = 1/15$ to obtain $\Gamma\theta/\omega^2 \approx 1$, as required by case 1. The evolution of A and ρ for case 1 given by the agent-based solver after non-dimensionalization are shown in Fig. 5. We see that, while A is not homogeneous in space like in Fig. 4, the attractiveness field has at most places a value close to $\bar{A}$. Any local variation in values disappears rather quickly. See Fig. 6 for A past $t = 200$. The same is true for ρ . We note that also in [47] a spatially homogeneous PDE solution in Fig. 4(a) corresponds to parameter values that give agent-based solutions not exactly spatially homogeneous, shown in Fig. 3(a) and 3(b).

The stability analysis from [47] suggests that to see the formation of hotspots, one should lower the value of η to satisfy condition (2.16). While that analysis assumes periodic boundary conditions, which we do not impose, we used condition (2.16) as an indication to design case 2. Note that case 2 has the same parameters as case 1, with the exception of η , which is lower than in case 1. This means that case 2 is characterized by weaker neighborhood effects.

Fig. 7 displays the evolution of the level of attractiveness and density for case 2 given by the PDE solver. Indeed, we observe that the system tends towards a steady state with large hotspots. In the top row of Fig. 7, we see large spots of high attractiveness surrounded by areas of low attractiveness. Variable ρ mirrors the pattern in A : in the bottom row of Fig. 7 we see the emergence of large spots with high concentrations of criminals surrounded by areas with much lower concentrations.

The noise in the initial data of case 2 affects the position of the hotspots, not on their existence or rough size, and the time needed for the emergence of the pattern. See Fig. 8, which shows A and ρ for case 2 with no noise computed by the PDE solver. The hotspots

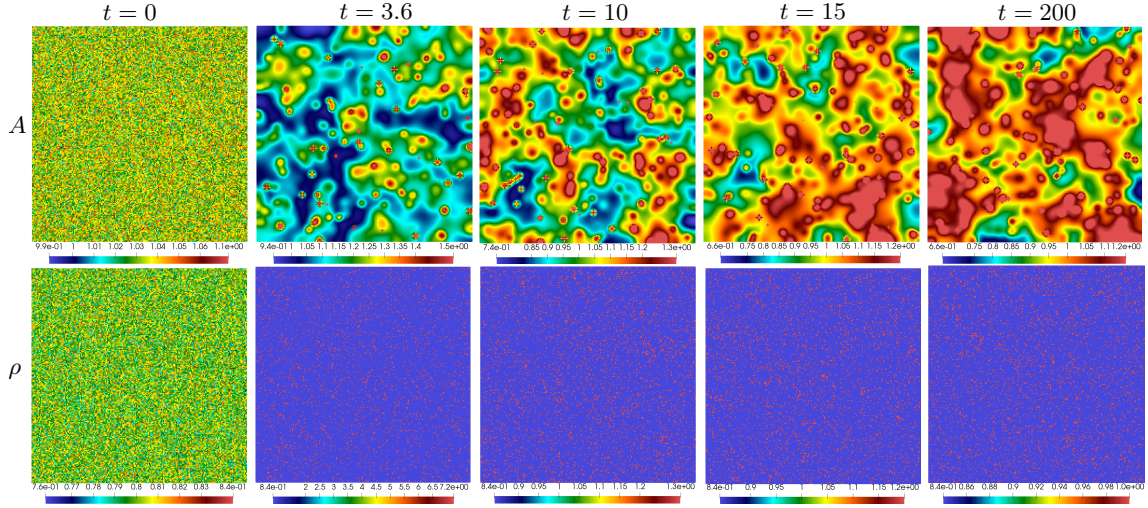


Figure 5: Case 1 - agent-based model: Evolution of the level of attractiveness (top) and criminals (bottom).

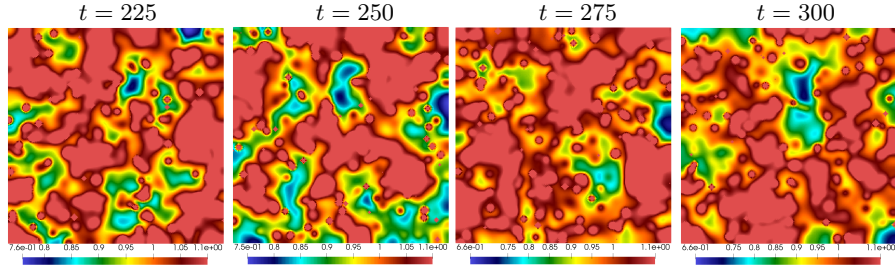


Figure 6: Case 1 - agent-based model: Evolution of A past $t = 200$.

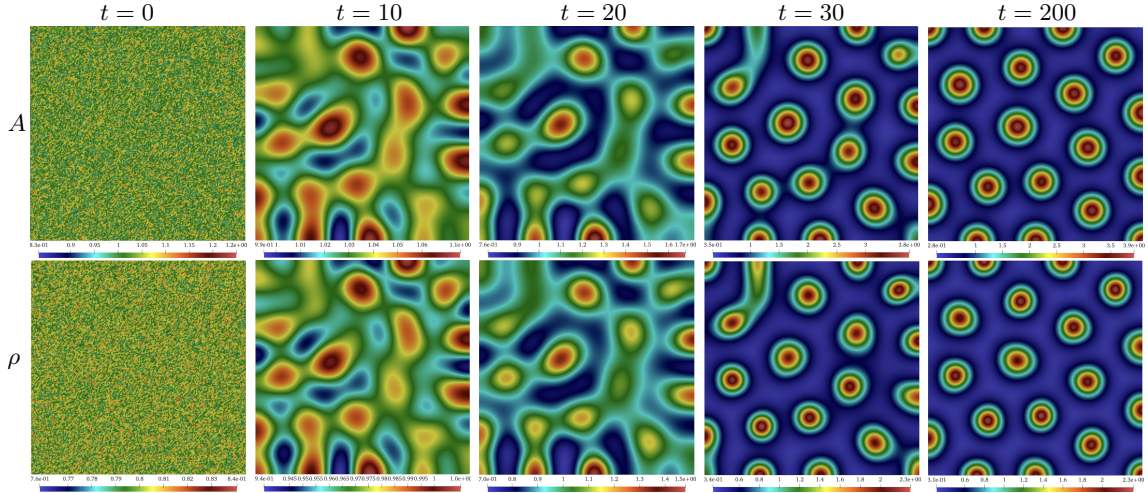


Figure 7: Case 2 - PDE model: Evolution of the level of attractiveness (top) and density (bottom).

have roughly the same size as those in Fig. 7, but their shape is more regular and the pattern is more structured. Additionally, we see that while at $t = 30$ a hotspot pattern has almost emerged in case 2 with noise, it takes around 130 time units for that to be the case when the noise is absent.

For comparison with Fig. 7, Fig. 9 shows the evolution of A and ρ for case 2 given by the agent-based solver after non-dimensionalization. For this case, we set $\Gamma = 0.019$,

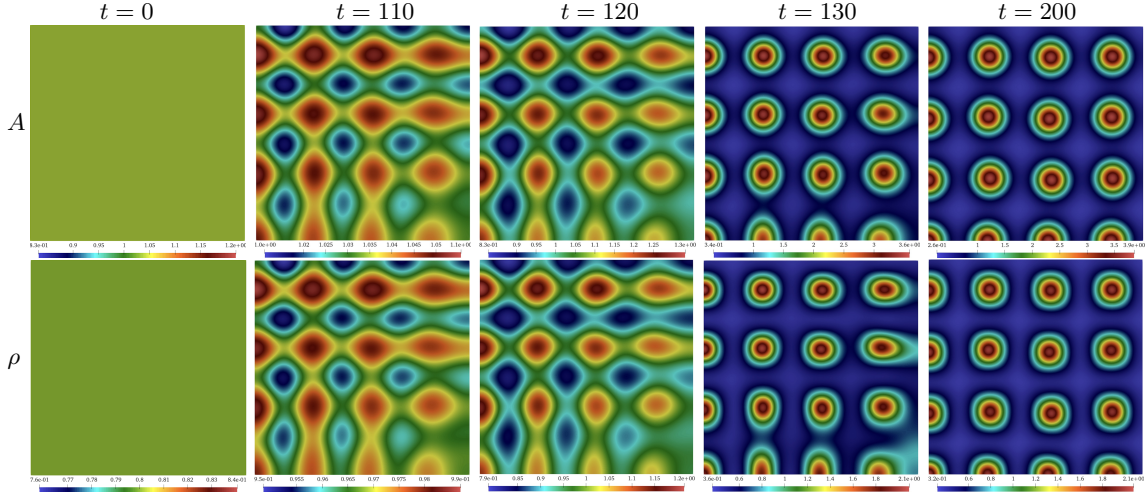


Figure 8: Case 2 without noise - PDE model: Evolution of the level of attractiveness (top) and density (bottom).

$\theta = 0.2339$, and $\omega = 1/15$ to obtain $\Gamma\theta/\omega^2 \approx 1$. We see that the magnitudes of A and ρ in Fig. 7 and 9 match and the sizes for the hotspots are comparable. The reason why we changed the values of Γ and θ with respect to case 1 is to make the hotspots in ρ more visible. For $\Gamma = 0.0077$ and $\theta = 0.58$, the hotspots in ρ are still there but look fainter in the plots.

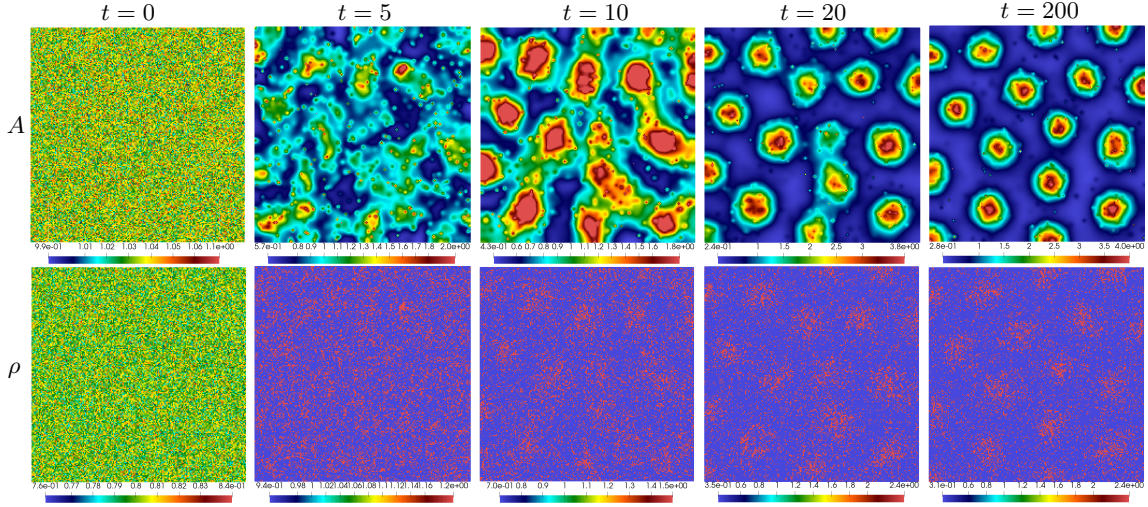


Figure 9: Case 2 - agent-based model: Evolution of the level of attractiveness (top) and criminals (bottom).

Further analysis from [47] indicates that hotspot separation increases with η . The hotspot separation is affected by their size, which means that the hotspot size decreases with η . For this reason, η is further decreased in case 3, while all the other parameters remain the same as in case 1 and 2.

Fig. 10 illustrates the evolution of attractiveness and density for case 3 given by the PDE solver. We observe the early emergence of small hotspots. Indeed, by $t = 10$, distinct small hotspots are already visible in Fig. 10, whereas in Fig. 7, the large-scale hotspot structures have not yet emerged. This comparison highlights that a smaller value of η leads to a faster onset of the hotspot pattern. For completeness, we report in Fig. 11 the evolution of A and ρ for case 3 with no noise. Again, we see that, in the absence of noise,

the hotspot pattern is structured and the pattern emergence is delayed.

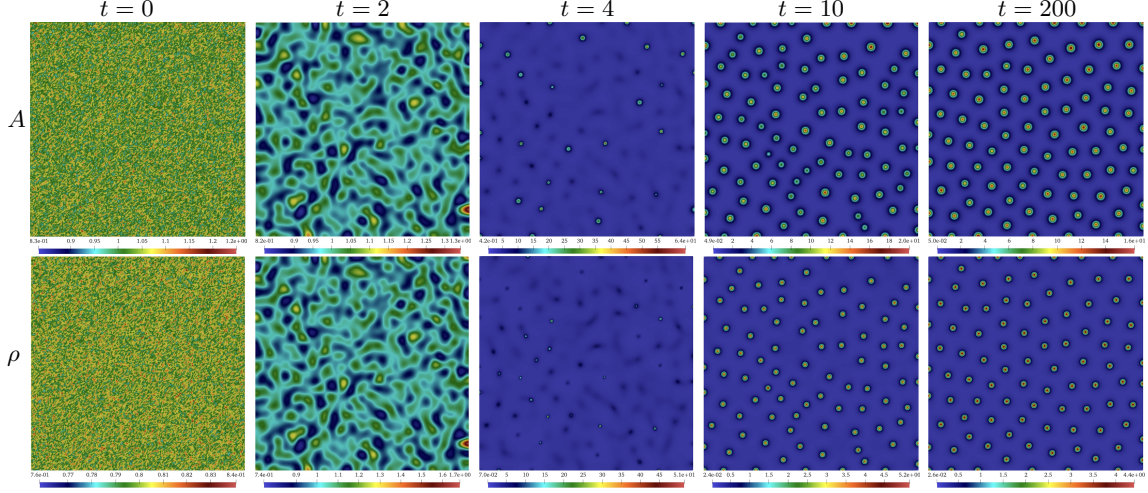


Figure 10: Case 3 - PDE model: Evolution of the level of attractiveness (top) and density (bottom).

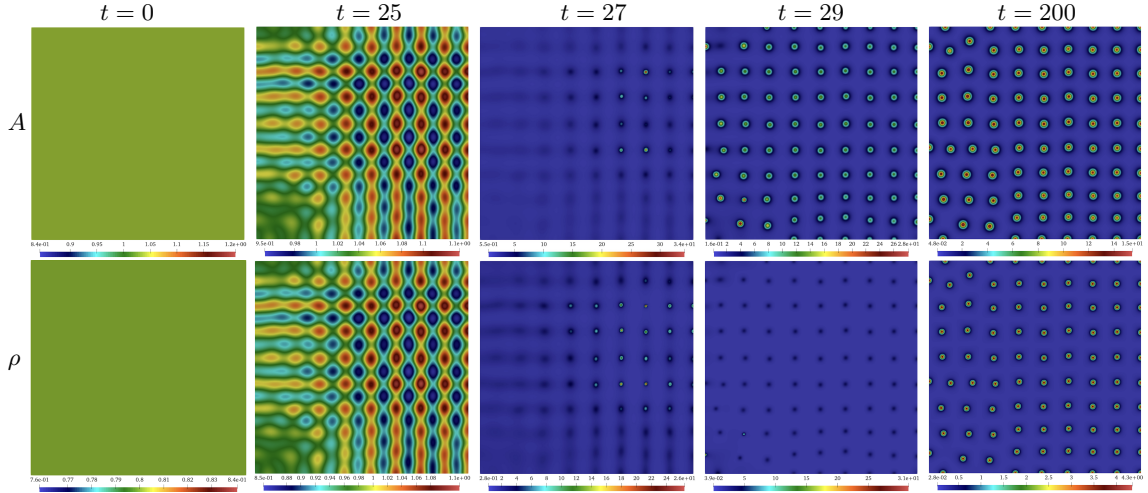


Figure 11: Case 3 without noise - PDE model: Evolution of the level of attractiveness (top) and density (bottom).

Finally, Fig. 12 shows the evolution of A and ρ for case 3 given by the agent-based solver after non-dimensionalization. For these results, we set again $\Gamma = 0.019$, $\theta = 0.2339$, and $\omega = 1/15$, which give $\Gamma\theta/\omega^2 \approx 1$. We observe that the hotspots in Fig. 10 and 12 are comparable in size and number. Additionally, the magnitudes of A and ρ match.

From the results presented in this section, we can conclude that the PDE solver output matches well the output from the agent-based simulations when the density of criminals is large.

4.2 Robustness and computational efficiency of the iterative partitioned solver

Now that the PDE solver has been validated, we use it to further investigate the relationship between η and the size of the hotspots. For this purpose, we vary the parameter η over the interval $[0.025, 0.45]$. Next, we examine the influence of discretization parameters and

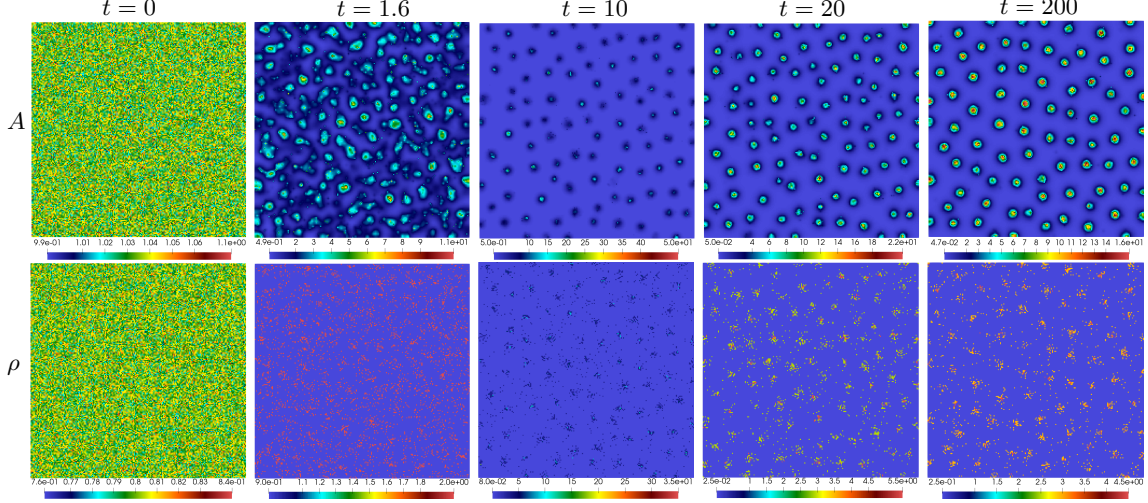


Figure 12: Case 3 - agent-based model: Evolution of the level of attractiveness (top) and criminals (bottom).

stopping tolerances for the three cases studied in Sec. 4.1, in order to better understand the robustness and computational efficiency of the iterative partitioned solver.

Fig. 13 reports the hotspot diameter and number of hotspots in domain Ω for $\eta \in [0.025, 0.45]$. The two quantities are measured from the computed results once the hotspot pattern has fully emerged and shows no visual changes over several time units. To reduce the computational time need for this study, we considered a coarser mesh ($h = 16/100$) and a larger time step ($\Delta t = 1/25$) than the ones used for the figures in the previous section and we reduced the stopping tolerances ($\text{tol}_1 = \text{tol}_2 = 10^{-6}$). Fig. 13 indicates a nontrivial dependence of both diameter and number of hotspots on η . A quantitative characterization of the hotspot number and diameter is achieved via empirical curve fitting. Specifically, the number of hotspots is well approximated by the exponential decay $N(\eta) \approx 96.82 * \exp(-16.14\eta) + 12.33$, while the hotspot diameter follows a quadratic law $D(\eta) \approx -5.24\eta^2 + 8.70\eta + 0.73$. We emphasize that these functional forms are obtained via numerical regression and currently lack a rigorous analytical derivation. Developing such theoretical justification would involve intricate analytical work, which lies beyond the scope of the present work. We refer the reader to related analytical studies in similar contexts, such as [5, 18] for the Cahn–Hilliard equation.

Let us now define η piecewise as follows:

$$\eta = \begin{cases} 0.03 \text{ (as in case 3)}, & \text{if } 12 \leq x_1 \leq 16, \\ 0.12, & \text{if } 8 \leq x_1 < 12, \\ 0.3 \text{ (as in case 2)}, & \text{if } 4 \leq x_1 < 8, \\ 0.9 \text{ (as in case 1)}, & \text{if } 0 \leq x_1 < 4, \end{cases} \quad (4.1)$$

i.e., η decreases with the vertical coordinate. We set all the other parameters and noise as in case 2 and 3. Fig. 14 displays the evolution of A and ρ computed with mesh $h = 16/200$ and time step $\Delta t = 1/50$. For a better visualization of the hotspots, we have manually set the maximum and minimum of A and ρ in the color bars. From Fig. 14, we see that, by the end of the simulation, the solution transitions from uniform in space (with $A = \bar{A}$ and $\rho = \bar{\rho}$, roughly in the bottom quarter) to hotspots with decreasing size as x_1 increases. This is expected from (4.1). Notice also that the smaller the hotspots the faster they emerge.

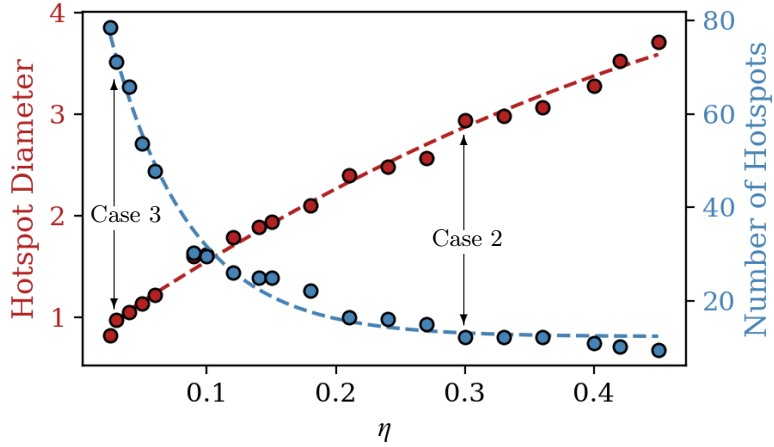


Figure 13: Hotspot diameter (left axis, red markers) and number of hotspots (right axis, blue markers) as the strength of the neighborhood effects η is varied in interval $[0.025, 0.45]$. The red and blue dashed lines fit the data with the corresponding color.

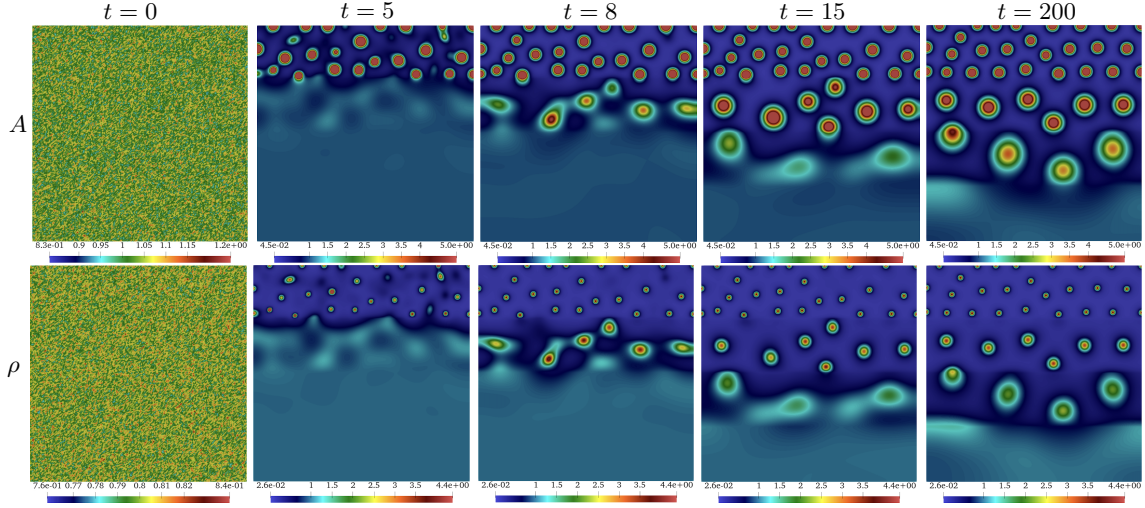


Figure 14: Case with η defined piecewise (4.1): Evolution of the level of attractiveness (top) and density (bottom).

Next, we vary the discretization parameters and stopping tolerances. Tab. 1 reports the average number of iterations over time required by the iterative partitioned algorithm to meet the stopping criterion (3.7) with $\text{tol}_1 = \text{tol}_2 = 10^{-3}, 10^{-6}, 10^{-9}$ for the three cases from Sec. 4.1 with different mesh sizes and time steps. We consider the combinations of $(h, \Delta t)$ used for the results presented thus far, i.e., $(16/200, 1/50)$ and $(16/100, 1/25)$, and added one finer level, i.e., $(16/400, 1/100)$. Additionally, we considered combination $(16/100, 2/25)$, which keeps the same mesh size as in one of the other combinations but doubles the time step. For case 1, the average number of iterations is around 1 for all $(h, \Delta t)$ combinations and tolerance values. This reflects the simplicity of the solution, which is spatially homogeneous and approaches a steady state rather quickly. Additionally, it indicates that even the algorithm in its non-iterative form, i.e., (3.3)-(3.4), could be stable and accurate for case 1. For case 2 and 3, the average number of iterations is around 1 only for $\text{tol}_1 = \text{tol}_2 = 10^{-3}$. In case 2, it less than doubles each time the tolerance values are reduced by 3 orders of magnitude, while in case 3 it almost doubles. This suggests that

case 3 is more challenging than case 2 for the partitioned algorithm. We speculate that the average number of iterations over time required by that algorithm grows with number of hotspots. Finally, we note that in case 2 the stopping criterion is always satisfied for every $(h, \Delta t)$ combination, while the partitioned algorithm does not reach convergence for case 3 when $h = 16/100$ and $\Delta t = 2/25$ are selected.

$(h, \Delta t)$	Case 1			Case 2			Case 3		
	10^{-3}	10^{-6}	10^{-9}	10^{-3}	10^{-6}	10^{-9}	10^{-3}	10^{-6}	10^{-9}
$(\frac{16}{400}, \frac{1}{100})$	1.01	1.10	1.14	1.02	1.67	2.65	1.04	2.73	4.88
$(\frac{16}{200}, \frac{1}{50})$	1.01	1.11	1.26	1.04	1.89	2.96	1.08	3.06	5.93
$(\frac{16}{100}, \frac{1}{25})$	1.02	1.13	1.31	1.07	2.11	3.39	1.17	3.16	6.04
$(\frac{16}{100}, \frac{2}{25})$	1.03	1.15	1.38	1.11	2.34	4.17	-	-	-

Table 1: Average number of iterations over time required by the iterative partitioned algorithm to meet the stopping criterion (3.7) with $\text{tol}_1 = \text{tol}_2 = 10^{-3}, 10^{-6}, 10^{-9}$ for the three cases from Sec. 4.1 with different values of mesh size h and time step Δt .

As mentioned in Sec. 3, the iterative partitioned algorithm (3.5)-(3.7) is a fixed point scheme that, at time t^{n+1} , starts the iterations from the solution at time t^n . Hence, if Δt is too large, the solution at time t^n is not a good guess for the solution at time t^{n+1} and the algorithm fails to converge. Tab. 1 shows that if one keeps the same mesh ($h = 16/100$) and halves the time step to $\Delta t = 1/25$, then algorithm (3.5)-(3.7) converges even in the more challenging case. We conclude that Tab. 1 indicates algorithm (3.5)-(3.7) is robust to (reasonable) variations of time step, mesh size, stopping tolerance and parameter η . The low numbers of iterations in Tab. 1 also indicate numerical efficiency, since only two linear systems have to be solved per iteration.

To complement the information reported in Tab. 1, Fig. 15 shows the number of iterations needed for the iterative partitioned algorithm (3.5)-(3.7) to satisfy the stopping criterion with $\text{tol}_1 = \text{tol}_2 = 10^{-3}, 10^{-6}, 10^{-9}$ over time for case 3 with no noise (see results in Fig. 11). We see that for most of the time interval under consideration, the numbers of iterations are close to the averages in Fig. 15. However, sharp peaks occur when the hotspot pattern is forming. In case 3 with no noise, that happens for $t \in [25, 35]$, as confirmed by Fig. 11.

4.3 Results for spatially heterogeneous parameters and a nontrivial geometry

All the results presented so far are associated to constant model parameters. When that is the case, even with the addition of noise, the hotspot distributions are fairly regular. See Fig. 7 and 10. The hotspots observed from visualization of real data (see Fig. 1) lack such regularity. From Fig. 1, it seems that persistent hotspots emerge along highways. This is corroborated by the analysis in [22] based on a dataset spanning all US counties between 1960-1993. It is found that a highway opening led to a 5% rise in the local index crime, driven by property crime. Below, we present cases where persistent hotspots form along hypothetical highways in the simple square geometry considered so far and in a more realistic geometry.

First, we stick to the domain Ω considered so far. The value of η and the initial density (ρ_0 and associated noise) are the same as in case 2. The presence of a highway can

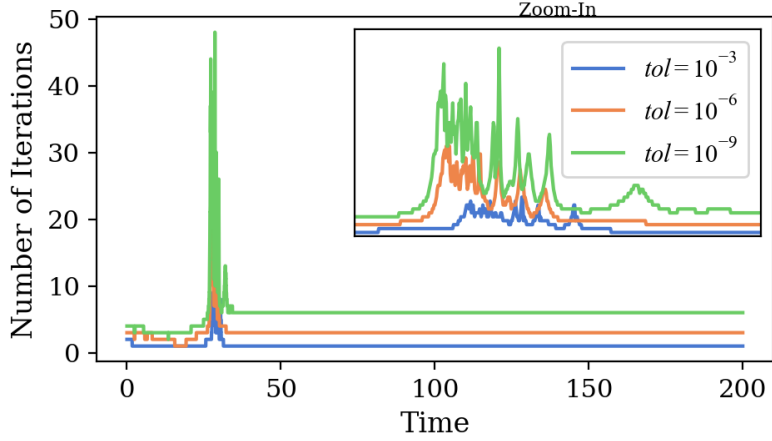


Figure 15: Case 3 without noise: number of iterations needed for the iterative partitioned algorithm (3.5)-(3.7) to satisfy the stopping criterion with $\text{tol}_1 = \text{tol}_2 = 10^{-3}, 10^{-6}, 10^{-9}$ over time. The zoomed-in view highlights the time interval $t \in [25, 35]$, where the peaks in iteration numbers occur.

be modeled through static component A^{st} and “burglar generation” Γ (hence, parameter combination $\Gamma\theta/\omega^2 = B_0$) heterogeneous in space $\mathbf{x} = (x_0, x_1)$. For example, we define them as:

$$A^{st}(\mathbf{x}) = \frac{1}{30} \cdot \exp\{-20 * (x_0 + x_1 - l)^2\} \cdot \left(x_0 > \frac{l}{2} \ \& \ x_1 < \frac{l}{2}\right) + \frac{1}{30} \cdot \chi_1(\mathbf{x}),$$

$$\frac{\Gamma(\mathbf{x})\theta}{\omega^2} = B_0(\mathbf{x}) = \exp\{-20 * (x_0 + x_1 - l)^2\} \cdot \left(x_0 > \frac{l}{2} \ \& \ x_1 < \frac{l}{2}\right) + 1 \cdot \chi_{0.1}(\mathbf{x}),$$

where $l = 16$ is the size of the computational domain. Note that the above A^{st} and B_0 take the values from case 2, introduce randomness through the χ_δ , and add the highway, which is only in the bottom right quadrant, i.e., where $x_0 > l/2$ and $x_1 < l/2$. For initial condition (2.19), we take $\delta = 0.1$ and $\xi_B(\mathbf{x})$ as in case 2.

We consider again mesh size $h = 16/200$ and time step to $\Delta t = 1/50$. Fig. 16 shows the evolution of A and ρ computed by the iterative partitioned algorithm (3.5)-(3.7), with stopping tolerances $\text{tol}_1 = \text{tol}_2 = 10^{-9}$. We observe that by $t = 9$ large hotspots have emerged along the highway, and smaller transient hotspots form in directions parallel to the highway. These hotspots remain transient for the entire duration of the time interval under consideration, while those along the highway persist.

Next, we consider a more realistic domain, with the same time interval of interest (i.e., $t \in (0, 200]$). Since the greater Houston area is mostly a circle, we take the geometry of Chicago [2], which is more interesting. This domain is meshed with 11218 nodes and 22466 elements. The time step is set to $\Delta t = 1/50$. Before adding highway I-90, we run a simulation with spatially homogeneous parameters: $\eta = 0.06$, $A^{st} = 1/30$, $\Gamma\theta/\omega^2 = B_0 = 1$, and $\rho_0 = 0.8$. Uniform Gaussian noise is added to the initial data, with $\chi_\delta(\mathbf{x}) = 1$, $\xi_B(\mathbf{x}) \sim \mathcal{N}(0, 0.05^2)$ and $\xi_\rho(\mathbf{x}) \sim \mathcal{N}(0, 0.01^2)$ in (2.19)–(2.20), sampled independently at each spatial point $\mathbf{x}$. Fig. 17 displays the evolution of A and ρ given by the PDE solver with $\text{tol}_1 = \text{tol}_2 = 10^{-9}$. As expected from the low value of η , we see the formation of persistent small hotspots. The averaged number of iterations over time required by the iterative partitioned algorithm is 5.46, which is in line with the number reported in Tab. 1 for a similar case (case 3, $(16/200, 1/50)$, 10^{-9}).

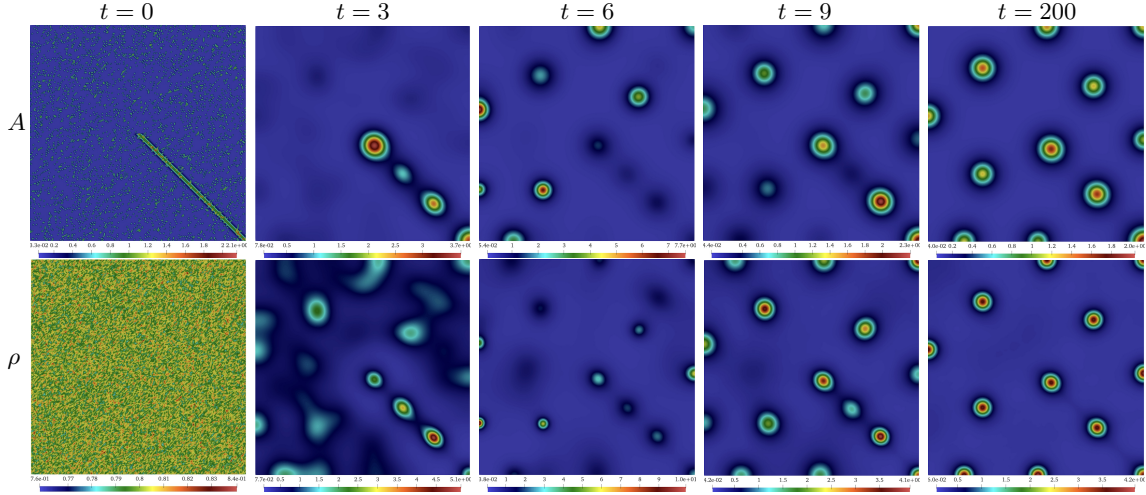


Figure 16: Simple highway case: Evolution of the level of attractiveness (top) and density (bottom).

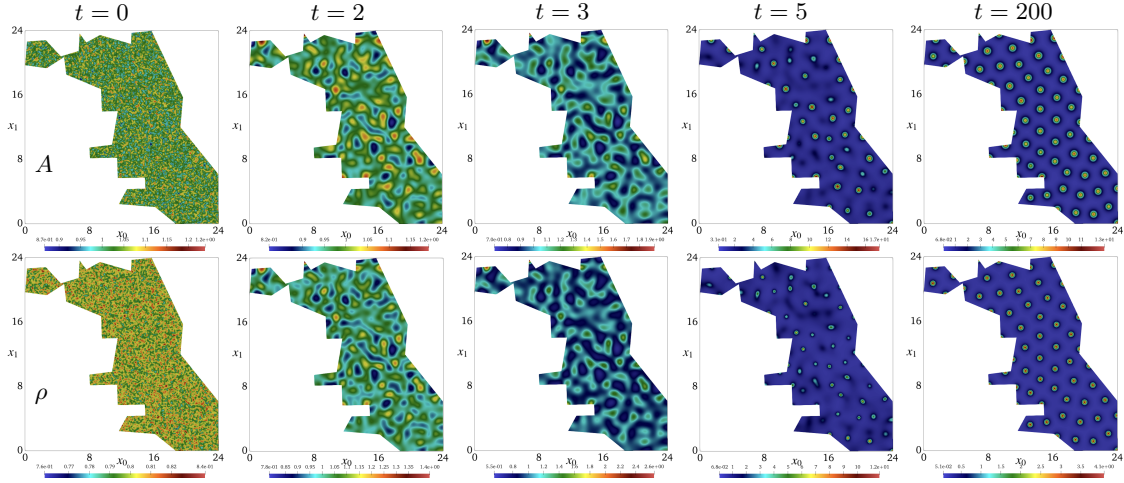


Figure 17: Chicago without highway I-90: Evolution of the level of attractiveness (top) and density (bottom).

In order to add highway I-90, we embed the geometry of Chicago in $[0, 24] \times [0, 24]$. See Fig. 18. We keep the same η and ρ_0 as in the case with no highway, but make A^{st} and Γ dependent on $\mathbf{x}$ as follows:

$$A^{st}(\mathbf{x}) = \begin{cases} \frac{1}{30} \cdot \chi_1(\mathbf{x}) + \frac{1}{30} \cdot \exp \left\{ -20 * (20.8 - x_1)^2 \right\}, & \text{if } x_0 < 9.5, \\ \frac{1}{30} \cdot \chi_1(\mathbf{x}) + \frac{1}{30} \cdot \exp \left\{ -20 * \left(32.4923 - \frac{16}{13}x_0 - x_1 \right)^2 \right\}, & \text{if } x_0 \geq 9.5. \end{cases}$$

$$\frac{\Gamma(\mathbf{x})\theta}{\omega^2} = B_0(\mathbf{x}) = \begin{cases} 1 \cdot \chi_{0.1}(\mathbf{x}) + \exp \left\{ -20 * (20.8 - x_1)^2 \right\}, & \text{if } x_0 < 9.5, \\ 1 \cdot \chi_{0.1}(\mathbf{x}) + \exp \left\{ -20 * \left(32.4923 - \frac{16}{13}x_0 - x_1 \right)^2 \right\}, & \text{if } x_0 \geq 9.5. \end{cases}$$

Note that to create the highway we have followed the same strategy used in the simple case (square geometry). Gaussian noise is added to the initial data, with $\chi_{\delta_B}(\mathbf{x}) = 0.1$, $\chi_{\delta_\rho}(\mathbf{x}) = 1$, $\xi_B(\mathbf{x}) \sim \mathcal{N}(0, 0.05^2)$ and $\xi_\rho(\mathbf{x}) \sim \mathcal{N}(0, 0.01^2)$ in (2.19)–(2.20), sampled independently at each spatial point $\mathbf{x}$. Fig. 18 shows the evolution of A and ρ given by the PDE solver

with $\text{tol}_1 = \text{tol}_2 = 10^{-9}$. We observe the formation of small persistent hotspots aligned with the highway, with smaller transient hotspots in directions parallel to the highway. The averaged number of iterations over time required by the iterative partitioned algorithm is 6.39, which is slightly higher than the case with no highway.

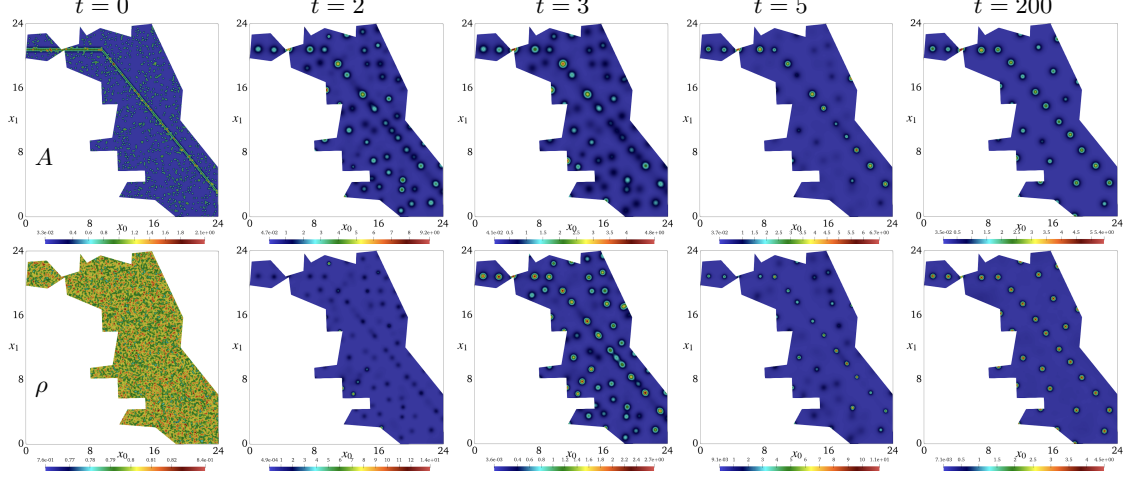


Figure 18: Chicago with highway I-90: Evolution of the level of attractiveness (top) and density (bottom).

In this paper, we have shown the evolution of the variables in the continuum model. However, from a practical point of view, one is interested in the number of burglaries over a certain period of time, which is the quantity shown in Fig. 1. Such number of b comes from the integral of $E_s(t)$ in (2.6) over the given time interval and it requires dimensional variables. This means that specific values of parameters θ and ω need to be assigned, as explained below, while the PDE model (2.12)-(2.13) needs only the value of $\Gamma\theta/\omega^2$.

Let us assume that we are interested in the number of burglaries n_B occurred over each element $Q_k \in \mathcal{T}_h$ during the entire interval $(0, T)$:

$$n_B|_{Q_k} = \int_0^T E_s(t)|_{Q_k} dt \approx \sum_{n=0}^{N_t-1} E_s(t^n)|_{Q_k} \Delta t = \sum_{n=0}^{N_t-1} n_s(t^n)|_{Q_k} p_s(t^n)|_{Q_k} \Delta t$$

Quantities $n_s(t^n)|_{Q_k}$ and $p_s(t^n)|_{Q_k}$ can be obtained from the finite element approximation of the continuum model:

$$n_s(t^n)|_{Q_k} = \int_{Q_k} \epsilon L^2 \rho_h^n d\mathbf{x} \approx \frac{\theta h^2}{\omega} \int_{Q_k} \rho_h^n d\mathbf{x}, \quad p_s(t^n)|_{Q_k} = 1 - e^{-A_h^n|_{Q_k} \Delta t}.$$

Thus, $n_B|_{Q_k}$ could be obtained from post-processing the finite element solution of the continuum model as follows:

$$n_B|_{Q_k} \approx \frac{\theta h^2 \Delta t}{\omega} \sum_{n=0}^{N-1} \left(1 - e^{-A_h^n|_{Q_k} \Delta t}\right) \int_{Q_k} \rho_h^n d\mathbf{x}.$$

This shows that ratio $\Gamma\theta/\omega^2$ is not enough to obtain n_B from the computed level of attractiveness and density of criminals. The ratio θ/ω needs to be provided as well.

5 Future perspectives

We believe the work presented in this paper is a stepping stone towards the application of the PDE model to predict residential burglary in real cities. Imminent future work will

be focused on making the simulations more and more realistic by, e.g., considering more realistic geometries, incorporating real data on property (see [53]), adding to the model police intervention and community investment, and developing data-driven methods for parameter estimation. As shown in Sec. 4, solutions are sensitive to key parameters η , A^{st} , and $\Gamma\theta/\omega^2$. However, as mentioned in Sec. 1, the scale of the PDE model (i.e., the macroscopic scale) is only one of three possible scales of observation and complex self-organizing, collective dynamics, such as crime, are best represented by a mathematical description covering all scales.

Systems like crime are made of living entities (as opposed to inert matter) that possess the behavioral ability to develop and adapt moving strategies in response to the context, which significantly complicates their mathematical modeling. A general mathematical structure able to capture the specific complexity features of crime as a living system at all scales is still missing. While there exists a sizeable body of literature on microscopic and macroscopic models (see Sec. 1), the mesoscopic scale has received much less attention. Taking inspiration from the kinetic theory of gases, mesoscopic (or kinetic) models derive a Boltzmann-type evolution equation for a statistical distribution function. In these models, people are seen as active particles, as opposed to classical particles in gases. This leads to a key difference: the interactions in kinetic models for gas dynamics are conservative and reversible, while the interactions in the kinetic models for living systems are irreversible, non-conservative and, in some cases, nonlocal and nonlinearly additive. The reader interested in the kinetic theory of active particles is referred to [9, 10]. In general, Ref. [10] provides the mathematical tools necessary for the approaches outlined in this section. To enable a *multiscale framework*, models must be derived at each scale using the same principles and similar parameters. It is worth noticing that substantial model development at the three different scales, even with the same principles and analogous parameters for an integrated multiscale vision, has been done for a different kind of collective social system: crowd dynamics [13, 14]. So, existing theory for crowd dynamics could be extended to crime dynamics.

The concept of *activity variable* mentioned in Sec. 1 belongs to the kinetic theory of active particles. It is a vector variable at the microscopic scale that collects a number of behavioral variables, such as emotional state, and changes through interactions with others. In the context of crowd dynamics, several works include the activity variable and some model its dynamics through a *transport equation* [13]. With the exception of [11], the concept of activity variable remains unexplored for models of crime dynamics. We believe the activity variable, and corresponding transport equation, is the key to modeling behavioral heterogeneity in criminals and policemen. Ideally, following H. A. Simon's theories, models should account for rules of interactions that depend on the environment where interactions occur, with the environment in turn evolving as a result of human intervention [48, 12, 10].

Finally, we would like to mention the importance of developing “*multi-physics*” models that couple crime dynamics to other factors correlated with crime, such as wealth inequalities [30] and level of education [8]. Ideas on how to do this at the macroscopic level are discussed in [15]. It is also imperative to think about safety checks for these models, and constant improvement through fact-based evidence, to avoid turning them into weapons of math destruction [37].

6 Conclusions

We introduced a finite element approximation of a well-known PDE model to predict residential burglary. For the numerical solution of the discretized PDE problem, we proposed an algorithm that decouples the computation of the model variables (attractiveness of residential sites and density of criminals) at each time step, resulting in the solution of two linear algebraic systems per iteration. We validated the numerical results obtained with our approach against the results provided by the agent-based model from which the PDE model was derived. We showed that the PDE solver output matches well the output from the agent-based simulations when the density of criminals is large, i.e., when the criminal population can be looked at as a continuum versus separate individuals. Through numerical experimentation, we have characterized the dependence of both diameter and number of crime hotspots on the strength of the neighborhood effects. Concerning the solution algorithm, we showed that the average number of iterations over time required to meet the stopping criterion ranges between 1 and 6 when model parameters, discretization and algorithmic parameters are varied by orders of magnitude. These low numbers of iterations indicate robustness to parameter variations and numerical efficiency, since only two linear systems have to be solved per iteration. Thanks to the flexibility allowed by the finite element method, we simulated crime dynamics with spatially heterogeneous model parameters and a realistic geometry.

Finally, while the robust and efficient code developed for this paper has a value per se, it is mostly intended as the solid base for the broader research program described in Sec. 5.

Acknowledgments

We thank Maria Lozano for the visualization of the 2024 Houston police data in Fig. 1. Zhong is supported by NSF-AoF grant #2225507. Zhong also gratefully acknowledges funding provided by the Oak Ridge Associated Universities Ralph E. Powe Junior Faculty Enhancement Award for FY2024.

References

- [1] A Finite Element and Agent-Based Modeling Framework for Urban Crime Dynamics. <https://github.com/baolihao/UrbanCrime-Sim>.
- [2] Chicago Data Portal - City Boundaries. <https://data.cityofchicago.org/Facilities-Geographic-Boundaries/Boundaries-City-Map/ewy2-6yfk>.
- [3] Houston Police Department - Monthly Crime Data by Street and Police Beat. https://www.houstontx.gov/police/cs/Monthly_Crime_Data_by_Street_and_Police_Beat.htm.
- [4] QGIS webpage. <https://qgis.org/project/overview/>.
- [5] N. D. Alikakos, G. Fusco, and G. Karali. Ostwald ripening in two dimensions—the rigorous derivation of the equations from the Mullins–Sekerka dynamics. *J. Differ. Equ.*, 205(1):1–49, 2004.
- [6] M. S. Alnaes, A. Logg, K. B. Ølgaard, M. E. Rognes, and G. N. Wells. Unified Form Language: a domain-specific language for weak formulations of partial differential equations. *ACM Trans. Math. Softw.*, 40, 2014.

- [7] I. A. Baratta, J. P. Dean, J. S. Dokken, M. Habera, J. S. Hale, C. N. Richardson, M. E. Rognes, M. W. Scroggs, N. Sime, and G. N. Wells. DOLFINx: the next generation FEniCS problem-solving environment. preprint, 2023.
- [8] E. Jason Baron, Joshua Hyman, and Brittany Vasquez. Public school funding, school quality, and adult crime. *Rev. Econ. Stat.*, pages 1–46, 04 2024.
- [9] N. Bellomo, D. Burini, G. Dosi, L. Gibelli, D. Knopoff, N. Outada, P. Terna, and M. E. Virgillito. What is life? a perspective of the mathematical kinetic theory of active particles. *Math. Models Methods Appl. Sci.*, 31(09):1821–1866, 2021.
- [10] N. Bellomo, D. Burini, and J. Liao. New trends in kinetic theory towards the complexity of living systems. *Math. Models Methods Appl. Sci.*, 36(02):to appear, 2026.
- [11] N. Bellomo, F. Colasuonno, D. Knopoff, and J. Soler. From a systems theory of sociology to modeling the onset and evolution of criminality. *Netw. Heterog. Media*, 2015.
- [12] N. Bellomo and M. Egidi. From Herbert A. Simon’s legacy to the evolutionary artificial world with heterogeneous collective behaviors. *Math. Models Methods Appl. Sci.*, 34(01):145–180, 2024.
- [13] N. Bellomo, L. Gibelli, A. Quaini, and A. Reali. Towards a mathematical theory of behavioral human crowds. *Math. Models Methods Appl. Sci.*, 32(02):321–358, 2022.
- [14] N. Bellomo, J. Liao, A. Quaini, L. Russo, and C. Siettos. Human behavioral crowds review, critical analysis and research perspectives. *Math. Models Methods Appl. Sci.*, 33(08):1611–1659, 2023.
- [15] N. Bellomo, N. Outada, J. Soler, Y. Tao, and M. Winkler. Chemotaxis and cross-diffusion models in complex environments: Models and analytic problems toward a multiscale vision. *Math. Models Methods Appl. Sci.*, 32(04):713–792, 2022.
- [16] A. Blumstein. Crime modeling. *Oper. Res.*, 50(1):16–24, 2002.
- [17] P. J. Brantingham and P. L. Brantingham. *Patterns in crime*. Macmillan, New York, 1984.
- [18] A. J. Bray. Theory of phase-ordering kinetics. *Adv. Phys.*, 51(2):481–587, 2002.
- [19] M. Burger, M. Di Francesco, and Y. Dolak-Struss. The Keller-Segel model for chemotaxis with prevention of overcrowding: linear vs. nonlinear diffusion. *SIAM J. Math. Anal.*, 38(4):1288–1315, 2006.
- [20] H. Byrne and M. Owen. A new interpretation of the Keller-Segel model based on multiphase modelling. *J. Math. Biol.*, 49:604–626, 2004.
- [21] Y. Cai, C. Wang, and Y. Zhang. A multiscale stochastic criminal behavior model and the convergence to a piecewise-deterministic-Markov-process limits. *Math. Models Methods Appl. Sci.*, 32(4):619–645, 2022.
- [22] F. Calamunci and J. Lonsky. Highway to hell? Interstate Highway System and crime. *IZA Inst. Labor Econ. Rep. No. 15800*, 2022.
- [23] D. Castelvechi. Mathematicians urge colleagues to boycott police work in wake of killings. *Natures News*, 2020.

- [24] S. Chainey and J. Ratcliffe. *GIS and crime mapping*. John Wiley & Sons, Chichester, 2005.
- [25] M. R. D’Orsogna and M. Perc. Statistical physics of crime: A review. *Phys. Life Rev.*, 12:1–21, 2015.
- [26] J. Eck and D. Weisburd. *Crime places in crime theory*, volume 4, pages 1–33. Criminal Justice Press, 1995.
- [27] F. Ersöz, T. Ersöz, F. Marcelloni, and F. Ruffini. Artificial intelligence in crime prediction: A survey with a focus on explainability. *IEEE Access*, 13:59646–59674, 2025.
- [28] M. Fuest and S. Heydari. A cross-diffusion system modeling rivaling gangs: Global existence of bounded solutions and FCT stabilization for numerical simulation. *Math. Models Methods Appl. Sci.*, 34(09):1739–1779, 2024.
- [29] C. Haskins. Academics confirm major predictive policing algorithm is fundamentally flawed. *Vice Tech*, 2019.
- [30] E. Itskovich and R. Factor. Economic inequality and crime: The role of social resistance. *J. Crim. Justice*, 86:102065, 2023.
- [31] S. D. Johnson, K. Bowers, and A. Hirschfield. New insights into the spatial and temporal distribution of repeat victimization. *Br. J. Criminology*, 37:224–244, 1997.
- [32] P. A. Jones, P. J. Brantingham, and L. R. Chayes. Statistical models of criminal behavior: the effects of law enforcement actions. *Math. Models Methods Appl. Sci.*, 20(supp01):1397–1423, 2010.
- [33] E. F. Keller and L. A. Segel. Initiation of slime mold aggregation viewed as an instability. *J. Theor. Biol.*, 26(3):399–415, 1970.
- [34] L. Mei and J. Wei. The existence and stability of spike solutions for a chemotaxis system modeling crime pattern formation. *Math. Models Methods Appl. Sci.*, 30(9):1727–1764, 2020.
- [35] G. Mohler, A. L. Bertozzi, J. Carter, M. B. Short, D. Sledge, G. E. Tita, C. D. Uchida, and P. J. Brantingham. Impact of social distancing during COVID-19 pandemic on crime in Los Angeles and Indianapolis. *J. Crim. Justice*, 68:101692, 2020.
- [36] G. Mohler, P. J. Brantingham, J. Carter, and M. B. Short. Reducing bias in estimates for the law of crime concentration. *J. Quant. Criminol.*, 35(4):747–765, 2019.
- [37] C. O’Neil. *Weapons of Math Destruction*. Crown New York, 2016.
- [38] C. Pan, B. Li, C. Wang, Y. Zhang, N. Geldner, L. Wang, and A. L. Bertozzi. Crime modeling with truncated Lévy flights for residential burglary models. *Math. Models Methods Appl. Sci.*, 28(9):1857–1880, 2018.
- [39] J. Park, F. P. Schoenberg, A. L. Bertozzi, and P. J. Brantingham. Investigating clustering and violence interruption in gang-related violent crime data using spatial–temporal point processes with covariates. *J. Am. Stat. Assoc.*, 116(536):1674–1687, 2021.
- [40] A. Quarteroni and A. Valli. *Numerical approximation of partial differential equations*. Springer, 1994.

- [41] N. Rodriguez and A. L. Bertozzi. Local existence and uniqueness of solutions to a pde model for criminal behavior. *Math. Models Methods Appl. Sci.*, 20:1425–1457, 2010.
- [42] N. Rodríguez and M. Winkler. Relaxation by nonlinear diffusion enhancement in a two-dimensional cross-diffusion model for urban crime propagation. *Math. Models Methods Appl. Sci.*, 30(11):2105–2137, 2020.
- [43] Y. Saad. *Iterative methods for sparse linear systems*. SIAM, 2003.
- [44] M. W. Scroggs, I. A. Baratta, C. N. Richardson, and G. N. Wells. Basix: a runtime finite element basis evaluation library. *J. Open Source Softw.*, 7(73):3982, 2022.
- [45] M. B. Short, A. L. Bertozzi, and P. J. Brantingham. Nonlinear patterns in urban crime: hotspots, bifurcations, and suppression. *SIAM J. Appl. Dyn. Syst.*, 9(2):462–483, 2010.
- [46] M. B. Short, P. J. Brantingham, A. L. Bertozzi, and G. E. Tita. Dissipation and displacement of hotspots in reaction-diffusion models of crime. *Proc. Natl. Acad. Sci. U.S.A.*, 107(9):3961–3965, 2010.
- [47] M. B. Short, M. R. D’Orsogna, V. B. Pasour, G. E. Tita, P. J. Brantingham, A. L. Bertozzi, and L. B. Chayes. A statistical model of criminal behavior. *Math. Models Methods Appl. Sci.*, 18(supp01):1249–1267, 2008.
- [48] H.A. Simon. *The science of the artificial*. MIT Press, Boston, 2019 (third edition).
- [49] B. Wang, P. Yin, A.L. Bertozzi, P.J. Brantingham, S.J. Osher, and J. Xin. Deep learning for real-time crime forecasting and its ternarization. *Chin. Ann. Math. Ser. B*, 40:949 – 966, 2019.
- [50] B. Wang, D. Zhang, D. Zhang, P. J. Brantingham, and A. L. Bertozzi. Deep learning for real time crime forecasting. In *Proc. 2017 Int. Symp. Nonlinear Theory Its Appl. NOLTA*, Dec 2017.
- [51] D. Weisburd. The law of crime concentration and the criminology of place. *Criminology*, 53(2):133–157, 2015.
- [52] J. Q. Wilson and G. L. Kelling. Broken windows and police and neighborhood safety. *Atlantic Mon.*, 249:29–38, 1982.
- [53] J.T. Woodworth, G.O. Mohler, A.L. Bertozzi, and P.J. Brantingham. Non-local crime density estimation incorporating housing information. *Phil. Trans. R. Soc. A*, 372:20130403, 2014.
- [54] R. Wright and S. Decker. *Burglars on the job*. Northeastern Univ. Press, 1994.
- [55] J. R. Zipkin, M. B. Short, and A. L. Bertozzi. Cops on the dots in a mathematical model of urban crime and police response. *Discrete Contin. Dyn. Syst., Ser. B*, 19(5):1479–1508, 2014.