From Agent-Based Models to PDEs: Coarse-Graining a Host-Pathogen Model

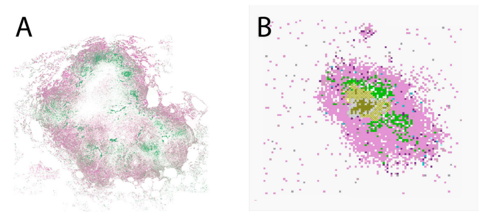
Wesley J. M. Ridgway

Department of Mathematics & Statistics
University of Saskatchewan

2026 Go20 Conference on Scientific Computing and Software
May 20

Motivation

- mathematical modelling of bacteria is often a cheaper and faster alternative than performing experiments *in vitro*
- but only if the model is not computationally expensive
- Agent-based models (ABMs) are used frequently



Granuloma in a nonhuman primate (A) and snapshot from an ABM (B). Adapted from (Marino *et al.*, 2015).

Motivation

Some Advantages of ABMs

- cell-level behaviour can be encoded explicitly through simple rules
- good software packages for implementing ABMs already exist, e.g. PhysiCell (Ghaffarizadeh *et al.*, 2018), GranSim (Segovia-Juarez, Ganguli, Kirschner, 2004)
- highly customizable

Motivation

ABMs have some drawbacks:

- computational cost scales with the number of agents
- many realizations are required to generate meaningful statistics (large ensembles)
- contains much more information than is needed

Motivation

Parameter sweeps with ABMs in high-dimensional parameter space + ensemble averaging is expensive. ⇒ Can we do better?

Idea:

Given a host-pathogen ABM:

1. develop a coarse-grained description of the ABM in terms of PDEs
2. perform low resolution parameter searches with the coarse-grained model
3. refine search as necessary with the ABM

Previous Modelling Efforts

Some recurring rules across various host-pathogen ABMs

- movement of cells and bacteria
- ingestion of bacteria by immune cells
- intracellular killing of bacteria
- intracellular reproduction/growth of bacteria
- bacteria reproduction
- immune cell death
- chemical signalling

- see e.g. (Segovia-Juarez, Ganguli, Kirschner, 2004), (An, *et al.*, 2009), (Dancik, Jones, Dorman, 2010), (Budak, *et al.*, 2025)

A Simple Agent-Based Model

We construct an ABM consisting of immune cell agents and bacteria agents with the following rules

- movement of cells and bacteria
- ingestion of bacteria
- phagocytosis (intracellular killing of bacteria)
- phagocytosis resistance (intracellular reproduction of bacteria)

Movement

In the absence of chemical signalling, its common for movement to be modelled with simple Brownian motion.

Movement Rules

Let $\mathbf{x}_k^{(b)}(t)$ and $\mathbf{x}_k^{(I)}(t)$ denote the position (in $\mathbb{R}^3$) of the immune cell and bacteria agents at time t. Then

$$\mathrm{d}\mathbf{x}_k^{(b)} = \sqrt{2D_b}\mathrm{d}W_k^{(b)}, \quad k = 1, \dots M_b, \text{ (bacteria)}$$
$$\mathrm{d}\mathbf{x}_k^{(I)} = \sqrt{2D_I}\mathrm{d}W_k^{(I)}, \quad k = 1, \dots, M_I, \text{ (immune cells)}$$

where $W_k^{(b)}$, $W_k^{(I)}$, are independent Weiner processes, and D_b and D_I are the associated diffusion coefficients.

We can ignore volume exclusion at low cell densities.

Phagocytosis Model

- Each immune cell agent is equipped with a state variable j that tracks the number of ingested bacteria
- Immune cell agents have a maximum capacity of N, i.e. $0 \leq j \leq N$.

Ingestion Rules

For each immune cell/bacterium pair, do the following:

1. Compute $d_{ik}(t) = |\mathbf{x}_i^{(b)}(t) - \mathbf{x}_k^{(I)}(t)|$.
2. While $d_{ik} < \bar{\rho}$ for some user-specified $\bar{\rho}$, delete (ingest) the bacterium agent with probability per unit time λ.
3. If the bacterium is ingested by an immune cell agent, increment this agent's state variable by one.

Phagocytosis Model

- after being ingested, bacteria are destroyed in a process called phagocytosis (Uribe-Querol, Rosales, 2020)
- some bacteria can survive phagocytosis, reproduce intracellularly, and eventually kill the host immune cell

Intracellular Rules

For each immune cell agent with state index j, do the following:

1. decrement the state variable j by one with probability per unit time q_j (death by phagocytosis)
2. increment the state variable j by one with probability per unit time r_j (intracellular reproduction)

Agent-Based Model

- stochastic rules (and initial conditions) imply that each realization will produce different results in general
- hence meaningful statistics require averaging over numerous independent runs (ensemble averaging)
- we simulate the model with the open source software PhysiCell (Ghaffarizadeh *et al.*, 2018)

Coarse Graining

Goal: Develop an alternative model that accurately reproduces ensemble averages.

We do this in two steps:

1. recast the ABM as a reaction-diffusion PDE system
2. take a suitable continuum limit

Coarse-Graining

Recast the ABM as a reaction-diffusion system.

- bacteria and cells are represented by continuous concentration fields $b(\mathbf{x}, t)$ and $u_j(\mathbf{x}, t)$, $j = 0, \ldots, N$.
- u_j denotes the concentration of immune cells with j internal bacteria

$$\frac{\partial b}{\partial t} = D_b \nabla^2 b - \sum_{j=0}^{N-1} k_j u_j b, \tag{1}$$

$$\frac{\partial u_0}{\partial t} = D_I \nabla^2 u_0 - k_0 u_0 b - r_0 u_0 + q_1 u_1, \tag{2}$$

$$\begin{aligned} \frac{\partial u_j}{\partial t} = {} & D_I \nabla^2 u_j + \left(k_{j-1} u_{j-1} - k_j u_j\right) b \\ & + r_{j-1} u_{j-1} - r_j u_j + q_{j+1} u_{j+1} - q_j u_j, \quad j = 1, \ldots N-1, \end{aligned} \tag{3}$$

$$\frac{\partial u_N}{\partial t} = D_I \nabla^2 u_N + k_{N-1} u_{N-1} b + r_{N-1} u_{N-1} - q_N u_N. \tag{4}$$

blue: ingestion

orange: phagocytosis

The reaction-diffusion system is semi-discrete

- cells are continuous in space and time; but
- discrete in internal state

Question: how do we calculate the ‘reaction rates’ k_j?

A Simple Model of a Bimolecular Reaction

Consider the (irreversible) chemical reaction between reactant molecules A and B:

$$\mathrm{A} + \mathrm{B} \to \mathrm{C}$$

The $\lambda - \bar{\rho}$ Model (Erban, Chapman, 2009)

- reactant molecules move randomly (Brownian Motion) with diffusion coefficients D_A and D_B
- react with probability per unit time λ when within a distance $\bar{\rho}$
- $\lambda = \infty$ means the molecules react whenever a collision occurs

A Simple Model of a Bimolecular Reaction

$$\mathrm{A} + \mathrm{B} \rightarrow \mathrm{C}$$

In a well-mixed system, we expect the reaction to be approximated by the ODE system

$$\begin{aligned}\frac{\mathrm{d}[\mathrm{A}]}{\mathrm{d}t} &= -k[\mathrm{A}][\mathrm{B}],\\ \frac{\mathrm{d}[\mathrm{B}]}{\mathrm{d}t} &= -k[\mathrm{A}][\mathrm{B}],\\ \frac{\mathrm{d}[\mathrm{C}]}{\mathrm{d}t} &= +k[\mathrm{A}][\mathrm{B}],\end{aligned}$$

where k is the reaction rate constant and $[(\cdot)]$ denotes the concentration of $(\cdot)$.

The Effective Reaction Rate k

Heuristic Procedure

- consider a single molecule of B in a concentration field of molecules of A
- in the reference frame of B, the diffusion coefficent of A is $D = D_A + D_B$.
- over sufficiently long timescales, the concentration of A satisfies

$$\begin{aligned} 0 &= D\nabla^2[\mathrm{A}], \quad |\mathbf{x}| > \bar{\rho}, \\ 0 &= D\nabla^2[\mathrm{A}] - \lambda[\mathrm{A}], \quad |\mathbf{x}| < \bar{\rho}, \\ [\mathrm{A}] &\to \mathrm{const.} \quad \text{as } |\mathbf{x}| \to \infty. \end{aligned}$$

Heuristic Procedure (continued)

- the reaction rate k is then calculated in terms of the flux of A out of the system, i.e.

$$k\left(\lim_{|\mathbf{x}|\to\infty}[\mathrm{A}]\right)=\int_{|\mathbf{x}|=\bar{\rho}} D\nabla[\mathrm{A}]\cdot\mathrm{d}\mathbf{x}.$$

- $k=4\pi D\bar{\rho}$ if $\lambda=\infty$
- $k=4\pi D\left(\bar{\rho}-\sqrt{\frac{D}{\lambda}}\tanh\left(\bar{\rho}\sqrt{\frac{\lambda}{D}}\right)\right)$ if $0<\lambda<\infty$

When $\lambda=\infty$, a more rigorous approach shows that this procedure is justified on timescales $t\gg(\bar{\rho})^2/D$; see (Doi, 1976).

Back to our Coarse-Grained System

$$\begin{aligned}
\frac{\partial b}{\partial t} &= D_b \nabla^2 b - \sum_{j=0}^{N-1} k_j u_j b, \\
\frac{\partial u_0}{\partial t} &= D_l \nabla^2 u_0 - k_0 u_0 b - r_0 u_0 + q_1 u_1, \\
\frac{\partial u_j}{\partial t} &= D_l \nabla^2 u_j + \left(k_{j-1} u_{j-1} - k_j u_j\right) b \\
&\quad + r_{j-1} u_{j-1} - r_j u_j + q_{j+1} u_{j+1} - q_j u_j, \quad j = 1, \dots N-1, \\
\frac{\partial u_N}{\partial t} &= D_l \nabla^2 u_N + k_{N-1} u_{N-1} b + r_{N-1} u_{N-1} - q_N u_N.
\end{aligned}$$

Employing a similar procedure with $\lambda = \infty$, $D = D_l + D_b$, we want to solve

$$\begin{aligned} D\nabla^2 u_0 &= r_0 u_0 - q_1 u_1, \\ D\nabla^2 u_j &= -r_{j-1} u_{j-1} + (r_j + q_j) u_j - q_{j+1} u_{j+1}, \quad j = 1, \dots N-1, \\ D\nabla^2 u_N &= -r_{N-1} u_{N-1} + q_N u_N, \end{aligned}$$

outside the sphere $|\mathbf{x}| > \bar{\rho}$ and

$$D\nabla^2 u_N = q_N u_N,$$

inside the sphere $|\mathbf{x}| < \bar{\rho}$.

- far-field conditions $u_j \to \mathrm{const.}$ as $|\mathbf{x}| \to \infty$
- near-field condition $u_j = 0$ on $|\mathbf{x}| = \bar{\rho}, \quad j = 0, \dots, N-1$

Seeking a solution in the form $(u_0, \dots, u_N) := \mathbf{u} = \frac{1}{\rho}\mathbf{v}$, where $\rho = |\mathbf{x}|$, we find the general solution

$$\mathbf{u}(\rho) = (A + B\rho^{-1})\mathbf{a}_0 + \sum_{k=1}^{N} \mathbf{a}_k \left(C_k e^{-\sqrt{\lambda_k}\rho} + \tilde{C}_k e^{\sqrt{\lambda_k}\rho} \right) \rho^{-1}, \quad \rho > \bar{\rho},$$

where A, B, C_k, and $\tilde{C}_k$ are constants; $(\lambda_k, \mathbf{a}_k)$ are eigenpairs of

$$\mathcal{M} = \frac{1}{D} \begin{pmatrix} r_0 & -q_1 & 0 & \dots & 0 \\ -r_0 & r_1 + q_1 & -q_2 & \dots & 0 \\ 0 & -r_1 & \ddots & & \vdots \\ \vdots & & & -q_{N-1} & 0 \\ 0 & & -r_{N-2} & r_{N-1} + q_{N-1} & -q_N \\ 0 & & 0 & -r_{N-1} & q_N \end{pmatrix}.$$

Inside the sphere, we solve

$$\begin{aligned} D\nabla^2 u_N &= q_N u_N, \\ u_N(\rho) &\text{ finite as } \rho \to 0^+, \end{aligned}$$

which yields

$$u_N(\rho) = \frac{C}{\rho}\sinh\left(\frac{\mu\rho}{\bar{\rho}}\right), \quad \mu = \sqrt{\frac{q_N}{D}}\bar{\rho}, \quad \rho < \bar{\rho}.$$

Enforcing boundary and regularity conditions gives a linear system for the unknown constants A, B, C, C_k (note $\tilde{C}_k = 0$).

$$a_{0N} B - C \sinh \mu + \sum_{k=1}^{N} C_k a_{kN} e^{-\mu_k} = -a_{0N} \bar{\rho} A,$$

$$a_{0N} B + (\mu \cosh \mu - \sinh \mu)\, C + \sum_{k=1}^{N} a_{kN} (1 + \mu_k) C_k e^{-\mu_k} = 0,$$

$$a_{0j} B + \sum_{k=1}^{N} a_{kj} C_k e^{-\mu_k} = -a_{0j} \bar{\rho} A, \quad j = 0, \ldots N-1.$$

where $\mu_k = \sqrt{\lambda_k} \bar{\rho}$.

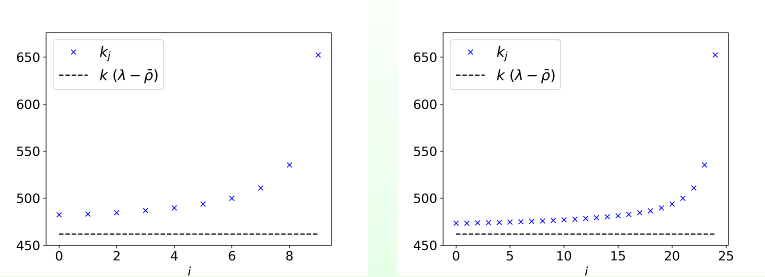
- The solution is unique up to an overall scaling

We then calculate the reaction rates k_j with

$$k_j\left(\lim_{|\mathbf{x}|\to\infty} u_j\right) = \int_{|\mathbf{x}|=\bar{\rho}} D\nabla u_j \cdot \mathrm{d}\mathbf{x}.$$

$$k_j = -\frac{4\pi D}{a_{0j}A}\left(a_{0j}B + \sum_{k=1}^{N} a_{kj}C_k(1+\mu_k)e^{-\mu_k}\right).$$

- we expect this approximation to be valid on sufficiently long timescales



Effective reaction rates k_j with $N = 10$ (left) and $N = 25$ (right)

.

Parameter values:

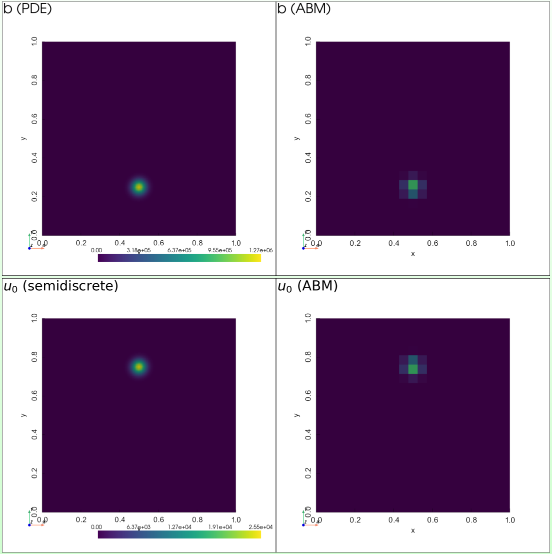
- $r_j = 0.2,\ q_j = 0.1,\ D_I = 5,\ D_b = 2$
- immune cell volume: $\frac{4}{3}\pi R_I^3 = 100$
- bacteria volume: $\frac{4}{3}\pi R_b^3 = 1$
- $\bar{\rho} = 1.5(R_b + R_I),\ \lambda = \infty$

Results

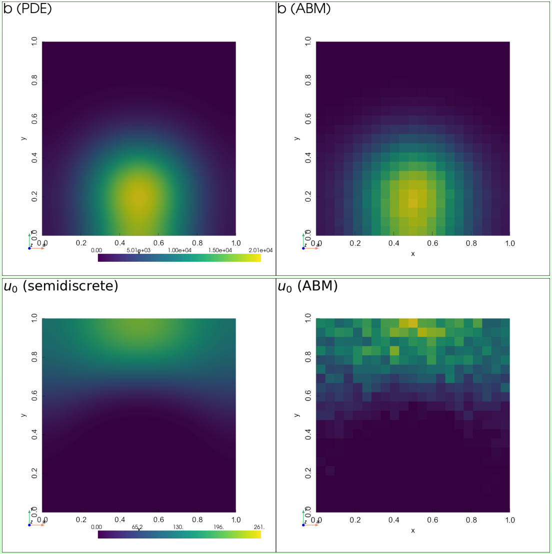
Comparison with ABM

- Define $U_j(t) := \int u_j(\mathbf{x}, t)\,\mathrm{d}\mathbf{x}$ and $B = \int b(\mathbf{x}, t)\,\mathrm{d}\mathbf{x}$
- U_j and B represent the cell and bacteria counts in the PDE system
- These can be compared with the ABM cell counts, averaged over many independent realizations.

Initial Condition



Solution Snapshots at $t = 10$

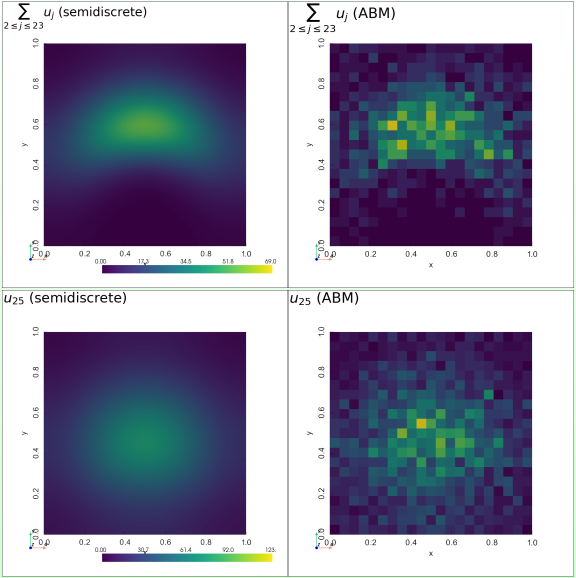


b (PDE)

b (ABM)

u_0 (semidiscrete)

u_0 (ABM)

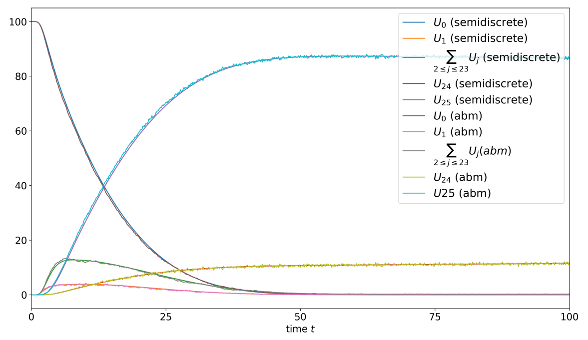


$\sum_{2 \le j \le 23} u_j$ (semidiscrete)

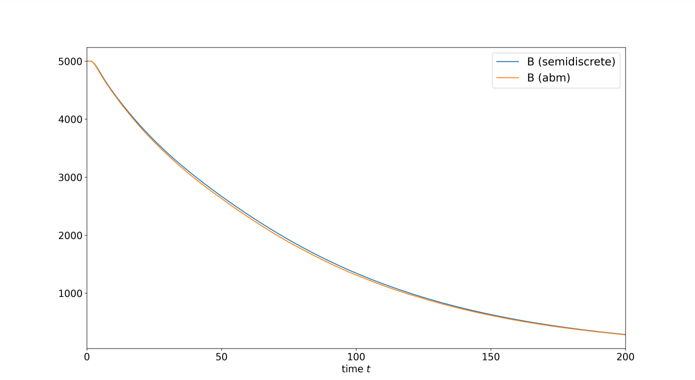
$\sum_{2 \le j \le 23} u_j$ (ABM)

u_{25} (semidiscrete)

u_{25} (ABM)



Immune cell counts. Parameter values: $N = 25$, $r_j = j\left(1 - \frac{j}{N+30}\right)$, $q_j = 2$, $D_I = 5$, $D_b = 2$. Immune cell volume $= 6 \times 10^{-7}$, bacteria volume $= 5 \times 10^{-9}$.


5000
4000
3000
2000
1000
0
50
100
150
200
time t
B (semidiscrete)
B (abm)

Continuum Approximation

The semidiscrete system

$$\begin{aligned}
\frac{\partial b}{\partial t} =& D_b \nabla^2 b - \sum_{j=0}^{N-1} k_j u_j b, \\
\frac{\partial u_0}{\partial t} =& D_I \nabla^2 u_0 - k_0 u_0 b - r_0 u_0 + q_1 u_1, \\
\frac{\partial u_j}{\partial t} =& D_I \nabla^2 u_j + \left(k_{j-1} u_{j-1} - k_j u_j\right) b \\
& + r_{j-1} u_{j-1} - r_j u_j + q_{j+1} u_{j+1} - q_j u_j, \quad j = 1, \ldots N-1, \\
\frac{\partial u_N}{\partial t} =& D_I \nabla^2 u_N + k_{N-1} u_{N-1} b + r_{N-1} u_{N-1} - q_N u_N.
\end{aligned}$$

looks like a method of lines discretisation of another PDE in state space.

In the regime $N \gg 1$, we identify the discrete variable j with a continuous version $s \sim j/N$.
We also define

$$u(s, \mathbf{x}, t) \sim u_j(\mathbf{x}, t), \quad k(s) \sim k_j, \quad q(s) \sim q_j, \quad r(s) \sim r_j. \tag{6}$$

Then the semidiscrete system can be reformulated as

$$\begin{aligned}\frac{\partial u}{\partial t} =& D_I \nabla^2 u + \left[k(s-\Delta s)b + r(s-\Delta s)\right] u(s-\Delta s) \\ & - \left[k(s)b + r(s) + q(s)\right] u(s) \\ & + q(s+\Delta s)u(s+\Delta s), \quad s \in (\Delta s, 1-\Delta s),\end{aligned}$$

where $\Delta s = 1/N$.

Taylor expanding in the regime $N \gg 1$, or equivalently $0 < \Delta s \ll 1$, we obtain

$$\frac{\partial u}{\partial t} = D_I \nabla^2 u + \frac{\partial^2}{\partial s^2}\left(\frac{f(s,b)}{2N^2} u\right) - \frac{\partial}{\partial s}\left(\frac{g(s,b)}{N} u\right),$$

for $(s, \mathbf{x}) \in (0,1) \times (0,1)^3$, where

$$f(s,b) = r(s) + q(s) + k(s)b, \quad g(s,b) = -q(s) + r(s) + k(s)b.$$

In addition to no-flux boundary conditions in physical space, we also impose no-flux boundary conditions in state space,

$$\frac{\partial}{\partial s}\left(\frac{f(s,b)}{2N} u\right) - g(s,b)u = 0, \quad s = 0, 1.$$

For the b equation, the sum is a discrete approximation of an integral:

$$\frac{\partial b}{\partial t} = D_b \nabla^2 b - b \left(N \int_0^1 k(s) u(s, \mathbf{x}, t) \, \mathrm{d}s - \frac{1}{2} k(s) u(s, \mathbf{x}, t) \Big|_{s=0}^{s=1} \right),$$

for $\mathbf{x} \in (0, 1)^3$.

- the boundary term arises from retainig higher order corrections; it is needed to avoid non-physical depletion of bacteria

An Illustrative Example

Consider $q = r = 0$ and $k = \mathrm{const.}$
The spatially uniform steady-states satisfy:

$$
\begin{aligned}
0 &= \frac{\partial^2}{\partial s^2}\left(\frac{kb}{2N}u\right) - \frac{\partial}{\partial s}\left(kbu\right), \\
0 &= \frac{\partial}{\partial s}\left(\frac{kb}{2N}u\right) - kbu, \quad s = 0, 1 \quad \text{(BC)}, \\
0 &= b\left(N\int_0^1 k(s)u(s,\mathbf{x},t)\,\mathrm{d}s - \frac{1}{2}k(s)u(s,\mathbf{x},t)\Big|_{s=0}^{s=1}\right).
\end{aligned}
$$

Two possibilities:

- $b = 0$, $u(s)$ arbitrary (all bacteria eliminated)
- $b = \mathrm{const.} > 0$, $u(s) \propto e^{2N(s-1)}$ (saturated immune cells)

Future Work

- comprehensive comparisons between continuum model and ABM
- can other ABM rules be coarse-grained
- $u(s, \mathbf{x}, t)$ and parameters $k(s)$ can be localised in state space, leading to boundary layers in the continuum system (e.g. all healthy cells) $\Rightarrow$ hybrid discrete/continuum model
- can we make ABM parameter searches/sweeps faster with coarse-grained models?

References I

J. Segovia-Juarez, S. Ganguli, and D. Kirschner.
Identifying control mechanisms of granuloma formation during *M. tuberculosis* infection using an agent-based model.
Journal of Theoretical Biology, 231(3):357–376, 2004.

M. Doi.
Stochastic theory of diffusion-controlled reaction
Journal of Physics A, 9(9):1479–1495, 1976.

R. Erban, and S. J. Chapman.
Stochastic modelling of reaction–diffusion processes: algorithms for bimolecular reactions
Physical Biology 6(4):046001, 2009.

E. Uribe-Querol, and C. Rosales.
Phagocytosis: our current understanding of a universal biological process
Frontiers in Immunology, 11:1066, 2020.