Mean field limits for interacting particle systems, their inference, and applications

Part 3 – Inference for more complicated interactions and some case studies, including control

Susana N. Gomes
University of Warwick
Susana.Gomes@warwick.ac.uk

UNIVERSITY OF WARWICK

From microscopic dynamics to continuum limits - ACDC Autumn School

23/10/2025

Outline

Motivation

Earlier today, we saw how mean-field limits helped us characterise the long-time behaviour of an interacting particle system. Now we will explore how to use this for inference.

Recall that we are considering the system

$$dX_t^i = V'\left(X_t^i\right)dt + \frac{1}{N}\sum_{j\neq i} K\left(X_t^i - X_t^j\right)dt + \sqrt{2\sigma}dW_t^i\,, \qquad X_0^i = x_0^i, \qquad i,j = 1,\dots,N.$$

Because of the interacting term, if we want to evaluate a MLE for a trajectory, we need to know **all the other trajectories**!

If we consider the mean-field limit, we can avoid that by instead solving a PDE for the density.

I will show you how to do this for two applications, and finish with some recent related work.

Case study 1

Pedestrian dynamics

SNG, A.M. Stuart, M.-T. Wolfram, SIAM Journal of Applied Mathematics 79(4), 1475-1500 (2019)

UNIVERSITY OF WARWICK

Crowd Dynamics

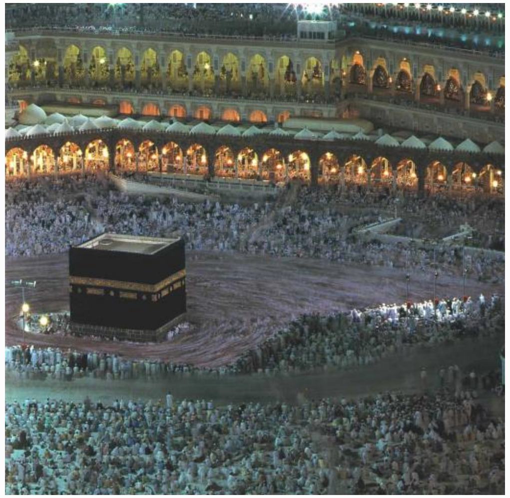
Understanding the individual dynamics as well as the evolution of a crowd is crucial in safety and transportation management.

Relevant questions

- How long would it take to evacuate a room in an emergency?
- How many people can one safely allow in a room/museum/stadium?
- This relies on correct modelling of a crowd, including knowing all the relevant parameters.

Challenges:

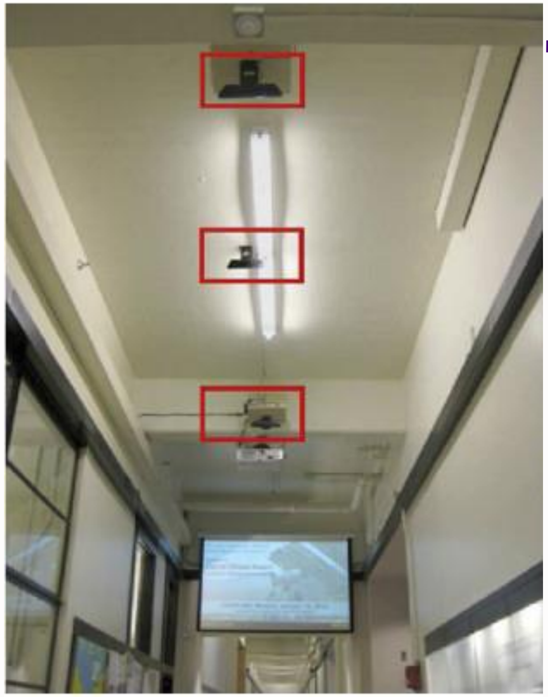
- Mathematical modelling: multi-scale nature, microscopic interactions not clearly defined.
- Analysis: highly nonlinear PDEs or complicated SDEs.
- Simulations: high computational complexity.



Individual trajectory data

Individual trajectories can be obtained, for example,

From overhead cameras [1]



From controlled experiments [2]

1 –Seer et al., Transportation Research Procedia, 2014.

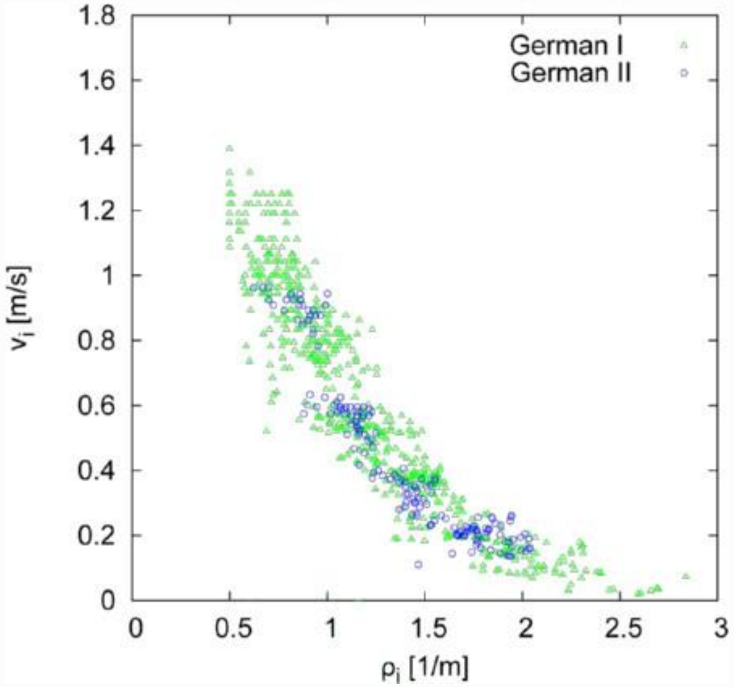
2 - Courtesy of Armin Seyfried (Forschungszentrum Jülich), BaSiGo experiments (5 days, 31 experiments, 200 runs, 28 industrial cameras, 2200 participants in total)

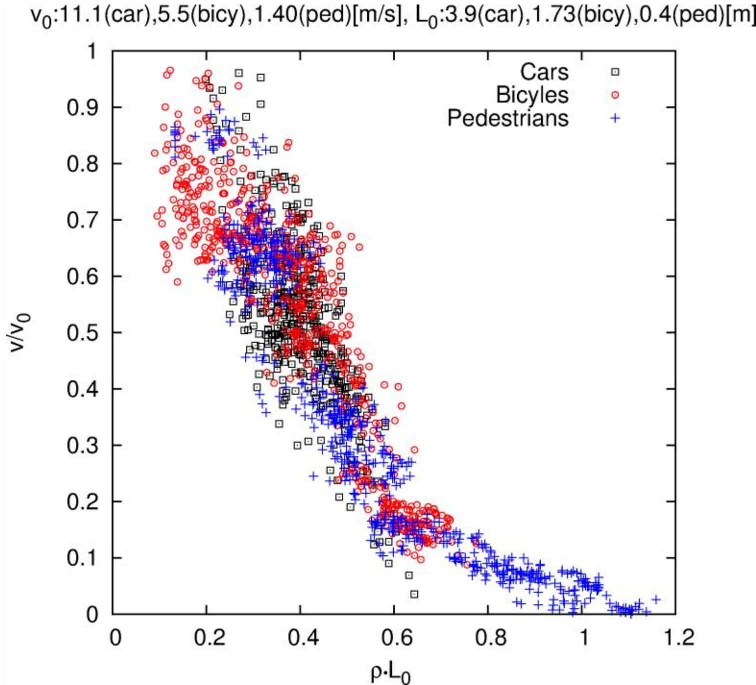
The fundamental diagram

The most widely used tool to characterise pedestrian dynamics is the fundamental diagram. [3]

- Relates the experimentally observed density of pedestrians to their velocity or outflow.
- There is a general agreement on its basic shape...
- ... but its parameterization depends on measurement and averaging techniques, and experimental setup.

Top figure: Students vs soldiers
Bottom figure: Cars vs pedestrians vs bicycles.





3 –Figures from Zhang and Seyfried, Procedia Engineering, 2013.

Existing models for pedestrian dynamics

Microscopic models (individual trajectories)

Agent-based models based on Newton's laws.

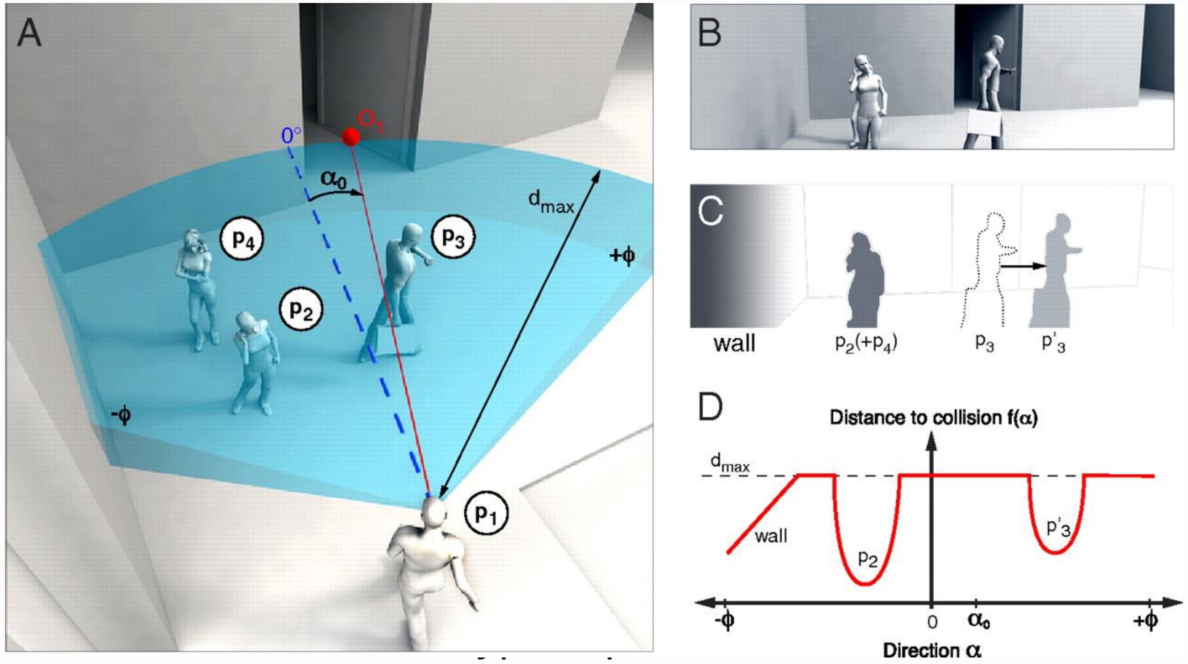
$$dX_t^i = V_t^i \, dt,$$
$$dV_t^i = F_i\left(X_t^i, \ldots, X_t^N, V_t^1, \ldots, V_t^N\right)dt + G\left(X_t^i\right) + \sigma_i dB_t^i$$

X_t^i, V_t^i are the position and speed of the i^{th} person, F, G are forces, and B_t^i stands for white noise.

Examples include

- Social force models [4]
- Cellular automata models [5]

Macroscopic models (pedestrian density)



- Hughes model.

4 – See, e.g., Helbing and Molnar, Physical Review E, 1995, Moussaïd, Helbing and Theraulaz, PNAS, 2011

5 – See, e.g., Burstedde et al, Physica A, 2001

Our model

Each trajectory solves a generalised **McKean–Vlasov equation**[6]

$$d\boldsymbol{X}_t^i = F\left(\rho(\boldsymbol{X}_t^i, t)\right)[1,0]^T\, dt + \sqrt{2\Sigma}\, d\boldsymbol{W}_t^i.$$

Here, $\boldsymbol{W}_t^i$ is a (2-dimensional) Brownian motion with diffusion $\Sigma = \mathrm{diag}(\sigma_1^2, \sigma_2^2)$, and

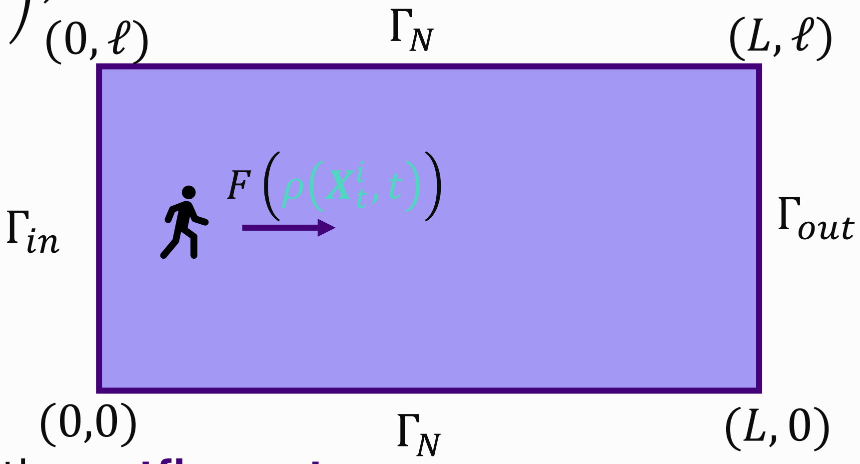
$$F\left(\rho(\boldsymbol{X}_t^i, t)\right) = v_{\max}\left(1 - \frac{\rho(\boldsymbol{x}, t)}{\rho_{max}}\right)[1,0]^T.$$

Note that: this **depends on the density of the process** $\rho(\boldsymbol{x}, t)$. The corresponding Fokker-Planck equation is

$$\rho_t = \nabla \cdot \left(\Sigma\nabla\rho - v_{max}\left(1 - \frac{\rho}{\rho_{max}}\right)\rho[1,0]^T\right),$$

coupled with the boundary conditions:

$$\begin{aligned}
(-\Sigma\nabla\rho + F(\rho)\rho[1,0]^T)\cdot\boldsymbol{n} &= -a(\rho_{max} - \rho), && \text{for } \boldsymbol{x} \in \Gamma_{in},\\
(-\Sigma\nabla\rho + F(\rho)\rho[1,0]^T)\cdot\boldsymbol{n} &= \quad b\rho, && \text{for } \boldsymbol{x} \in \Gamma_{out},\\
(-\Sigma\nabla\rho + F(\rho)\rho[1,0]^T)\cdot\boldsymbol{n} &= \quad 0, && \text{for } \boldsymbol{x} \in \Gamma_{N},
\end{aligned}$$

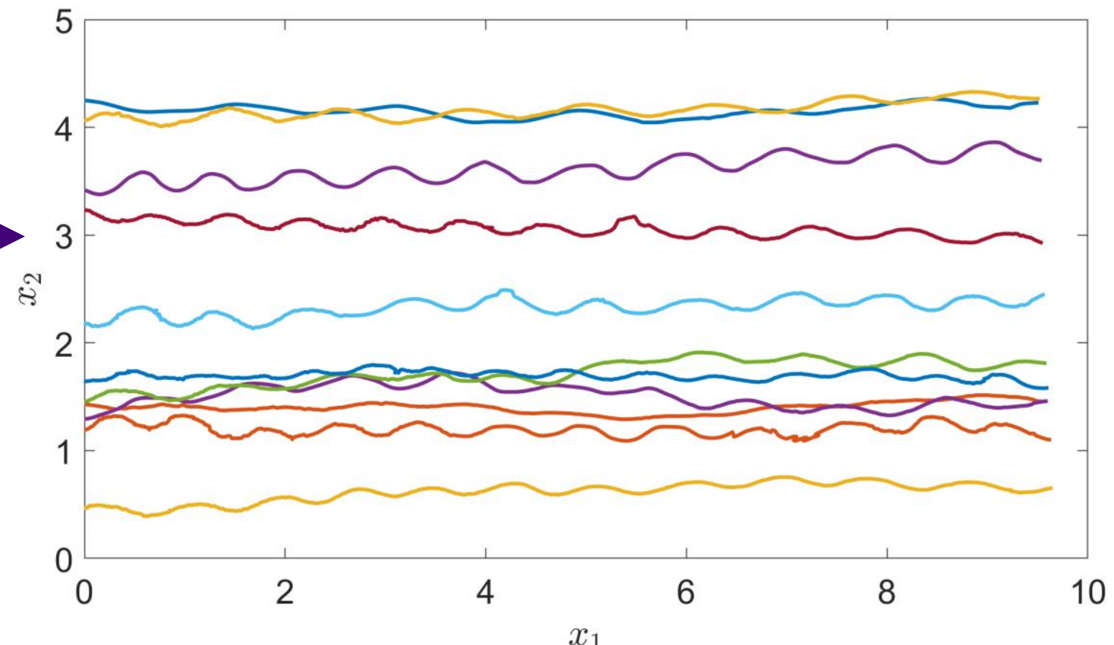


Where $\boldsymbol{n}$ is the outward pointing unit normal, a is the **inflow rate**, and b is the **outflow rate**.

6 – We consider the mean-field limit, but can write this as IPS using $K(x) = v_{\max}\left(1 - \frac{\delta(x)}{\rho_{max}}\right)$.

The inverse problem:

Given a set of J trajectories, $\left\{X_t^j\right\}_{j=1}^J$, either generated from the model or collected from controlled experiments, estimate v_{max} (and ρ_{max}).



Parameter estimation: usual inverse problem setting

Recall our goal: estimate v_{max} given a set of trajectories $\left\{\boldsymbol{X}_t^j\right\}_{j=1}^J$.

Recall that $\boldsymbol{X}_t^j$ solves an SDE, and therefore we **cannot solve the inverse problem in the usual way**, i.e. by minimising

$$\phi(v;\boldsymbol{X_t}) = \int_0^T \left|\dot{\boldsymbol{X}}_t - F\big(\rho(\boldsymbol{X}_t;v)\big)\right|_\Sigma^2 dt,$$

over all values of v, because it would be infinite almost surely.
Instead, we use the results from this morning to show that it is equivalent to minimise the *log-likelihood function* of the process:

$$\Psi(v;X_t) = \log L(\boldsymbol{X}_t;v) = \frac{1}{4}\int_0^T \left(\left|F\big(\rho(\boldsymbol{X}_t;v)\big)\right|_\Sigma^2 dt \; - 2\underbrace{\left\langle F\big(\rho(\boldsymbol{X}_t;v)\big), d\boldsymbol{X}_t\right\rangle}_{\text{Stochastic integral}}{}_\Sigma\right).$$

Since $\frac{1}{4}\int_0^T \left|\dot{\boldsymbol{X}}\right|^2 dt$ does not depend on v, minimising ψ is equivalent to minimising ϕ!

Possible issue: No idea if is convex, differentiable, …

Including prior information and the Bayesian approach

Instead of minimising ψ, we can include *prior information* and consider instead the functional

$$\mathcal{J}(v;\boldsymbol{X}_t) = \psi(v;\boldsymbol{X}_t) + \frac{1}{2}|v-m|_c^2.$$

Using Bayes' and Girsanov's theorems, we can show that the function

$$e^{-\mathcal{J}(v;\boldsymbol{X}_t)}\chi(v>0),$$

appropriately normalised, is the **posterior distribution** $\mathbb{P}(v|\boldsymbol{X})$ of v given the observation $\boldsymbol{X}$.

In summary...

- Minimising ψ over all possible v gives the **maximum likelihood estimator** (MLE) for v, which is also the mean of $\mathbb{P}(v|\boldsymbol{X})$.
- Minimising $\mathcal{J}$ gives the **maximum a posteriori estimator** (MAP) for v. This is also the mode of $\mathbb{P}(v|\boldsymbol{X})$.
- The full Bayesian approach involves sampling from $\mathbb{P}(v|\boldsymbol{X})$ and provides the whole posterior distribution.

Parameter estimation approaches

We want to use $J = 20$ trajectories $\left\{X_t^j\right\}_{t\in[0,T]}^{j=1,\dots,20}$ to compute our estimate.

Optimisation approach: the Nelder-Mead algorithm [7]

- Few function evaluations
- Good improvement in objective function in few iterations
- Result can be used to identify good initial guess for MCMC
- Quick convergence (when it does converge)
- Convergence only guaranteed for 1d **and** strictly convex functions

Bayesian approach: MCMC with pCN algorithm [8]

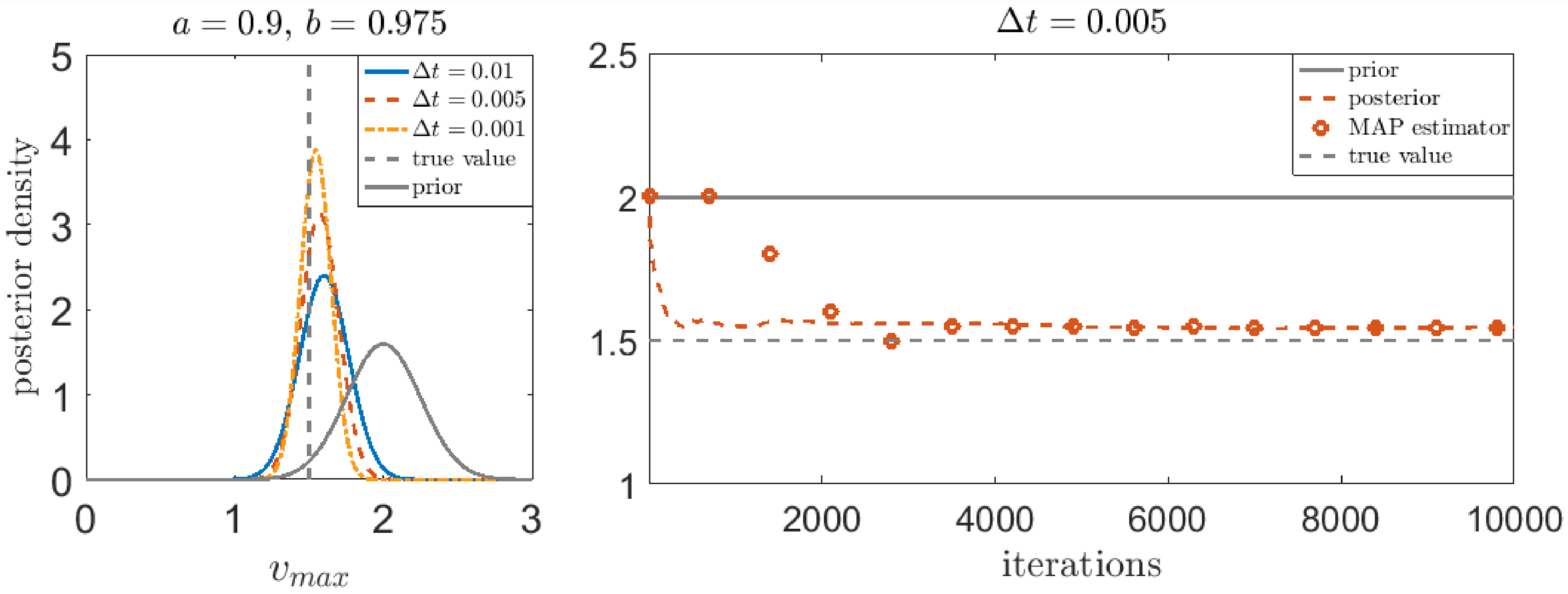
- Provides posterior distribution
- Allows for non-parametric (functional) estimation
- Has a single tuneable parameter - can be used to maximize efficiency
- Allows for uncertainty quantification
- Simple interpretation
- Requires many iterations (= PDE solutions!)

7 – Nelder and Mead, The Computer Journal, 1965.

8 - Cotter et al., Statistical Science, 2013.

Benchmarking: Bayesian approach vs optimisation

(maximal current) $v_{max} = 1.5, a = 0.9, b = 0.975, \sigma_1 = \sigma_2 = 0.05.$



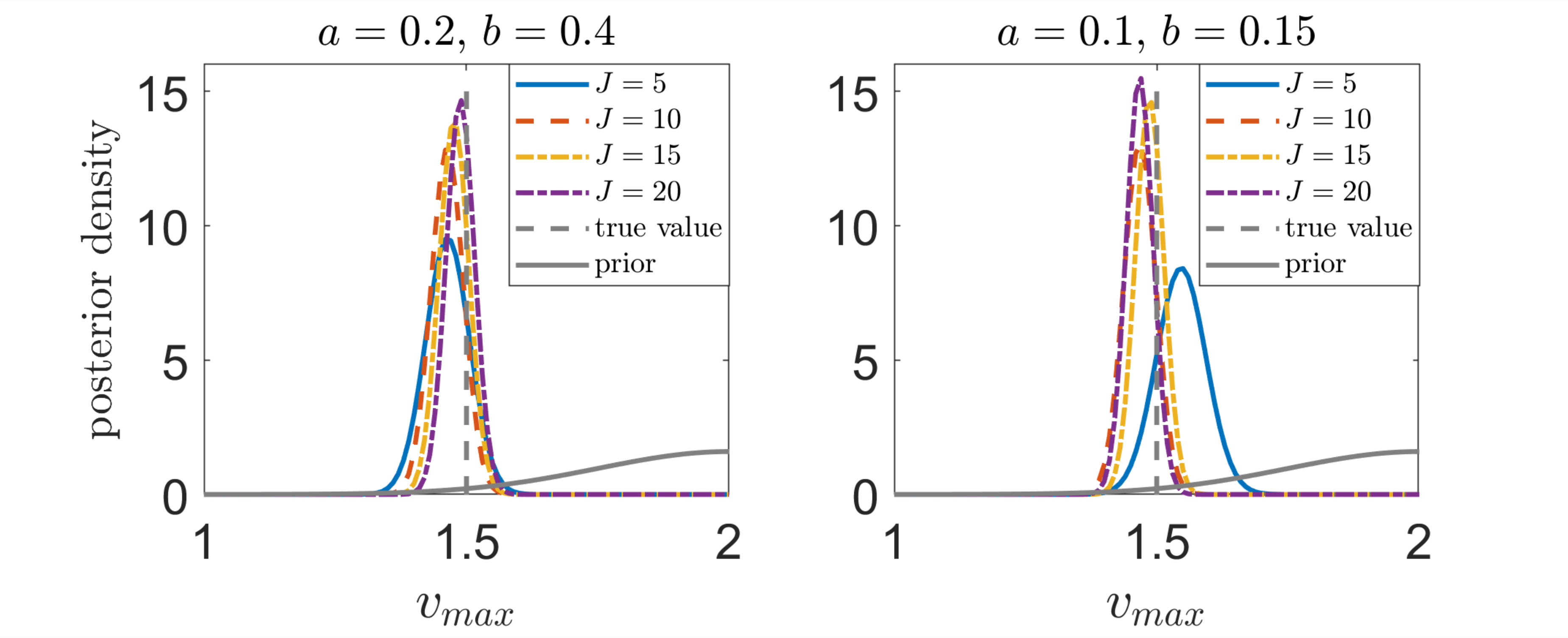
Left: Influence of the PDE time step on the posterior distribution. **Right:** Comparison between Bayesian and optimization approach for t = 0:005.

MAP estimator obtained by derivative-free optimisation coincides with the posterior mean for all test cases – suggests posterior distribution is Gaussian.

Estimates are independent of all model independent parameters (prior mean and variance, inital guess, parameters of MCMC methodology)

Dependence on amount of information

Influence of number of trajectories for two **influx limited** regimes (i.e. inflow < outflow).

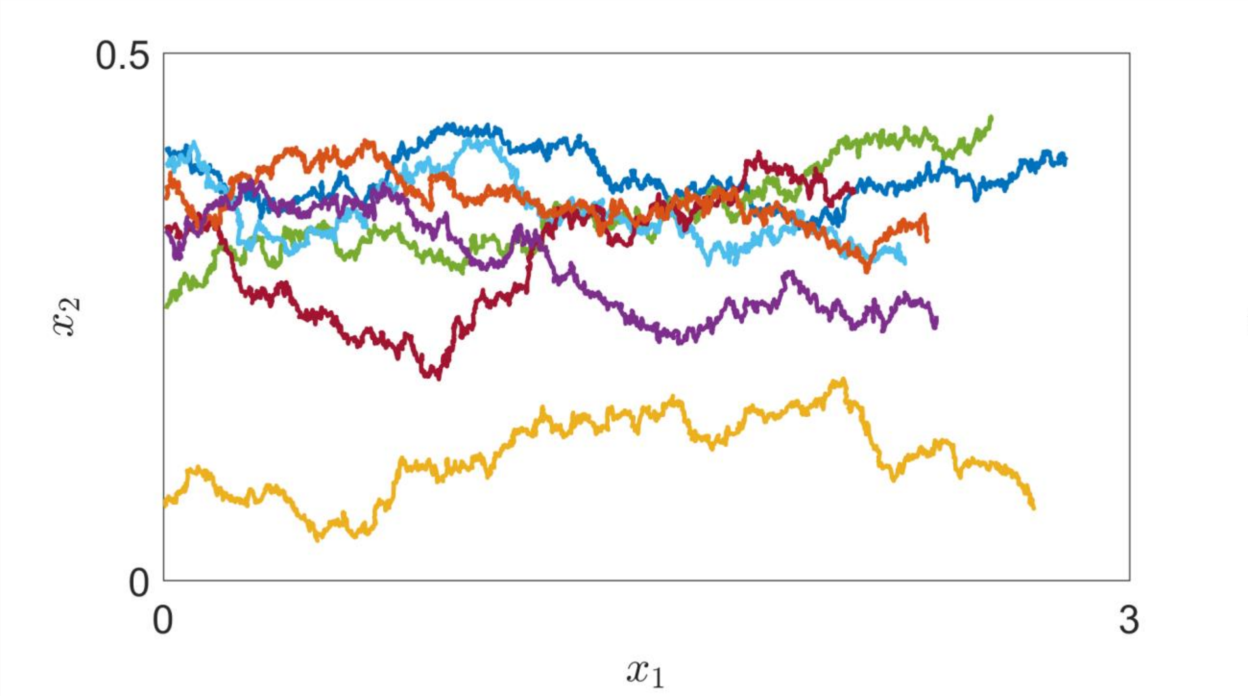


A small (5) number of trajectories gives biased estimates, but for more than 10 trajectories the posterior distributions concentrate around the true value.

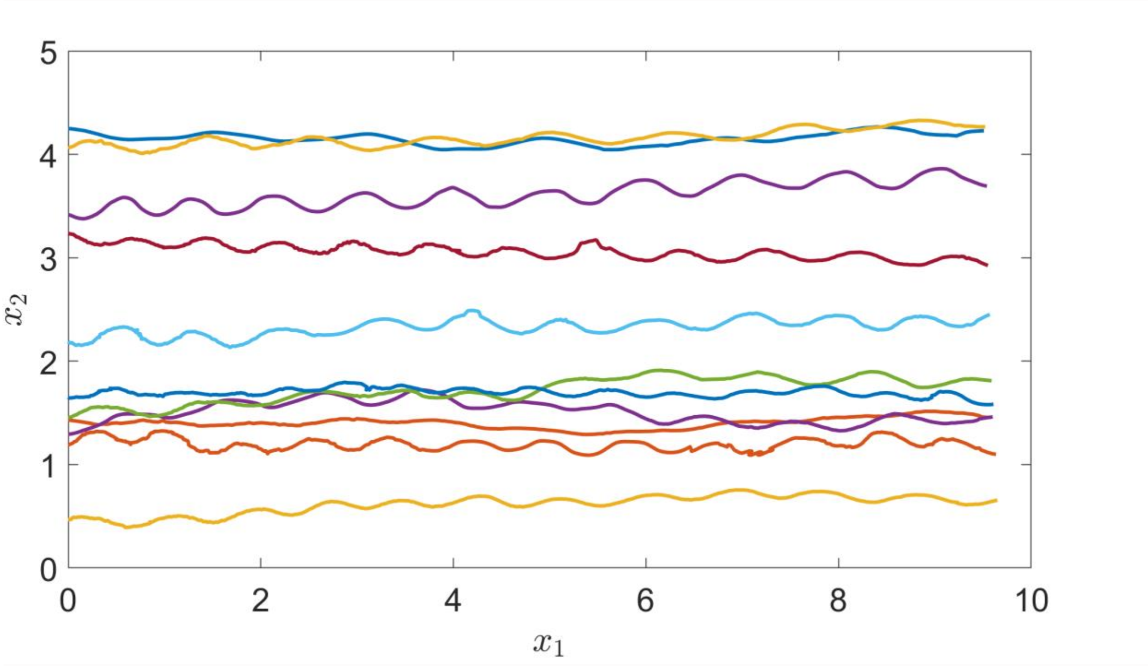
What about real data?

We tested the methodology using data from the BaSiGo experiments (10x5 corridor).[9] We used experiment 5, corresponding to a maximal current regime ($\min(a,b) \geq \frac{v_{max}}{2}$).

While it is not clear that the model will work…
For example, the model generated trajectories do not look realistic:



Left: Model.



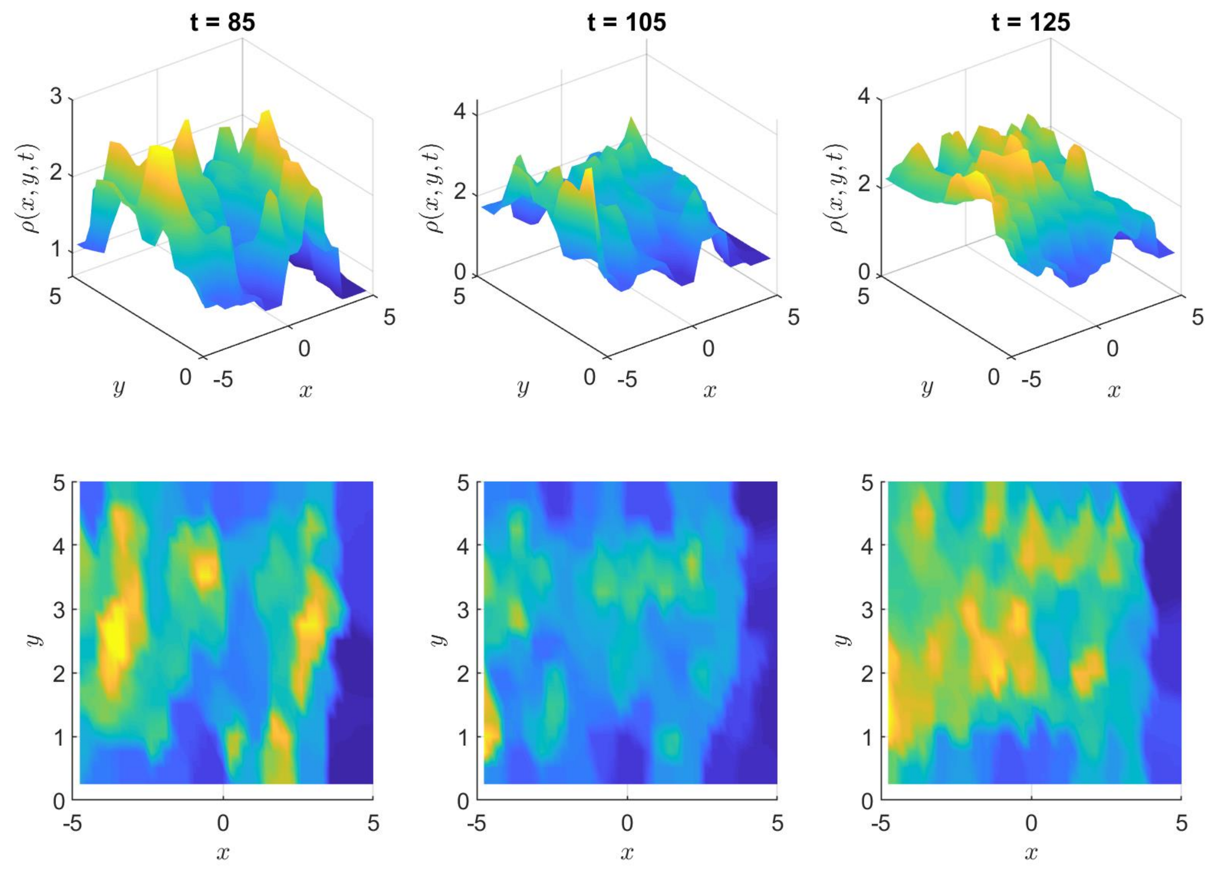
Right: Data.

9 - https://ped.fz-juelich.de/db/doku.php?id=corridor5

What about real data?

We tested the methodology using data from the BaSiGo experiments (10x5 corridor).[9] We used experiment 5, corresponding to a maximal current regime ($\min(a, b) \geq \frac{v_{max}}{2}$).

While it is not clear that the model will work…
…and the observed density does not appear constant:

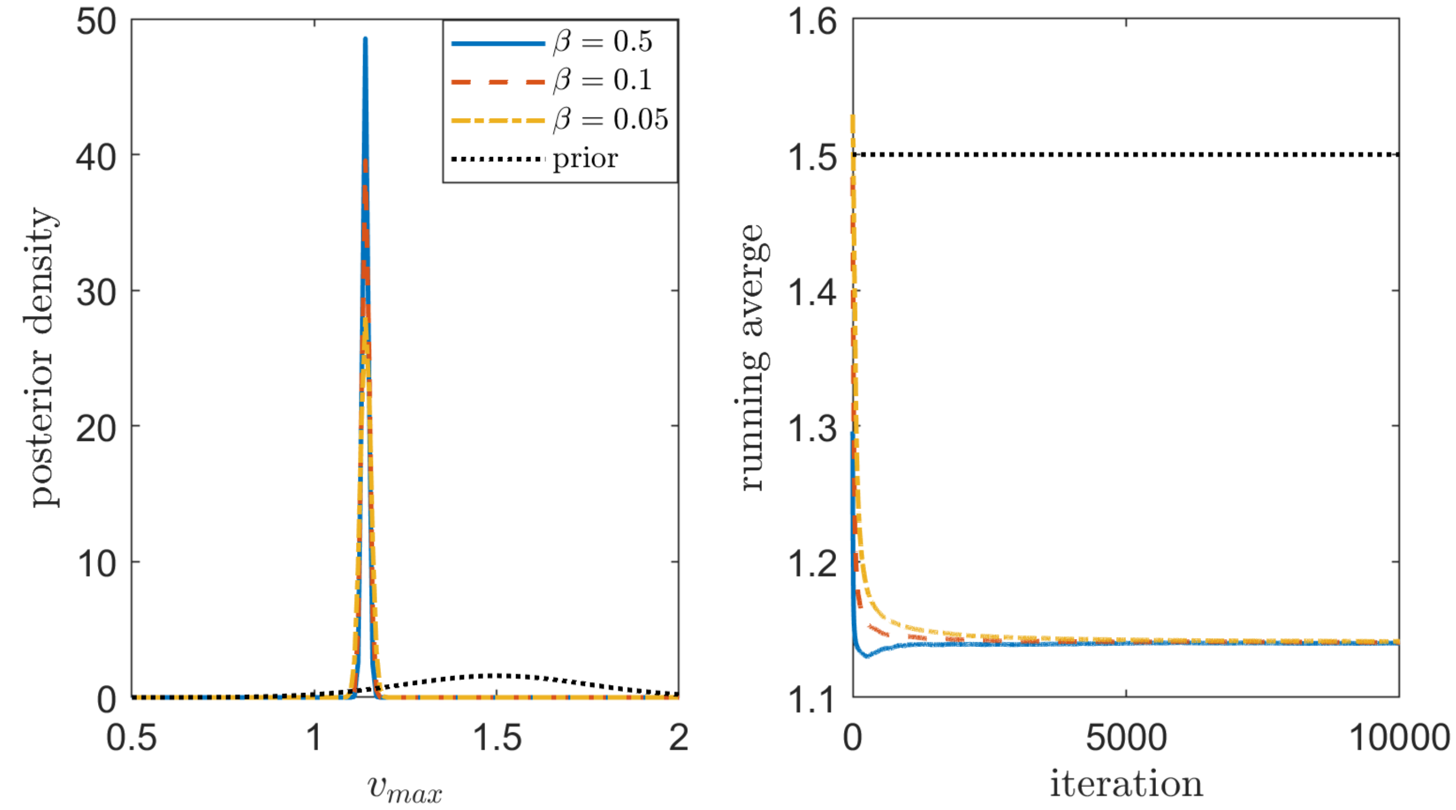


9 - https://ped.fz-juelich.de/d

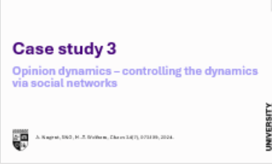
What about real data?

We tested the methodology using data from the BaSiGo experiments (10x5 corridor).[9] We used experiment 5, corresponding to a maximal current regime ($\min(a,b) \geq \frac{v_{max}}{2}$).

While it is not clear that the model will work…
Preliminary results show consistent estimates!



9 - https://ped.fz-juelich.de/db/doku.php?id=corridor5



Case study 2

Cell population dynamics

Work in preparation, joint with J.A. Carrillo (Oxford) and G. Estrada (UPC Barcelona).

UNIVERSITY OF WARWICK

Motivation

Understanding how cell populations interact has several practical applications such as tissue and organ formation.

One of the fundamental biological phenomena which explains how cells interact with each other is the mechanism of **cell-cell adhesion**.

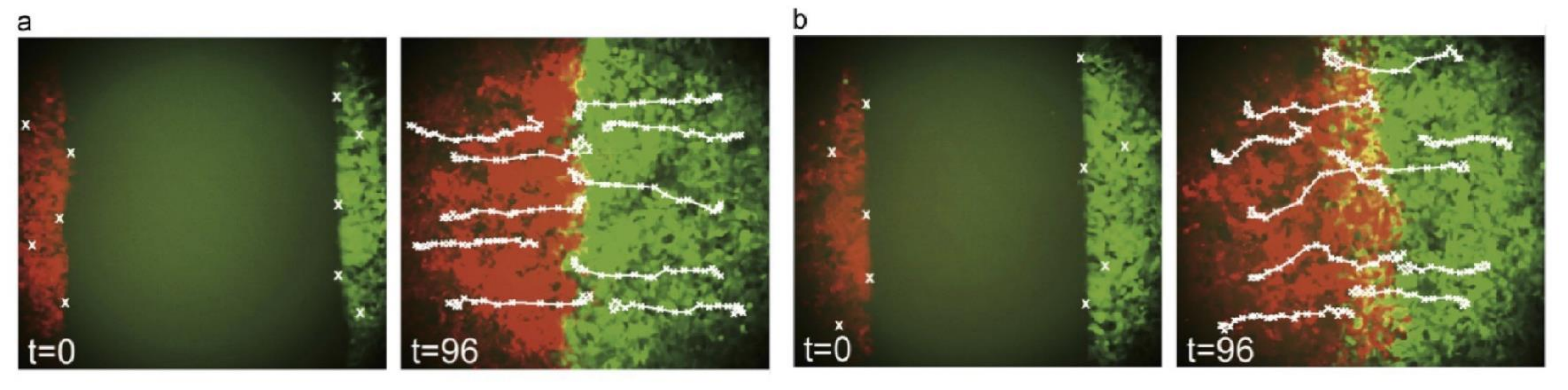
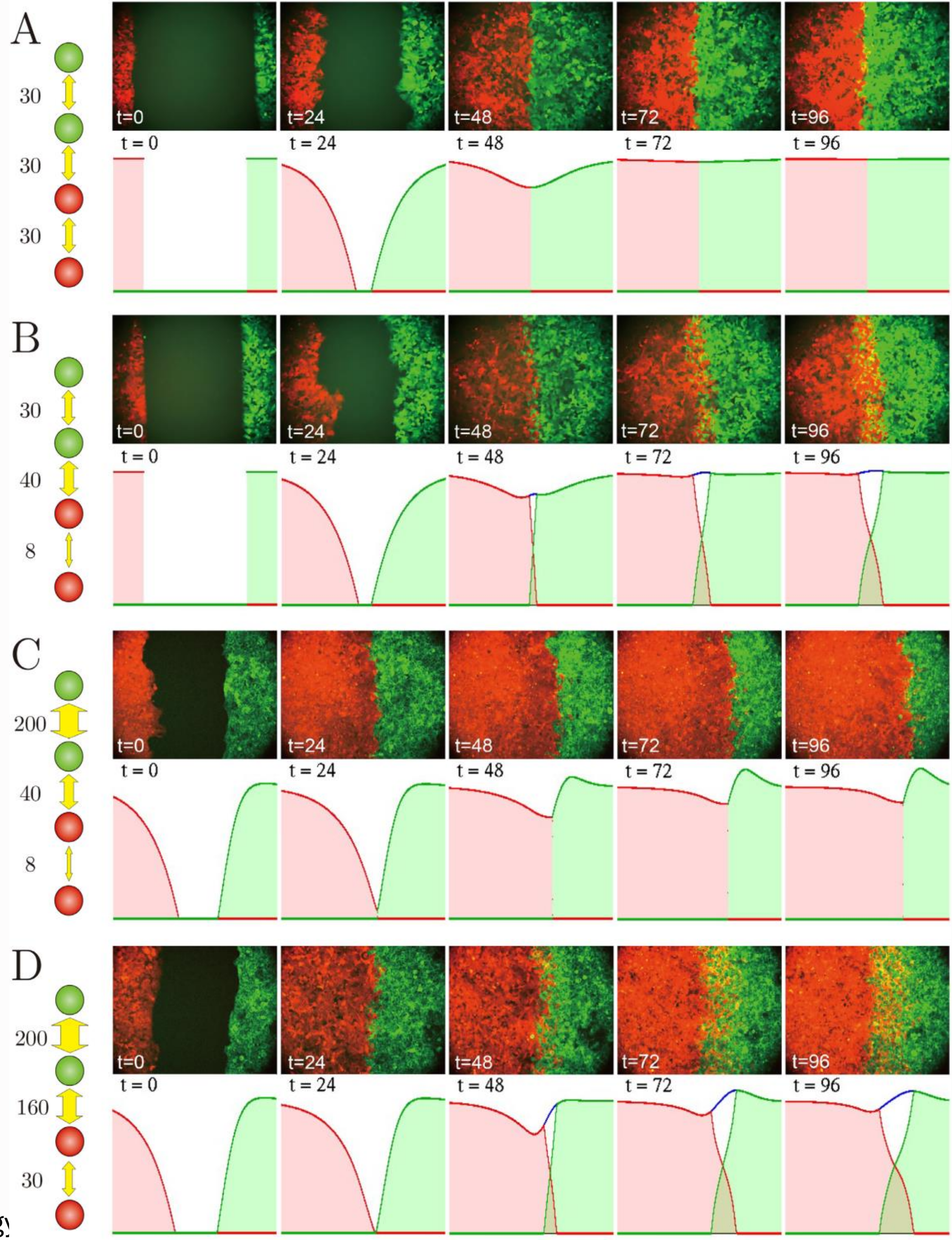


Figure: Cell-cell adhesion experiments, taken from Murakawa and Togashi, JTB, 2015.

Motivation

Depending on cell properties, when two populations meet, they can **form a barrier** (case A) or **mix at the boundary**, exhibiting different behaviours of invading each other (“bidirectional invasion” in B, “pushing back” in C, or “asymmetric invation” in D).

There are several experimental works showing this behaviour, and this has recently attracted mathematical interest.[10]



10 - see, e.g., J.A. Carrillo, H. Murakawa, M. Sato, H. Togashi, O. Trush, Journal of Theoretical Biology

The microscopic model (trajectories)

We consider two populations of cells $\boldsymbol{x}$ and $\boldsymbol{y}$, where each cell interacts with other cells both in its own population and outside it, according to certain rules:

$$d\boldsymbol{x}_i = -\frac{1}{N_{\boldsymbol{x}}}\sum_{j\neq i}^{N_{\boldsymbol{x}}} \nabla\mathcal{F}_{11}(\boldsymbol{x}_i - \boldsymbol{x}_j)\,dt - \frac{1}{N_{\boldsymbol{y}}}\sum_{j\neq i}^{N_{\boldsymbol{y}}} \nabla\mathcal{F}_{12}(\boldsymbol{x}_i - \boldsymbol{y}_j)\,dt + \sqrt{2\Sigma_{\boldsymbol{x}}}\,dW_i^{\boldsymbol{x}}(t)$$

$$d\boldsymbol{y}_i = -\frac{1}{N_{\boldsymbol{y}}}\sum_{j\neq i}^{N_{\boldsymbol{y}}} \nabla\mathcal{F}_{22}(\boldsymbol{y}_i - \boldsymbol{y}_j)\,dt - \frac{1}{N_{\boldsymbol{x}}}\sum_{j\neq i}^{N_{\boldsymbol{x}}} \nabla\mathcal{F}_{21}(\boldsymbol{y}_i - \boldsymbol{x}_j)\,dt + \sqrt{2\Sigma_{\boldsymbol{y}}}\,dW_i^{\boldsymbol{y}}(t)$$

Here, the function $\mathcal{F}_{mn}$ encodes all interactions:

$$\nabla\mathcal{F}_{mn}(\boldsymbol{z}) = \nabla\mathcal{W}_{mn}(\boldsymbol{z}) + \nabla\mathcal{M}_{mn}(\boldsymbol{z}) = \nabla\left(\underbrace{a_{mn}\big((|\boldsymbol{z}| - \ell_{mn})^2}_{\text{Short range repulsion}} - \underbrace{(R - \ell_{mn})^2}_{\text{Long range attraction}}\big)\right) + \underbrace{\frac{b_{mn}}{4\pi\epsilon_{mn}^2}\nabla e^{-\frac{|z|^2}{4\epsilon_{mn}^2}}}_{\text{Local repulsion (volume exclusion)}}.$$

The macroscopic model (population densities)

The previous equations assume every cell **can see every other cell**. To overcome this, we can pass to the mean-field limit, and instead consider:

$$d\boldsymbol{X}_{\boldsymbol{i}} = -\nabla(\mathcal{F}_{11} \star \rho_1)(\boldsymbol{X}_{\boldsymbol{i}})\, dt - \nabla(\mathcal{F}_{12} \star \rho_2)(\boldsymbol{X}_{\boldsymbol{i}})\, dt + \sqrt{2\Sigma_{\boldsymbol{x}}}\, dW_i^{\boldsymbol{x}}(t),$$

$$d\boldsymbol{Y}_{\boldsymbol{i}} = -\nabla(\mathcal{F}_{22} \star \rho_2)(\boldsymbol{Y}_{\boldsymbol{i}})\, dt - \nabla(\mathcal{F}_{21} \star \rho_1)(\boldsymbol{Y}_{\boldsymbol{i}})\, dt + \sqrt{2\Sigma_{\boldsymbol{y}}}\, dW_i^{\boldsymbol{y}}(t),$$

where $\star$ denotes convolution, and ρ_1 and ρ_2 are the densities of the two cell populations.

We have that ρ_1 and ρ_2 solve the following **aggregation-diffusion** PDEs:

$$\partial_t \rho_1 = \nabla \cdot (\rho_1 \nabla(\mathrm{b}_1(\rho_1 + \rho_2) + \mathcal{W}_{11} \star \rho_1 + \mathcal{W}_{12} \star \rho_2)) + \Sigma \Delta \rho_1$$
$$\partial_t \rho_2 = \nabla \cdot (\rho_2 \nabla(\mathrm{b}_2(\rho_1 + \rho_2) + \mathcal{W}_{22} \star \rho_2 + \mathcal{W}_{21} \star \rho_1)) + \Sigma \Delta \rho_2$$

We want to estimate all parameters in the equations, i.e., $\theta = (b_1, b_2, a_{11}, a_{12}, a_{21}, a_{22}, \ell_{11}, \ell_{12}, \ell_{21}, \ell_{22})$ based on cell trajectories.

Inference strategy

To estimate these parameters, given a set of trajectories $\mathcal{X} = (x_1, \dots, x_{J_x})$ and $\mathcal{Y} = (y_1, \dots, y_{J_y})$,we would want to estimate the parameters in the potentials.

The log-likelihood function ψ now is

$$\Psi(\theta; \mathcal{X}, \mathcal{Y}) = \frac{1}{4}\int_0^T \sum_{i=1}^{J_x} |\nabla(\mathcal{F}_{11}\star \rho_1)(\boldsymbol{x_i}) + \nabla(\mathcal{F}_{12}\star \rho_2)(\boldsymbol{x_i})|^2 \, dt + 2 \langle \nabla(\mathcal{F}_{11}\star \rho_1)(\boldsymbol{x_i}) + \nabla(\mathcal{F}_{12}\star \rho_2)(\boldsymbol{x_i}), d\boldsymbol{x_i}(t)\rangle + \frac{1}{4}\int_0^T \sum_{i=1}^{J_y} |\nabla(\mathcal{F}_{22}\star \rho_2)(\boldsymbol{y_i}) + \nabla(\mathcal{F}_{21}\star \rho_1)(\boldsymbol{y_i})|^2 dt + 2 \langle \nabla(\mathcal{F}_{22}\star \rho_2)(\boldsymbol{y_i}) + \nabla(\mathcal{F}_{21}\star \rho_1)(\boldsymbol{y_i}), d\boldsymbol{y_i}(t)\rangle.$$

As before, Instead of minimising Ψ, we can include **prior information** and consider instead the functional

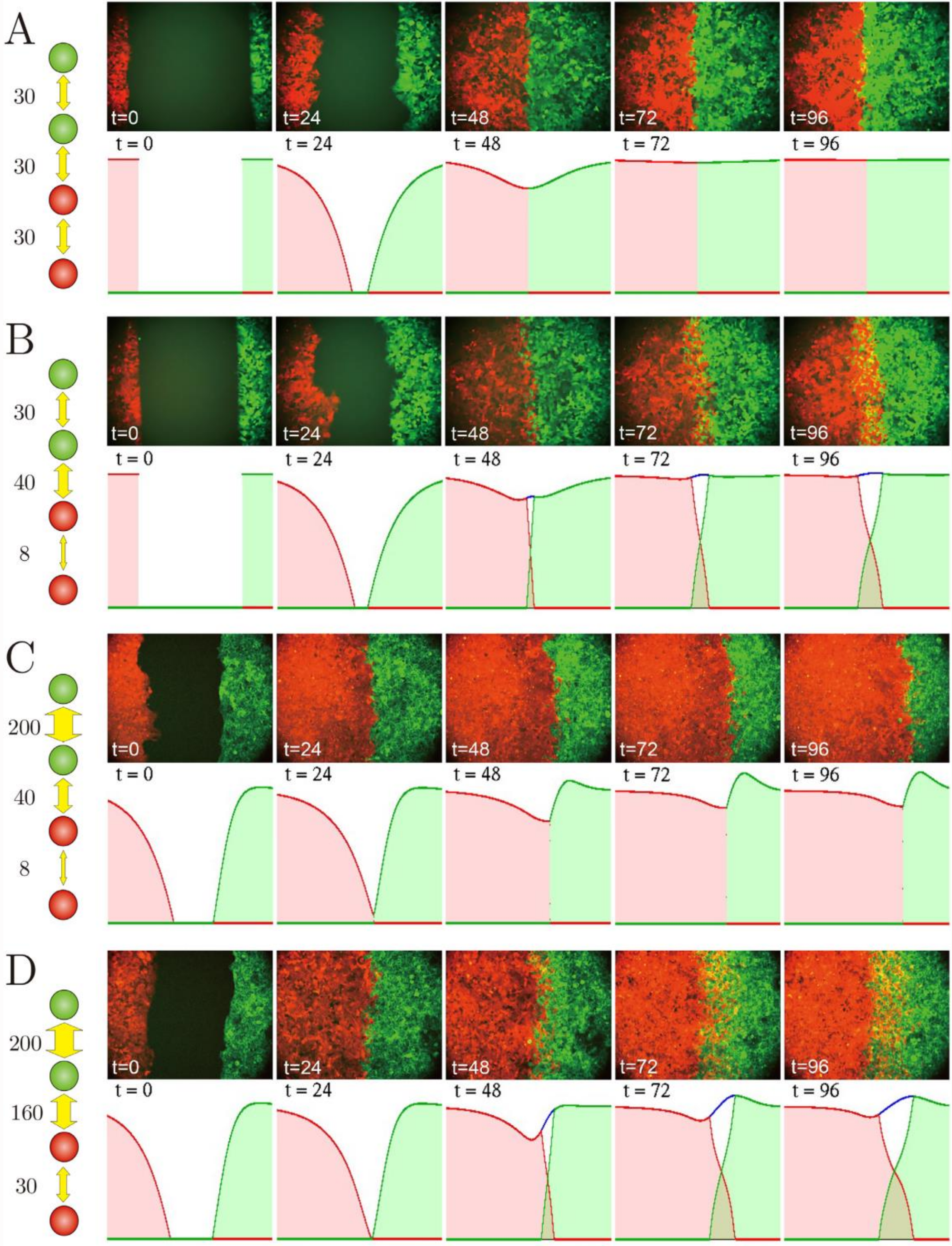
$$\mathcal{J}(\theta; \mathcal{X}, \mathcal{Y}) = \psi(\theta; \mathcal{X}, \mathcal{Y}) + \frac{1}{2}\langle \theta - m, C^{-1}(\theta - m)\rangle,$$

for a given m and C. This is equivalent to adding a Gaussian prior with mean m and covariance C.

Numerical setup

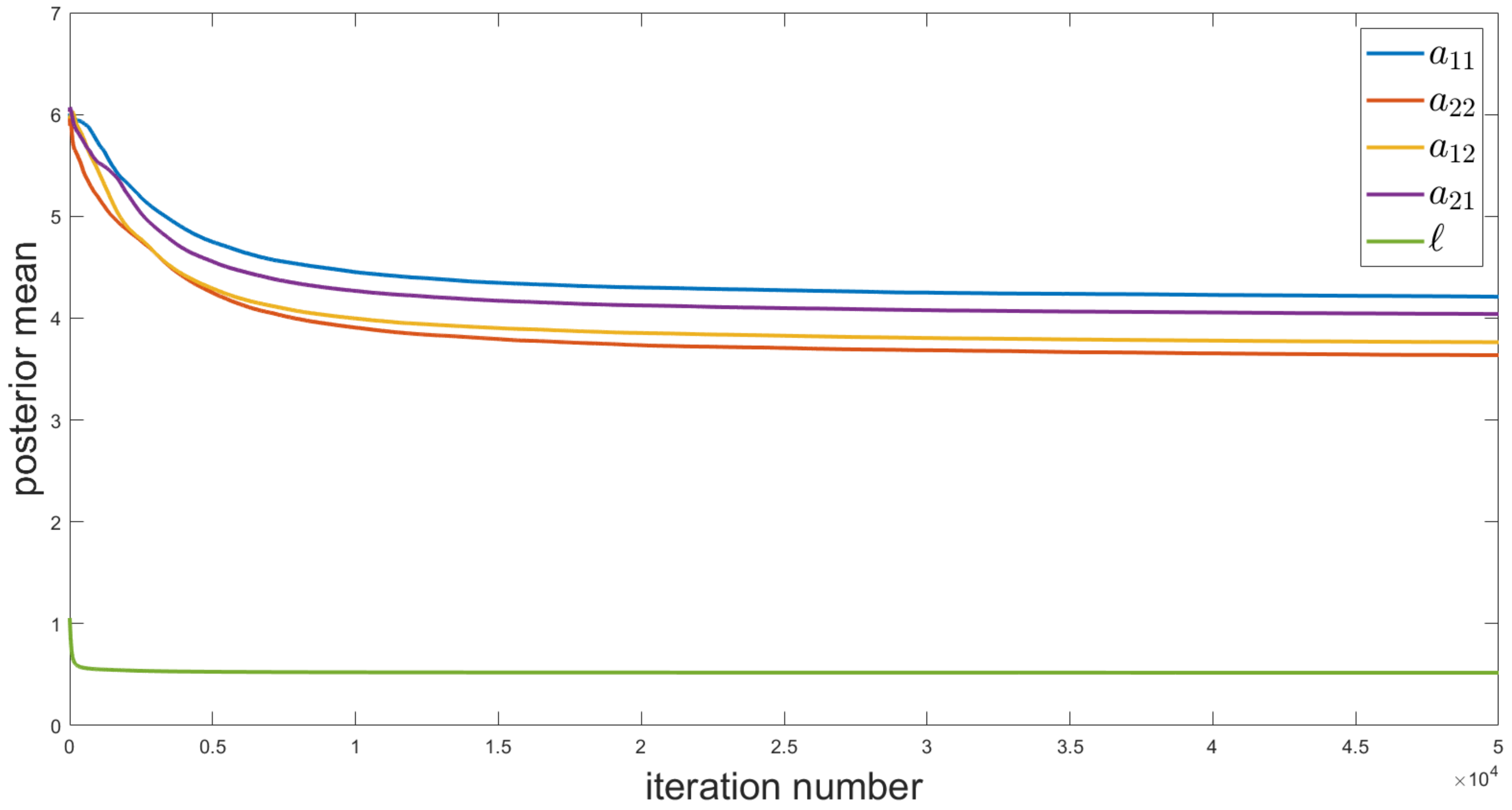
We consider four cases which are representative of the characteristic behaviours of the cell populations

	Case A	Case B	Case C	Case D
a_{11}	4	4	6	6
a_{12}	4	1	3	1
a_{21}	4	1	3	1
a_{22}	4	4	1	1
ℓ_1	0.5	0.5	0.5	0.5
ℓ_2	0.5	0.5	0.5	0.5
b_1	0.1	0.1	0.1	0.1
b_2	0.1	0.1	0.1	0.1
R	1	1	1	1



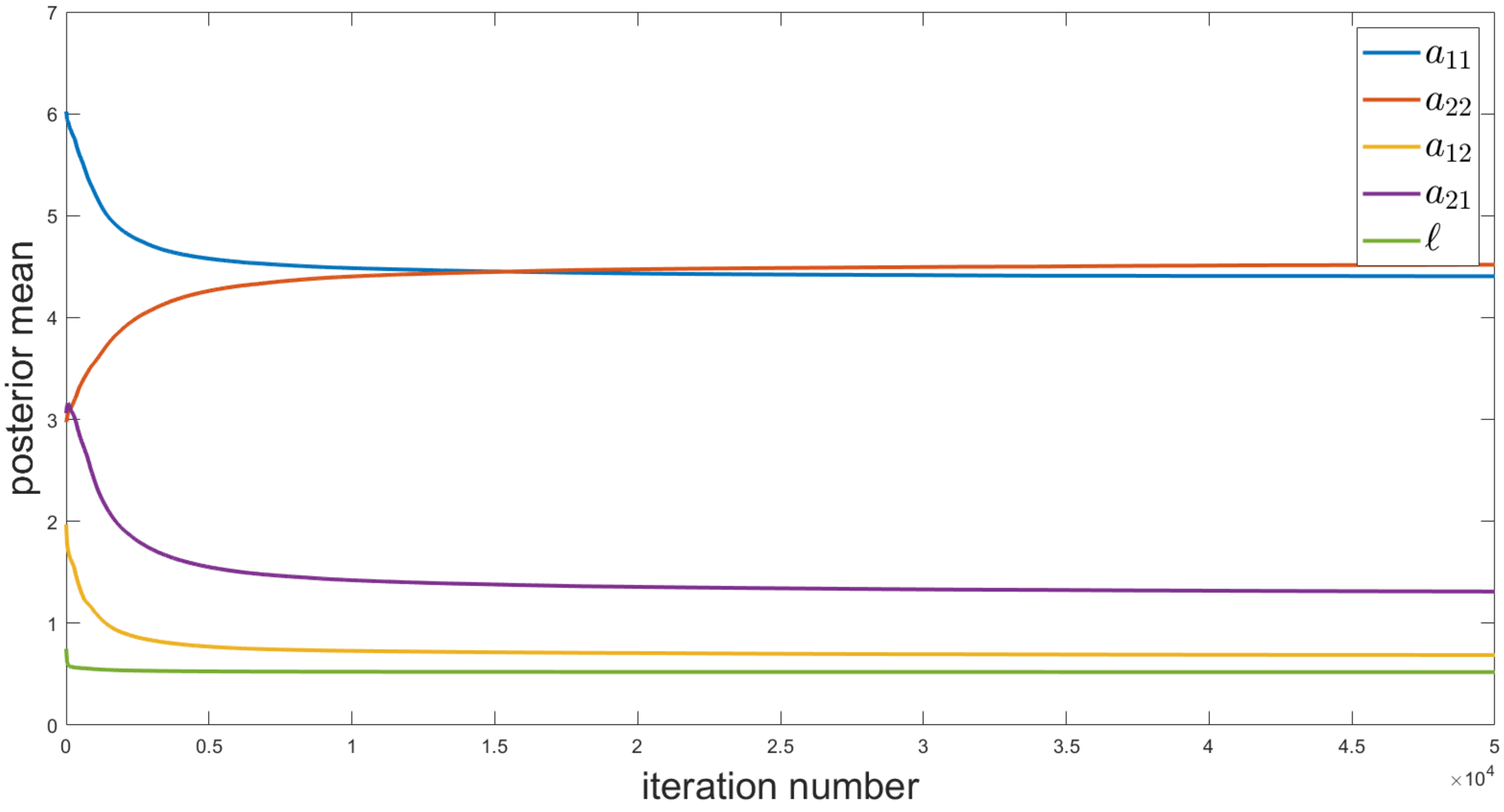
Numerical results

Representative results for case A ($a_{11} = a_{22} = a_{12} = a_{21} \ = 4, \ell = 0.5$).



Numerical results

Representative results for case B ($a_{11} = a_{22} = 4, a_{12} = a_{21} = 1, \ell = 0.5$).



a_{11} a_{22} a_{12} a_{21} ℓ

Recent related work

There is a lot of research activity surrounding these topics, both from the theoretical and the applications sides… Some examples (that I know of) include

- Inference specifically for McKean-Vlasov equations[11]
- Likelihood free inference, for when it is not possible to find likelihood functions. This is the basis of **Approximate Bayesian Computation (ABC)**[12]
- Sequential Monte Carlo techniques for when one has positions but not trajectories[13]
- Modelling and inference for pedestrians and traffic flow models[14]
- Work on the mean-field approximations for cell dynamics models[15]
- Several other interesting research avenues (happy to discuss later)

11 – see Sharrock et al, Stoch Proc Appl 2023, Pavliotis and Zanoni, SIAM J Appl Dyn Syst 2022, Amorino et al, Stoch Proc Appl, 2023.

12 – see Toni et al., Journal of the Royal Society Interface, 2009, or Lintusaari, et al., Systematic biology, 2017.

13 – see Cheng, Wen and Li, Roy Soc Open Sci, 2023.

14 – see Wurth et al, Adv Comp Mat 2022, Corbetta and Toschi, Annual Review of Condensed Matter Physics, 2023, Gödel et al, Safety Science 2022.

15 – see Morale, Capasso, Oelschläger, J Math Bio, 2005, Burger, Capasso, Morale, Nonlinear Analysis: Real World Applications, 2007.

Case study 3

Opinion dynamics – controlling the dynamics via social networks

A. Nugent, SNG, M.-T. Wolfram, Chaos 34(7), 073109, 2024.

UNIVERSITY OF WARWICK

Controlling opinion dynamics via the underlying network

We saw this morning that we can model opinion dynamics co-evolving with an underlying social network. We will see how we can use this to control the dynamics of the population.

$$\frac{dx_i}{dt} = \frac{1}{k_i} \sum_{j=1}^{N} w_{ij} \phi\left(\left|x_j - x_i\right|\right)(x_j - x_i)$$

Controlling opinion dynamics via the underlying network

We saw this morning that we can model opinion dynamics co-evolving with an underlying social network. We will see how we can use this to control the dynamics of the population.

Common approaches **add a control** to each agent, with the goal of promoting consensus.

$$\frac{dx_i}{dt} = \frac{1}{k_i}\sum_{j=1}^{N} w_{ij}\phi\left(\left|x_j - x_i\right|\right)\left(x_j - x_i\right) + \boldsymbol{u_i}$$

Controlling opinion dynamics via the underlying network

We saw this morning that we can model opinion dynamics co-evolving with an underlying social network. We will see how we can use this to control the dynamics of the population.

Common approaches **add a control** to each agent, with the goal of promoting consensus.

We propose a different approach, consisting of controlling the social network instead:

$$\frac{dx_i}{dt} = \frac{1}{k_i}\sum_{j=1}^{N} w_{ij}\phi(|x_j - x_i|)(x_j - x_i)$$

$$\frac{dw_{ij}}{dt} = f(\boldsymbol{u_{ij}}, w_{ij}).$$

Controlling opinion dynamics via the underlying network

The function $f(\boldsymbol{u_{ij}}, w_{ij})$ describes the effect of controls in the network. It has to be bounded and integrable, and needs to satisfy:

- $f(\mathbf{0}, w_{ij}) = 0$ for all $w_{ij} \in [0,1]$, so weights remain constant if uncontrolled,
- $f(\boldsymbol{u_{ij}}, 0) \geq 0$ for all $u_{ij} \in \mathbb{R}$, so weights remain nonnegative,
- $f(\boldsymbol{u_{ij}}, 1) \leq 0$ for all $u_{ij} \in \mathbb{R}$, so weights do not exceed 1.

We will use $f(\boldsymbol{u_{ij}}, w_{ij}) = s(\boldsymbol{u_{ij}})(\ell(\boldsymbol{u_{ij}}) - w_{ij})$ throughout this talk.

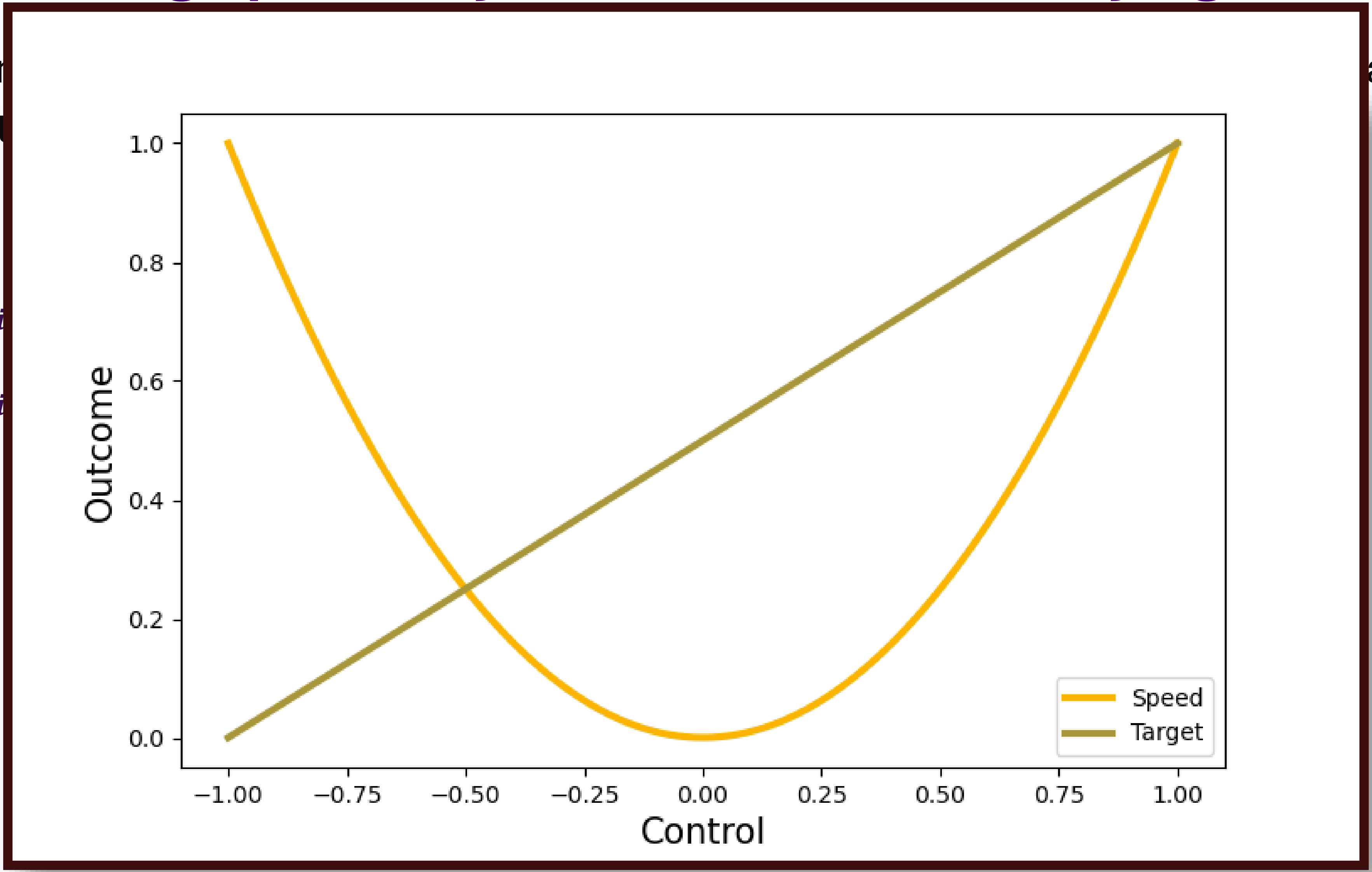
Goal: For given initial conditions, find a control $u = (u_{ij}) \in L^{\infty}(\mathbb{R}^{+}; \mathbb{R}^{N \times N})$ that brings the whole population to consensus at a pre-selected target opinion x^{*}.

Controlling opinion dynamics via the underlying network

The fu… as to be bounded and int…

- $f(0,$…
- $f(u_i$…
- $f(u_i$…

We will…



Controllability questions

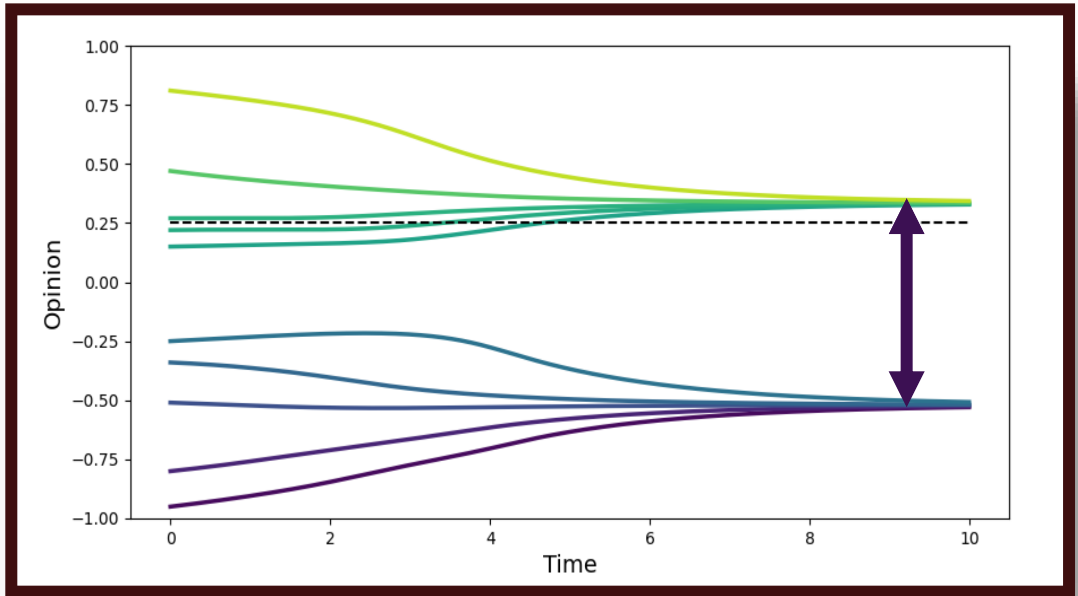
We first consider existence of controls for this problem. In particular, we ask ourselves:

1. For what combination of initial conditions and targets **does such a control exist**?
2. If it exists, **how do we find it**?
3. If we can find such a control, **how well does it work**?
4. If we can't find such a control, **how close can we get**?

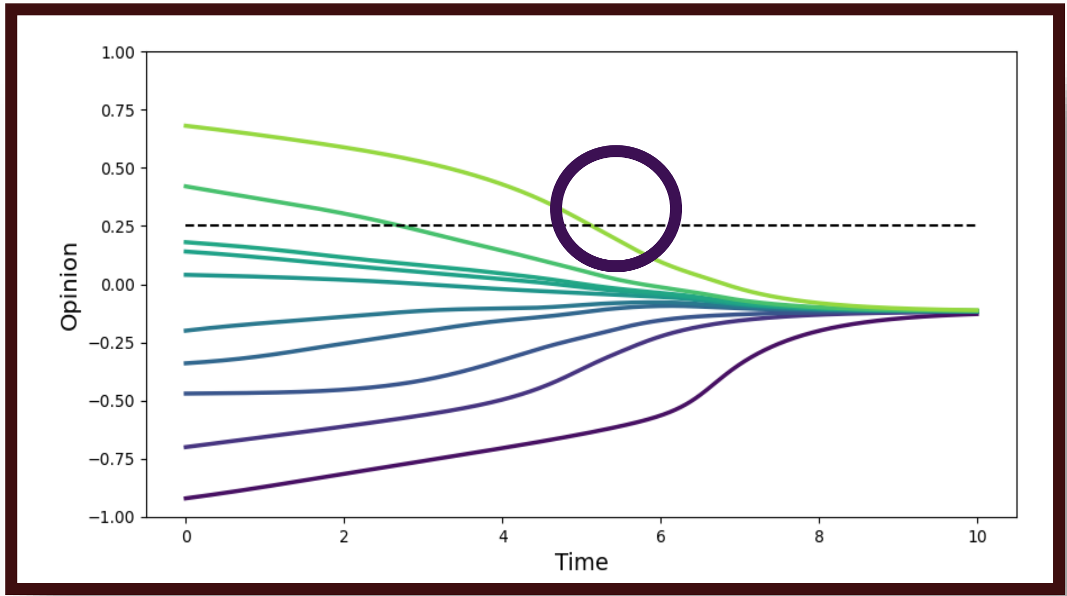
Controllability

Proposition: Let $N > 1$ and let the interaction function ϕ and the controls $f(u_{ij}, w_{ij})$ be sufficiently nice.

Then, for any $x^* \in [-1,1]$, there exist initial opinions $x(0) \in [-1,1]^N$ and initial edge weights $w(0) \in [0,1]^{N\times N}$ for which control to consensus at x^* is not possible.



Issue 1: Opinion fragmentation prevents consensus.



Issue 2: The target opinion is not inside the opinion interval.

Controllability

If we know the initial conditions $x(0)$ and $w(0)$, we can choose a consensus target x^* so that the system can be controlled. To show this, we introduce the following definitions:

We reorder our agents such that $x_1(0) \leq x_2(0) \leq \cdots \leq x_N(0)$, and define

$$x_m(t) = \min_i x_i(t), \qquad x^M(t) = \max_i x_i(t).$$

We say that an opinion vector $x(t) = \big(x_1(t), \dots, x_N(t)\big)$ is an **r-chain** if

$$|x_{i+1}(t) - x_i(t)| < r, \quad \text{for} \quad i = 1, \dots, N.$$

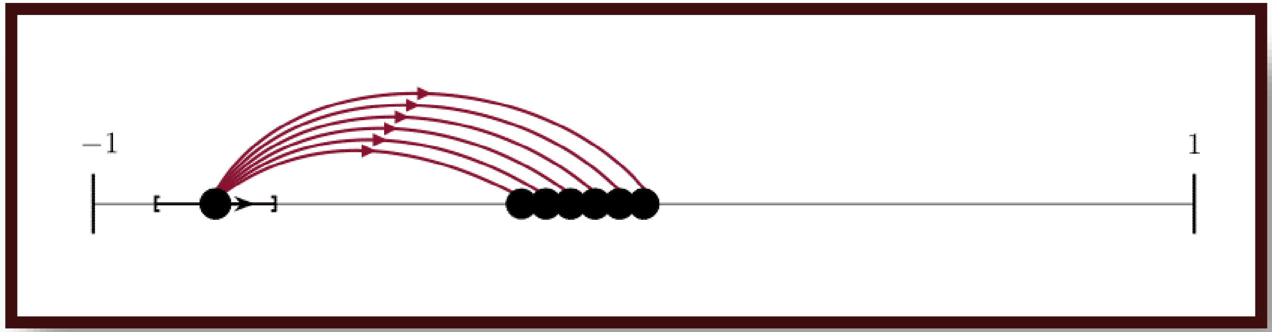
Proposition: Let $\mathrm{w}(0)$ be an empty network, $x(0)$ be an r-chain, and choose $x^* \in [x_m(0), x^M(0)]$. Assume also that there exists a control value $\mathrm{u}^+ \in \mathbb{R}$ for which

$$s^+ = s(u^+) > 0, \qquad \ell^+ = \ell(u^+) > 0,$$

that is, the control can be used to create new edges. Then, there exists a control $\mathrm{u} \in L^\infty(\mathbb{R}^+; \{0, u^+\}^{N\times N})$ such that the system reaches consensus at x^*.

Controllability

The proof is based on identifying the two individuals closer to x^* on either side, and gather individuals towards them, and once this is done, control these two individuals to move towards x^*.



From this result, we can conclude that **if the control acts *sufficiently quickly***, we can prevent individuals crossing over the target opinion: if we remove all edges connecting an individual to the rest of the population, we can limit their overall opinion change.

Finally, for an interaction function $\phi: [0,2] \to [0,1]$, we denote the set of its roots by $\mathcal{R}_\phi = \{r \in [0,2]: \phi(r) = 0\}$, and define $r^* = \inf \mathcal{R}_\phi$.

Controllability

Theorem: Let ϕ be a smoothed bounded confidence function with radius R.

Let the initial conditions be $x(0) \in [-1,1]^N$ and $w(0) \in [0,1]^{N\times N}$, such that $x(0)$ is an r-chain,

and choose a consensus target $x^* \in [x_m(0), x^M(0)]$.

Furthermore, assume that there exist u^+ as before and u^- such that

$$s^- = s(u^-) > 0 \qquad \text{and} \qquad \ell(u^-) = 0.$$

Then, for s^- sufficiently large, there exists a control $\mathrm{u} \in L^\infty(\mathbb{R}^+; \{0, u^+\}^{N\times N})$ such that the system reaches consensus at x^*.

Specifically, we can define

$$d_1 = \max_{i\in\{1,\ldots,N\}} \frac{k_i(0)+k_{i+1}(0)-2}{r^*-|x_i(0)-x_{i+1}(0)|}, \text{ and } d_2 = \max_{i\in\{1,\ldots,N\}, x_i(0)\neq x^*} \frac{k_i(0)-1}{|x_i(0)-x^*|},$$

and we have

$$s^- > D(0)\max\left\{d_1, d_2, \frac{4(N-1)}{R}\right\}.$$

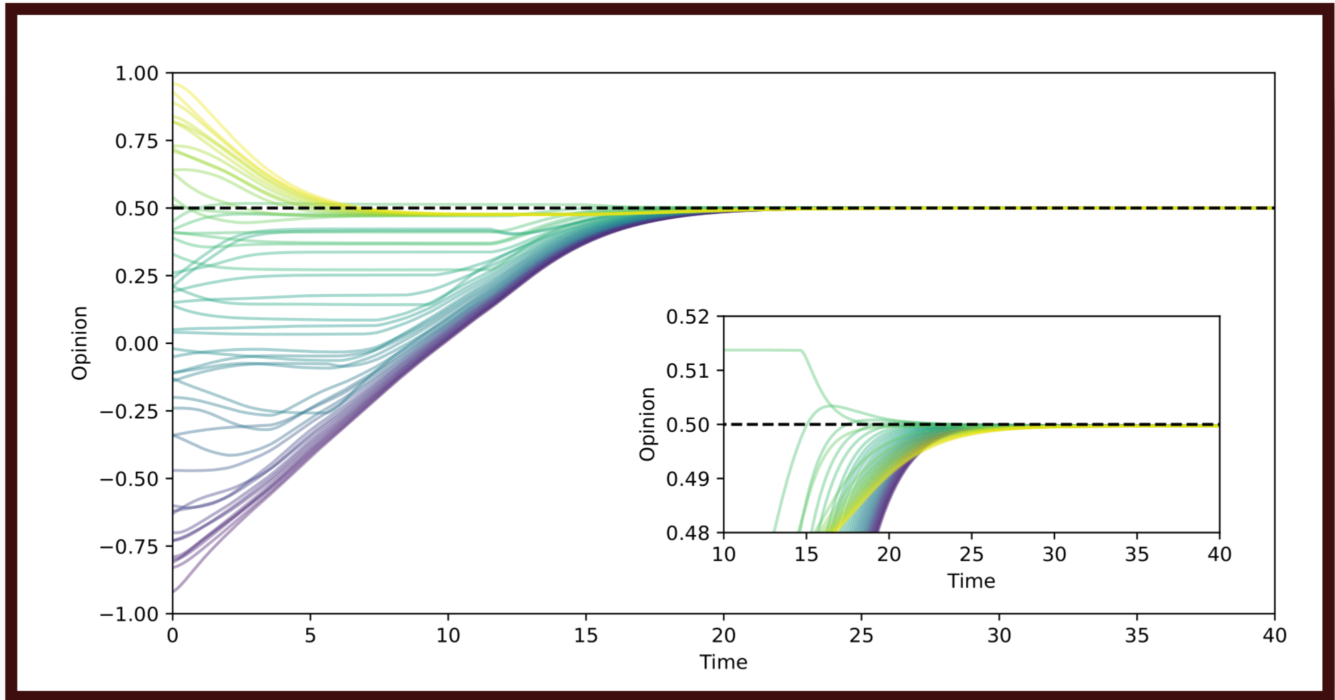
Controllability

In other words, if we assume that:

- the **target opinion** is in the interior of the initial points,
- the **initial opinions** are not already clustered, and
- the control **can create and remove edges anywhere within the network**,

Then if the control acts sufficiently quickly, consensus can be achieved at the desired target opinion.

Figure: Explicit control strategy used to prove previous theorem. Starting with a non-empty network, edges are created to gather the population closer to individuals near the target opinion $x^* = 0.5$ and removed to prevent crossing this value or splitting the population into multiple clusters. Individuals closest to x^* are then controlled to consensus precisely at this point.



Optimal control

Finally, we consider the optimal control problem, where we minimise a cost functional

$$\mathcal{C}(u, x(0), x^*) = \int_0^T \alpha \sum_{i=1}^{N} \sum_{j=1}^{N} u_{ij}(s)^2 + \beta \sum_{i=1}^{N} (x_i(s) - x^*)^2 \, ds,$$

subject to the opinion and controlled weight dynamics, and with the box constraints given by $u_{ij}(t) \in [-M, M]$, for some positive constant M.

By introducing a Hamiltonian, we can show that the adjoint dynamics are given by

$$\frac{dp_i}{dt} = 2\beta(x_i - x^*) - \sum_{j=1}^{N} \Big(\phi'(x_j - x_i)(x_j - x_i) + \phi(x_j - x_i)\Big)\left(\frac{w_{ji}}{k_j} p_j - \frac{w_{ij}}{k_i} p_i\right),$$

$$\frac{dq_{ij}}{dt} = q_{ij} s(u_{ij}) - \frac{p_i}{k_i} \phi(x_j - x_i)(x_j - x_i) + \frac{p_i}{k_i^2} \sum_{r=1}^{N} w_{ir} \phi(x_r - x_i)(x_r - x_i).$$

Optimal control

If we focus on the specific case when $M = 1$, so $u_{ij} \in [-1,1]$, and choose

$$s(u) = \mathcal{S}u^2, \qquad \ell(u) = \frac{1}{2}(u+1),$$

this problem has an analytical solution given by the **bang-bang control** $\tilde{u} = (\tilde{u}_{ij})$:

$$\tilde{u}_{ij} = \begin{cases} -M & if \quad q_{ij} < b_L(w_{ij}), \\ 0 & if \quad b_L(w_{ij}) \leq q_{ij} \leq b_U(w_{ij}), \\ M & if \quad q_{ij} > b_U(w_{ij}), \end{cases}$$

where $\mathrm{b_L(w)} = \begin{cases} -\infty & if\ w = 0, \\ -\frac{\alpha}{\mathcal{S}w} & if\ w > 0, \end{cases}$ and $\mathrm{b_L(w)} = \begin{cases} \infty & if\ w = 1, \\ \frac{\alpha}{\mathcal{S}(1-w)} & if\ w < 1. \end{cases}$

Optimal control

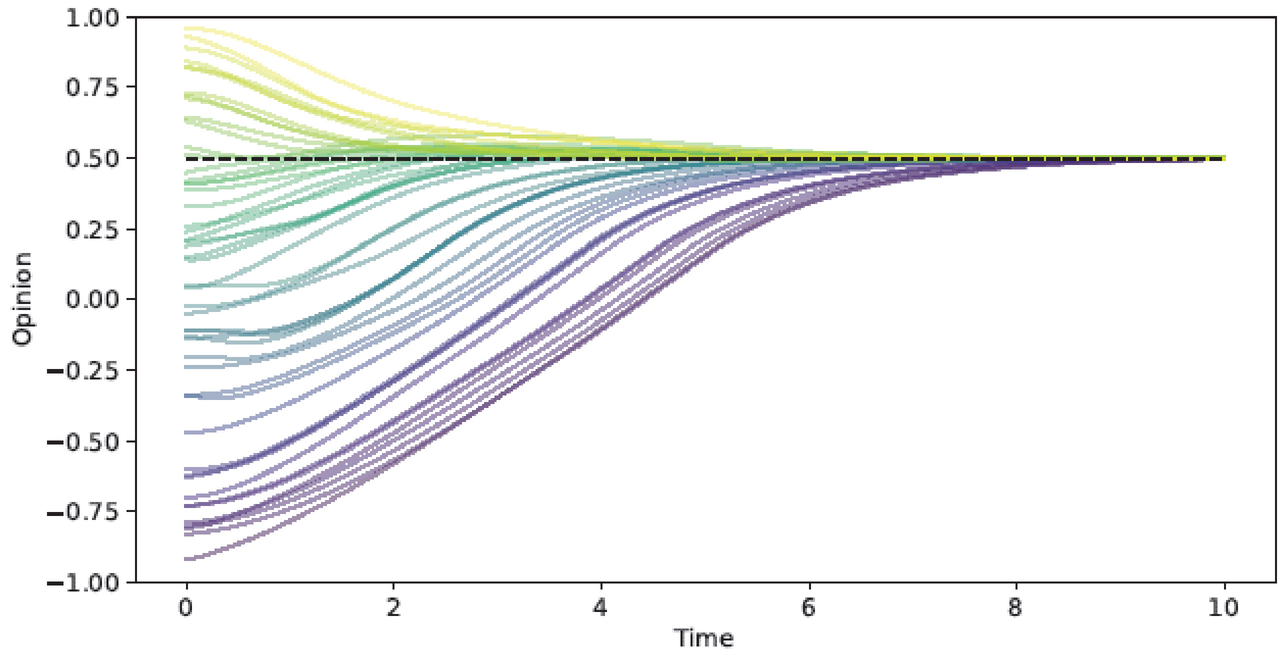


Figure: Controlled opinion dynamics under the optimal controls obtained by minimising the cost functional $\mathcal{C}$.

For the control video, (and all the other videos on the opinion dynamics models, see Andrew's webpage: https://warwick.ac.uk/fac/sci/mathsys/people/students/mathsysii/nugent/

Other related topics I am interested in...

- Have opinions evolving in a **higher dimensional space**
  - e.g. multiple opinions influencing each other, agents having an opinion on other agents' reliability
- **Mean-field limit** for large populations (without a social network)[16]
  - Already discussed control in this context this morning
- Mean-field limits **including social networks**
  - Includes investigating dynamics on graphons[17]
  - How does the geometry of the network influence the dynamics
- Making control more realistic/applicable by using, e.g., **sparse controls**
- Having other effects, such as lying agents, or age/duration of interactions, ...

16 – see Goddard, Gooding, Pavliotis and Short, IMA Journal of Applied Mathematics 2022.
17 – see Throm, European Journal of Applied Mathematics 2024.

Discussion

Conclusions

Collective dynamics is a rich field of mathematics, which can be explored from different perspectives. Today, we considered:

- Agent-based models – good for very specific, detailed models, ideally with not too many agents; can be computationally expensive.
- Interacting particle systems – more amenable to analysis, but can also be computationally expensive, e.g. when controlling via network weights.
- Mean-field limits – in cases when we can model the population as a whole, we can sometimes predict long-time behaviour and use this for control.

And our ability to use connections between these models to explore

- Inference based on information from individual trajectories,
- Control using social networks.

There's still a lot to explore, including several open questions. In the Exercises session, Andrew will explore some of the results we discussed, as well as the computational side of these systems.

Thank you for your attention!

Susana.gomes@warwick.ac.uk