WARWICK
THE UNIVERSITY OF WARWICK

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

Part 2: Inference for more complicated interactions, and some applications.

Susana Gomes (University of Warwick)

Woudschoten Conference 2023

Outline

Motivation

Application 1: Pedestrian dynamics

Application 2: Cell population dynamics

Discussion

WARWICK
THE UNIVERSITY OF WARWICK

Motivation

Today's talk

Yesterday we saw how to pass to the mean-field limit and explored one of the benefits of having a Fokker–Planck equation: characterising the long time behaviour of the system.

Today we will explore how to use this for inference. Recall that we are considering the system

$$dX_t^i = V'\left(X_t^i\right)\,\mathrm{d}t + \frac{1}{N}\sum_{j\neq i} K\left(X_t^i - X_t^j\right)\,\mathrm{d}t + \sqrt{2\sigma}\,\mathrm{d}W_t^i, \quad X_0^i = x_0^i.$$

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 this for two applications, and finish with some recent related work.

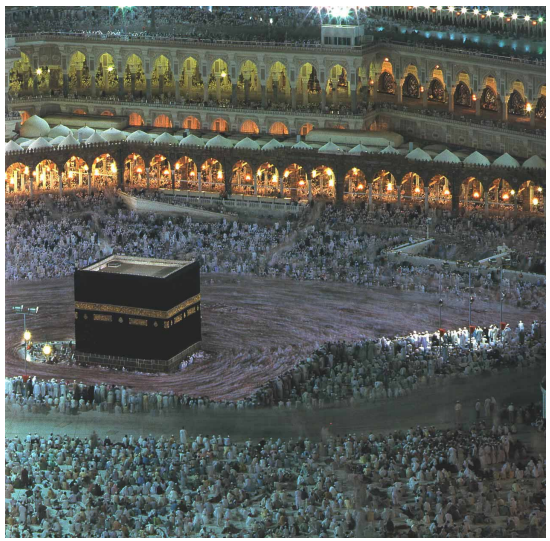
WARWICK
THE UNIVERSITY OF WARWICK

Application 1: Pedestrian dynamics

Joint work with Andrew Stuart (Caltech) and Marie-Therese Wolfram (Warwick)

SNG, A.M. Stuart, M.-T. Wolfram, SIAM J. Appl. Math 79(4), 1475-1500 (2019)

Crowd dynamics



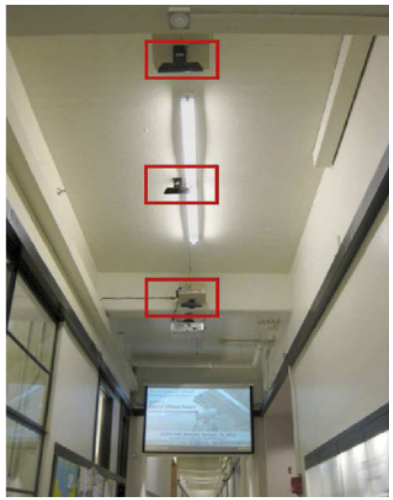
Understanding the individual dynamics as well as the evolution of a crowd is crucial in safety and transportation management.

Challenges:

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

Individual trajectories are obtained

...from overhead cameras, [1]



... or from controlled experiments.[2]

[1] Seer *et al.*, Transportation Research Procedia, 2014

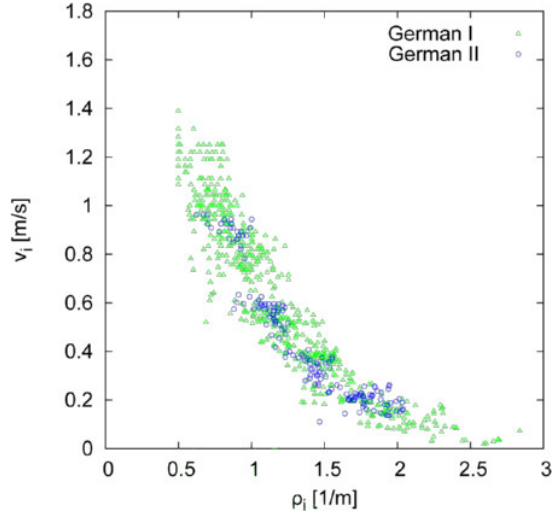
[2] BaSiGo experiments (Forschungszentrum Jülich)

Fundamental diagram[3]

The most widely used tool to characterize pedestrian dynamics is the fundamental diagram.

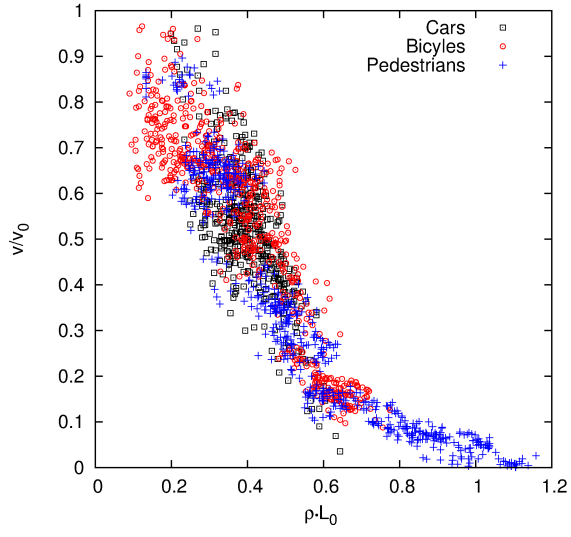
- Relates the experimentally observed density of pedestrians to their velocity or outflow.
- There is a general agreement on its basic shape...
- ... but its parameterization depends on measurement and averaging techniques, and experimental setup.

Students vs Soldiers



Cars vs Pedestrians vs Bikes

v_0:11.1(car),5.5(bicy),1.40(ped)[m/s], L_0:3.9(car),1.73(bicy),0.4(ped)[m]



[3] Figures from Zhang and Seyfried, Procedia Engineering, 2013.

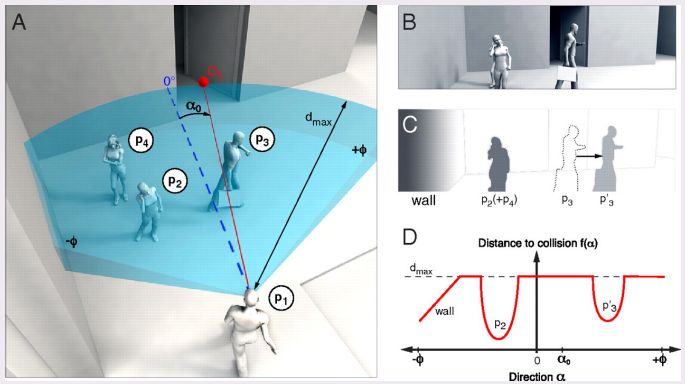
Modelling approaches

Microscopic approach

Agent based models based on Newton's laws.

$$\begin{aligned} dX_t^i &= V_t^i\, dt \\ dV_t^i &= F_i\left(X_t^1, \ldots, X_t^N, V_t^1, \ldots, V_t^N\right)\, dt \\ &\quad + G(X_t^i)\, dt + \sigma_i\, dB_t^i, \end{aligned}$$

X^i, V^i: position and speed of $i-$th particle, F, G: forces, B_t^i: white noise.



Examples include the **Social force model**[4] or the **Cellular automata model**[5]

Macroscopic approach

Usually nonlinear PDEs or conservation laws which describe an averaged quantity (usually density of pedestrians ρ). These PDEs can also be coupled with the eikonal equation.

Examples include the **Fokker–Planck equation** or the **Hughes model**.

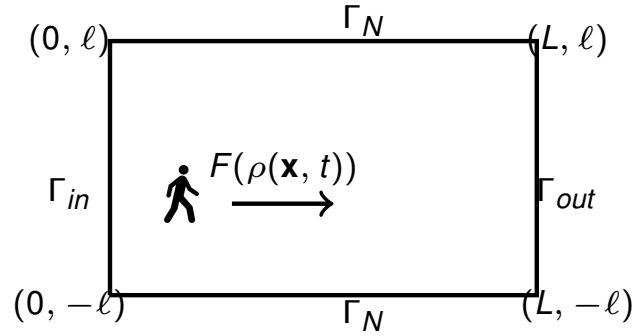
[4] Helbing and Molnar, Physical Review E, 1995, Moussaïd, Helbing and Theraulaz, PNAS, 2011

[5] Burstedde *et al*, Physica A, 2001

Our model

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

$$d\mathbf{X}_t = F(\rho(\mathbf{X}_t, t))\, dt + \sqrt{2\Sigma}\, d\mathbf{W}_t.$$



Here, $W(\cdot)$ is a $2D$ Brownian motion with diffusion $\Sigma = \mathrm{diag}(\sigma_x^2, \sigma_y^2)$, and

$$F(\rho(\mathbf{x}, t)) = v_{\max}\left(1 - \frac{\rho(\mathbf{x}, t)}{\rho_{\max}}\right)\mathbf{e_1}.$$

Importantly, $\rho(\mathbf{X}_t, t)$ for $\mathbf{x} = (x, y)$, is the density of the process.

The corresponding Fokker–Planck equation is

$$\rho_t = \nabla \cdot (\Sigma \nabla \rho - F(\rho)\rho\, \mathbf{e_1}),$$

with **boundary conditions**:

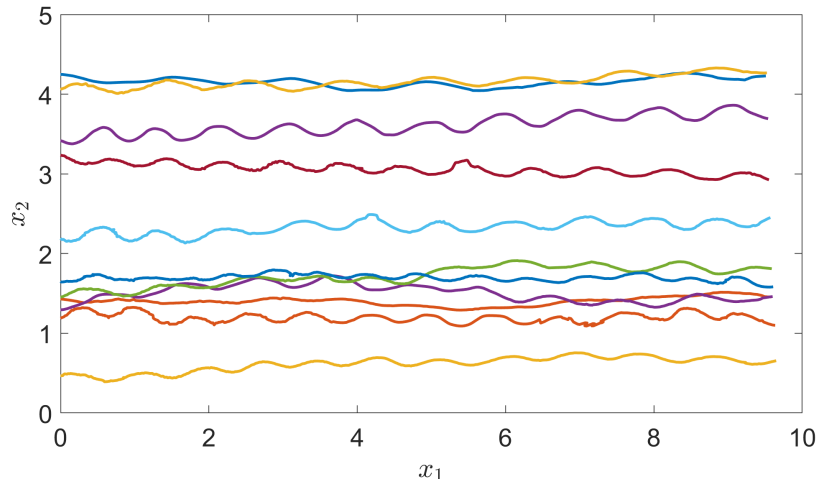
$$\begin{aligned}
(-\Sigma\nabla\rho + F(\rho)\rho)\cdot n &= -a(\rho_{\max} - \rho), && \text{for } \mathbf{x} \in \Gamma_{in},\\
(-\Sigma\nabla\rho + F(\rho)\rho)\cdot n &= b\rho, && \text{for } \mathbf{x} \in \Gamma_{out},\\
(-\Sigma\nabla\rho + F(\rho)\rho)\cdot n &= 0, && \text{for } \mathbf{x} \in \Gamma_N.
\end{aligned}$$

~~where n is the unit normal, a is the~~ **inflow rate** and b the **outflow rate**.

[6] We consider the mean-field limit, but can write this as IPS using $K(x) = v_{\max}(1 - \delta(x)/\rho_{\max})$

The inverse problem

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



Parameter estimation - usual inverse problem setting

Goal: Given a set of J trajectories $\left\{\mathbf{X}_t^j\right\}_{j=1}^J$, estimate $v_{\max}$.

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

$$\Phi(v; \mathbf{X}_t) = \frac{1}{4}\int_0^T |\dot{\mathbf{X}}_t - F\left(\rho(\mathbf{X}_t, t); v\right)|_\Sigma^2,$$

over all values of v, because it would be infinite almost surely.

Instead, we use the results from yesterday to show that it is equivalent to minimise the *log-likelihood* function of the process:

$$\Psi(v; \mathbf{X}_t) := L(\mathbf{X}_t; v) = \frac{1}{4}\int_0^T \Big(|F\left(\rho(\mathbf{X}_t, t); v\right)|_\Sigma^2 dt - 2\underbrace{\langle F\left(\rho(\mathbf{X}_t, t); v\right), d\mathbf{X}_t\rangle_\Sigma}_{\text{stochastic integral!}} \Big).$$

Possible issue: No idea if Ψ is convex, differentiable, ...

Including prior information and the Bayesian approach

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

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

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

$$e^{-\mathcal{J}(v;\mathbf{X}_t)\mathbb{1}(v>0)},$$

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

Summary

- Minimising Ψ over all possible v gives the *maximum likelihood estimator* (MLE) for v, which is also the *mean* of the distribution $\mathbb{P}(v|\mathbf{X}_t)$.

- Minimising $\mathcal{J}$ gives the *maximum a posteriori estimator* (MAP) for v. This is also the *mode* of $\mathbb{P}(v|\mathbf{X}_t)$.

- The full Bayesian approach involves sampling from $\mathbb{P}(v|\mathbf{X}_t)$ and provides the whole posterior distribution.

Parameter estimation approaches

We use $J = 20$ trajectories $\left\{\mathbf{X}_t^j\right\}_{t\in[0,T]}^{j=1,\ldots,20}$ 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 1*d* **and** strictly convex functions

Bayesian approach: MCMC with pCN algorithm[8]

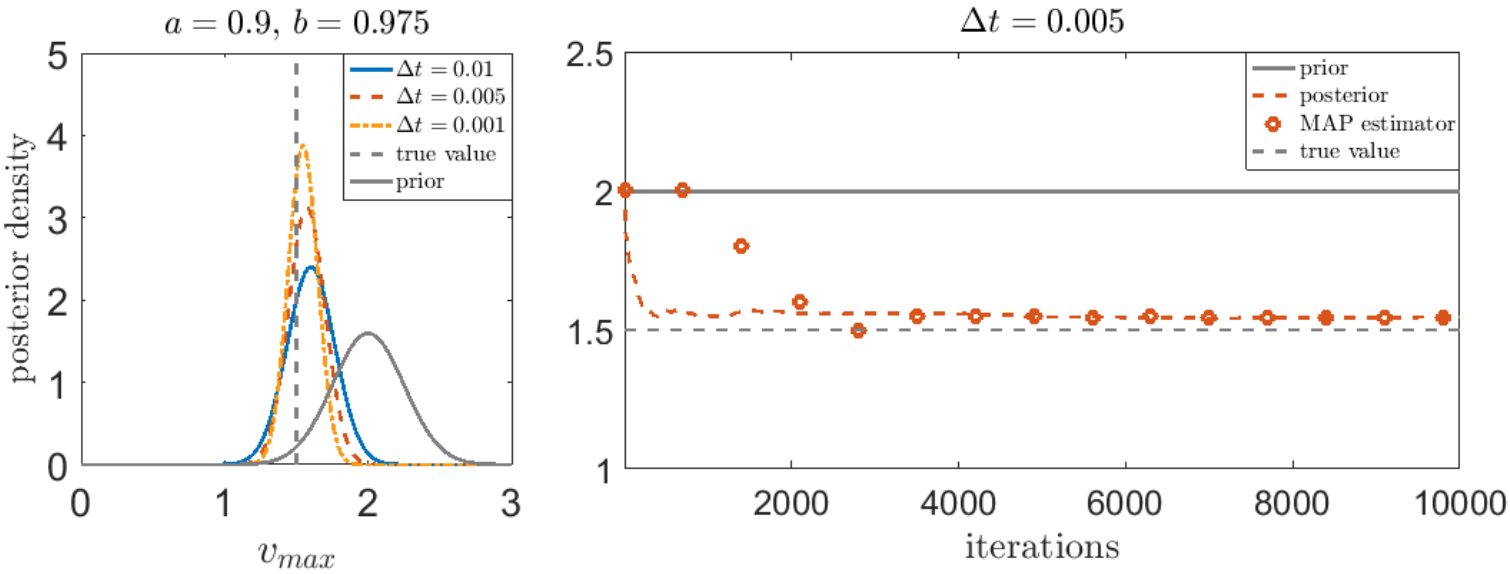
- \+ Provides posterior distribution
- \+ Allows for non-parametric (functional) estimation
- \+ Has a single tuneable parameter β - can be used to maximize efficiency
- \+ Allows for uncertanty quantification
- \+ Simple interpretation
- \- Requires many iterations

[7] Nelder and Mead, The Computer Journal, 1965.

[8] Cotter *et al.*, Statistical Science, 2013.

Bayesian methodology and Optimization

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



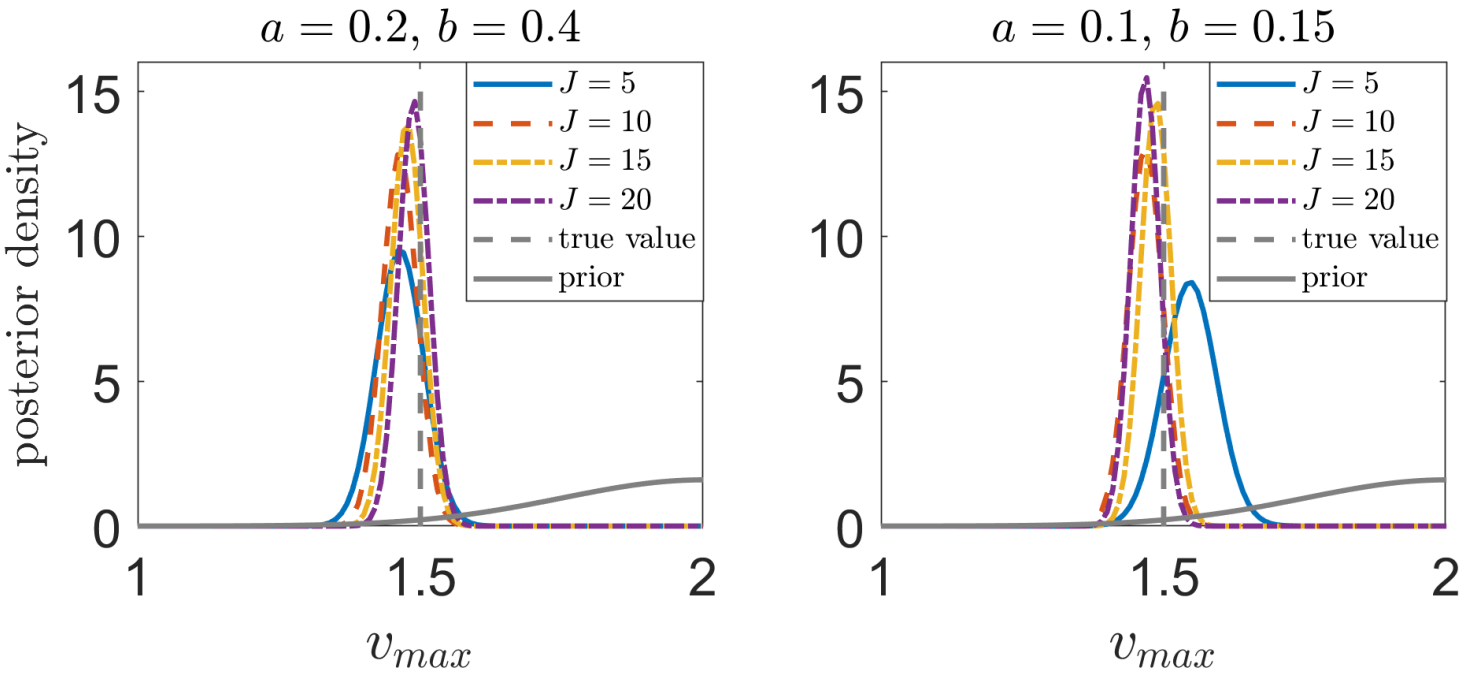
Left: Influence of the PDE time step on the posterior distribution.
Right: Comparison between Bayesian and optimisation approach for $\Delta t = 0.005$.

MAP estimator obtained by derivative-free optimization coincides with 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)

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[9] (10×5 corridor). We used experiment 5, corresponding to maximal current regime ($\min(a, b) \geq v_{\max}/2$).

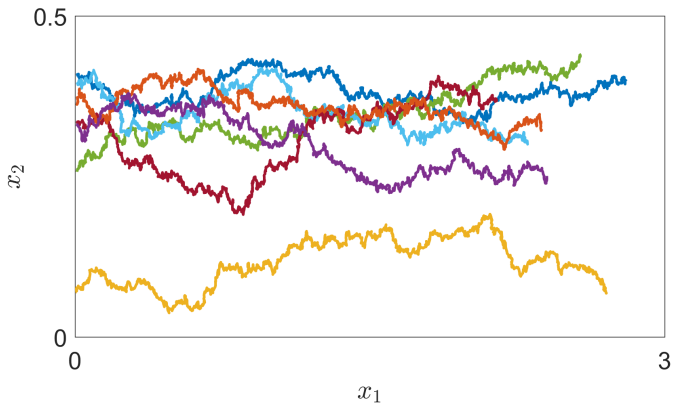
[9] https://ped.fz-juelich.de/db/doku.php?id=corridor5

What about real data?

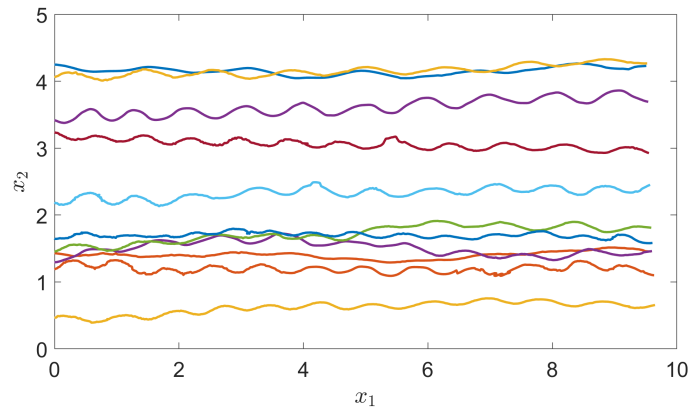
We tested the methodology using data from the BaSiGo experiments[9] (10×5 corridor). We used experiment 5, corresponding to maximal current regime ($\min(a, b) \geq v_{\max}/2$).

While it is not immediately clear that the method will work...

For example, the model generated trajectories do not look realistic:



data generated from our model



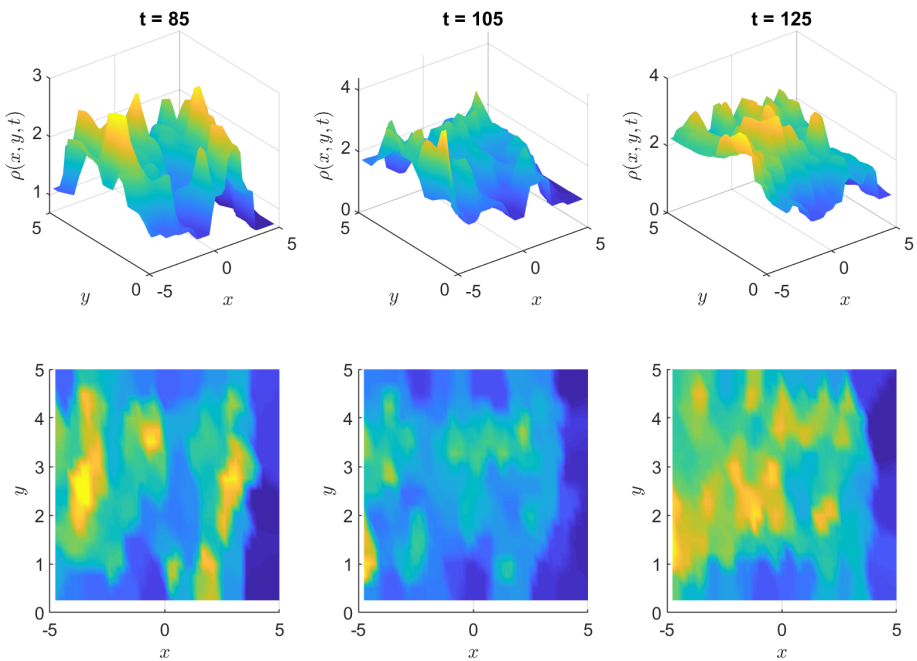
data from experiments Video

[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[9] (10×5 corridor). We used experiment 5, corresponding to maximal current regime ($\min(a, b) \geq v_{\max}/2$).

While it is not immediately clear that the method will work...
... and the observed density does not appear constant



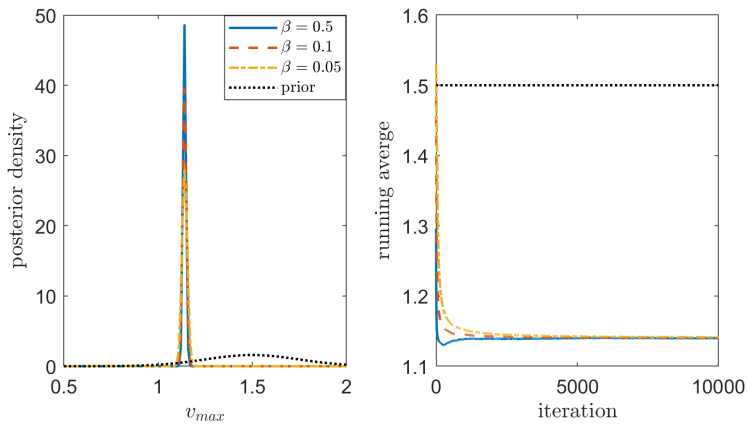
[9] https://ped.fz-juelich.de/db/doku.php?id=corridor5

What about real data?

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

While it is not immediately clear that the method will work...

Preliminary results show consistent estimates!



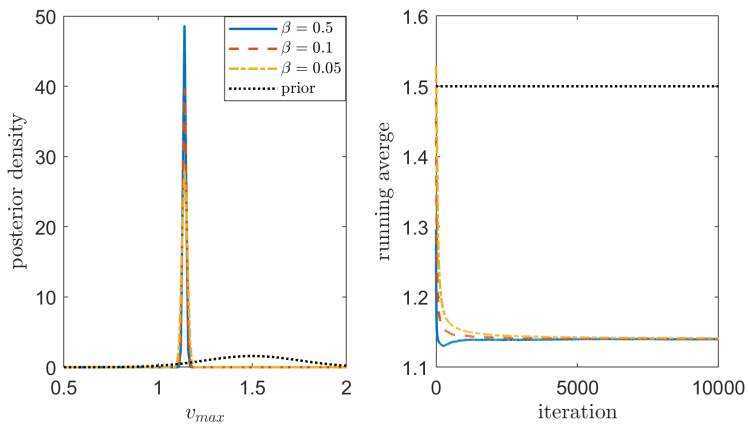
[9] https://ped.fz-juelich.de/db/doku.php?id=corridor5

What about real data?

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

While it is not immediately clear that the method will work...

Preliminary results show consistent estimates!



But still a lot of work to do!

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

WARWICK
THE UNIVERSITY OF WARWICK

Application 2: Cell population dynamics

Joint work with José Carrillo (Oxford) and Gissell Estrada-Rodriguez (Universitat Politècnica de Catalunya)

J.A. Carrillo, G. Estrada, SNG, in preparation

Motivation from cell dynamics

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**

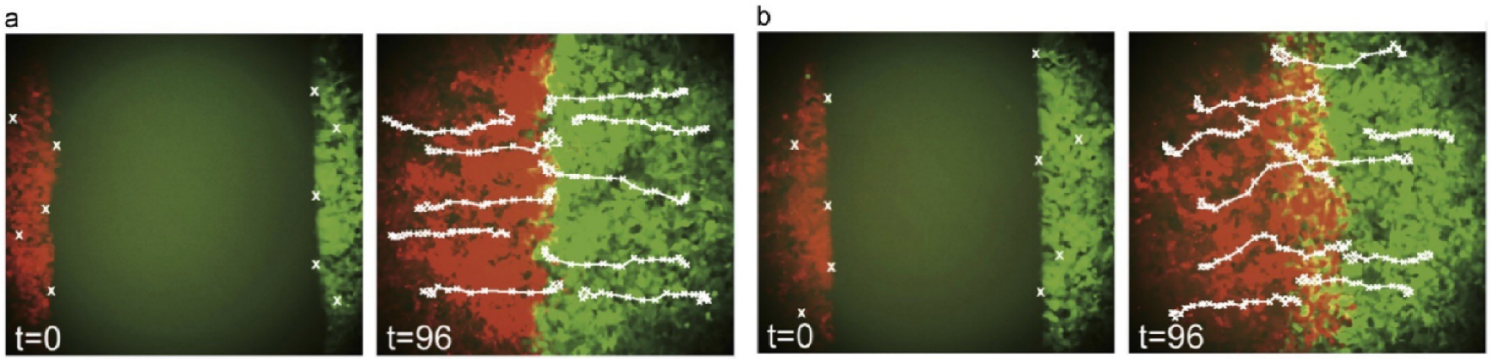


Figure: Cell–cell adhesion experiments, taken from Murakawa and Togashi 2015.[10]

Depending on cell properties, when two populations meet, they can **form a barrier**, or **mix at the boundary**, invading each other.

There are several experimental works showing this behaviour, and this has recently attracted mathematical interest.[11]

[10] H. Murakawa, H. Togashi, Journal of Theoretical Biology, 2015

[11] see, e.g., J.A. Carrillo, H. Murakawa, M. Sato, H. Togashi, O. Trush, Journal of Theoretical Biology, 2019.

A model for two-population cell dynamics

We consider two populations of cells **x** and **z**, where each cell interacts with other cells both in its own population and outside it, acccording to certain rules:

$$\mathrm{d}\mathbf{x}_i(t) = -\frac{1}{N_x}\sum_{i\neq j}\nabla\mathcal{F}_{11}(\mathbf{x}_i(t)-\mathbf{x}_j(t))\mathrm{d}t - \frac{1}{N_z}\sum_{j}\nabla\mathcal{F}_{12}(\mathbf{x}_i(t)-\mathbf{z}_j(t))\mathrm{d}t + \sqrt{2\Sigma_x}\mathrm{d}\mathcal{W}_t^{x,i}$$

$$\mathrm{d}\mathbf{z}_i(t) = -\frac{1}{N_z}\sum_{j\neq i}\nabla\mathcal{F}_{22}(\mathbf{z}_i(t)-\mathbf{z}_j(t))\mathrm{d}t - \frac{1}{N_x}\sum_{j}\nabla\mathcal{F}_{21}(\mathbf{z}_i(t)-\mathbf{x}_j(t))\mathrm{d}t + \sqrt{2\Sigma_z}\mathrm{d}\mathcal{W}_t^{z,i}$$

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

$$\nabla\mathcal{F}_{mn}(\mathbf{y}) = \nabla W_{mn}(\mathbf{y}) + \nabla M_{mn}^{b,\epsilon}(|\mathbf{y}|)$$

$$= \nabla\left(a_{mn}\left(\underbrace{(|\mathbf{y}|-\ell_{mn})^2}_{\text{short range repulsion}} - \underbrace{(R-\ell_{mn})^2}_{\text{long range attraction}}\right)\right) + \underbrace{\frac{b_{mn}}{(4\pi\varepsilon_{mn}^2)^{1/2}}\nabla e^{-\frac{|\mathbf{y}|^2}{4\varepsilon_{mn}^2}}}_{\text{local repulsuion (volume filling)}}$$

Mean-field limit and corresponding PDE

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

$$\mathrm{d}\mathbf{x}_i(t) = -\nabla(\mathcal{F}_{11} * \rho_1)(\mathbf{x}_i(t))\,\mathrm{d}t - \nabla(\mathcal{F}_{12} * \rho_2)(\mathbf{x}_i(t))\,\mathrm{d}t + \sqrt{2\Sigma_x}\mathrm{d}\mathcal{W}_t^{x,i}\ ,$$
$$\mathrm{d}\mathbf{z}_i(t) = -\nabla(\mathcal{F}_{22} * \rho_2)(\mathbf{z}_i(t))\,\mathrm{d}t - \nabla(\mathcal{F}_{21} * \rho_1)(\mathbf{z}_i(t))\,\mathrm{d}t + \sqrt{2\Sigma_z}\mathrm{d}\mathcal{W}_t^{z,i}\ .$$

where $*$ denotes convolution, and ρ_1 and ρ_2 are the two populations of cells.

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

$$\begin{aligned}\partial_t\rho_1 &= \nabla\cdot(\rho_1\nabla(b_1(\rho_1+\rho_2) + W_{11}*\rho_1 + W_{12}*\rho_2)) + \Sigma\Delta\rho_1\ ,\\ \partial_t\rho_2 &= \nabla\cdot(\rho_2\nabla(b_2(\rho_1+\rho_2) + W_{22}*\rho_2 + W_{21}*\rho_1)) + \Sigma\Delta\rho_2\ .\end{aligned}$$

In this case, we would like to estimate all parameters in the equations, i.e., $\theta = (b_1,\ b_2,\ a_{mn},\ \ell_{mn},\ \epsilon_{mn})$.

Inference framework

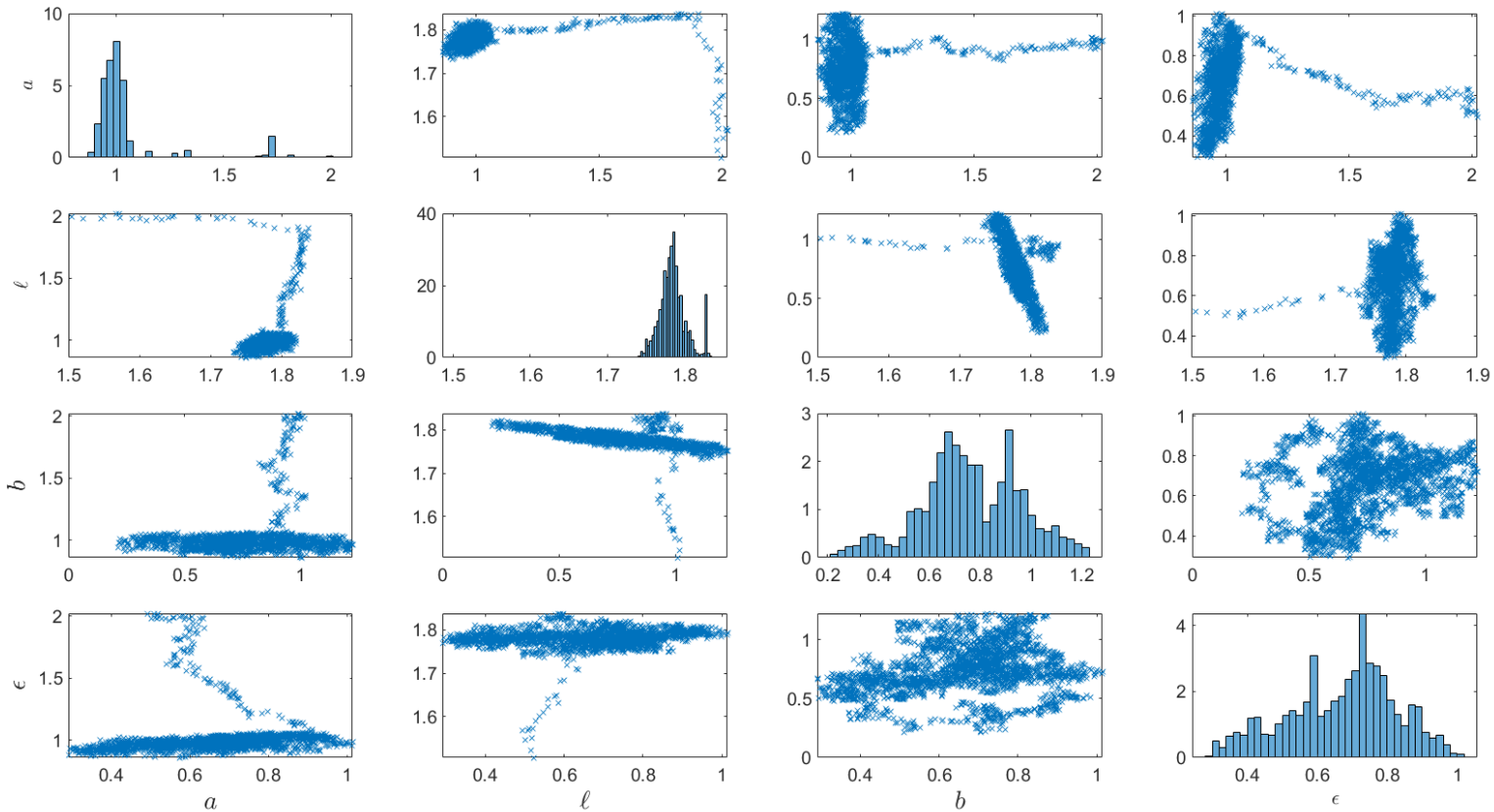
As before, we wish to minimise the negative log-likelihood function, which in this case is

$$
\begin{aligned}
\Psi(\theta; \{\mathbf{x}(t), \mathbf{z}(t)\}) \quad := \quad & \sum_{i \in \mathcal{P}_1} \frac{1}{4} \int_0^T \left| \left(\mathcal{F}_{11}^{\theta} * \rho_1\right)\left(\mathbf{x}_i(t)\right) \right|_{\Sigma_x}^2 \, \mathrm{d}t \\
& + \sum_{i \in \mathcal{P}_1} \frac{1}{4} \int_0^T 2 \left\langle \left(\mathcal{F}_{11}^{\theta} * \rho_1\right)\left(\mathbf{x}_i(t)\right) + \left(\mathcal{F}_{12}^{\theta} * \rho_2\right)\left(\mathbf{x}_i(t)\right), d\mathbf{x}_t \right\rangle_{\Sigma_x} \\
& + \sum_{i \in \mathcal{P}_2} \frac{1}{4} \int_0^T \Big(\left| \left(\mathcal{F}_{22}^{\theta} * \rho_2\right)\left(\mathbf{z}_i(t)\right) \right|_{\Sigma_z}^2 \, \mathrm{d}t \\
& + \sum_{i \in \mathcal{P}_2} \frac{1}{4} \int_0^T 2 \left\langle \left(\mathcal{F}_{22}^{\theta} * \rho_2\right)\left(\mathbf{z}_i(t)\right) + \left(\mathcal{F}_{21}^{\theta} * \rho_1\right)\left(\mathbf{z}_i(t)\right), \mathrm{d}\mathbf{z}_t \right\rangle_{\Sigma_z},
\end{aligned}
$$

where we sum over all available trajectories, $\mathcal{P}_1$ is the population of cells of type $\mathbf{x}$ and $\mathcal{P}_2$ is the population of cells of type $\mathbf{z}$.

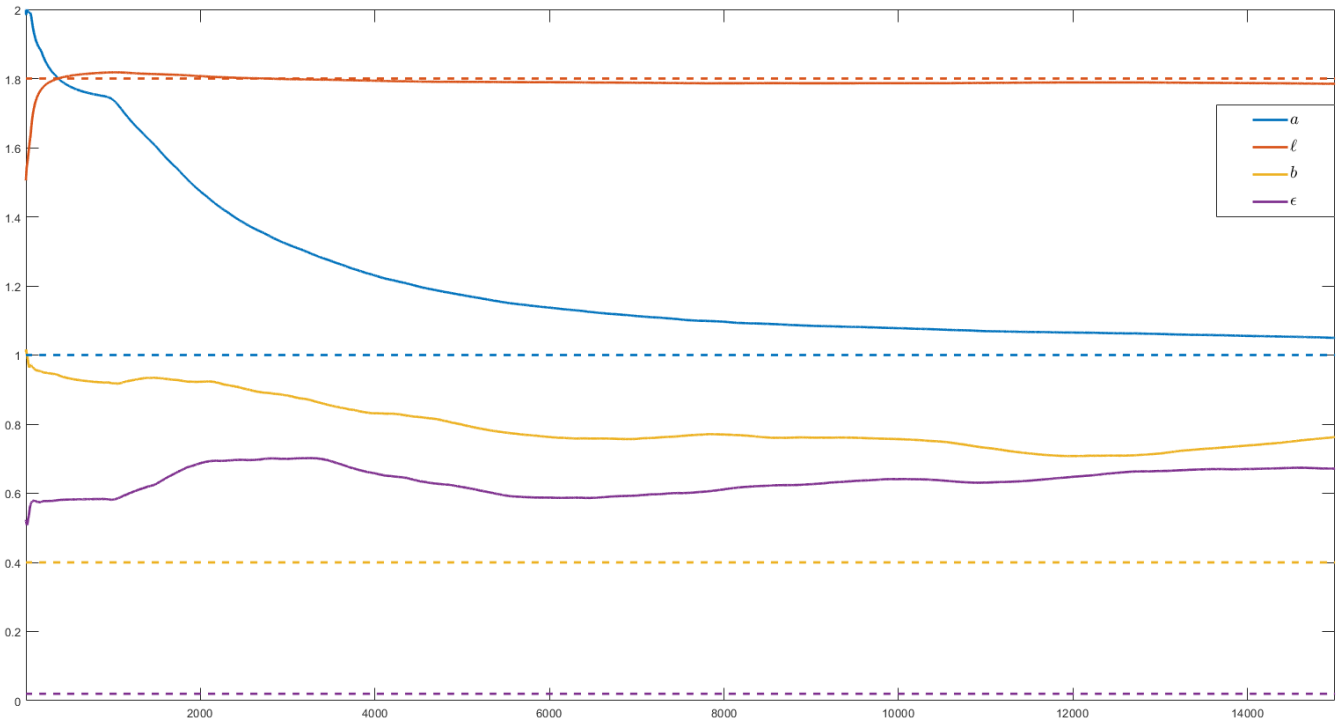
Results using generated data

Generated data for **one population** of cells, with true parameters $a = 1$, $\ell = 1.8$, $b = 0.4$, $\epsilon = 0.0208$.



Posterior distributions after 15000 interations of the MCMC algorithm.

Results using generated data

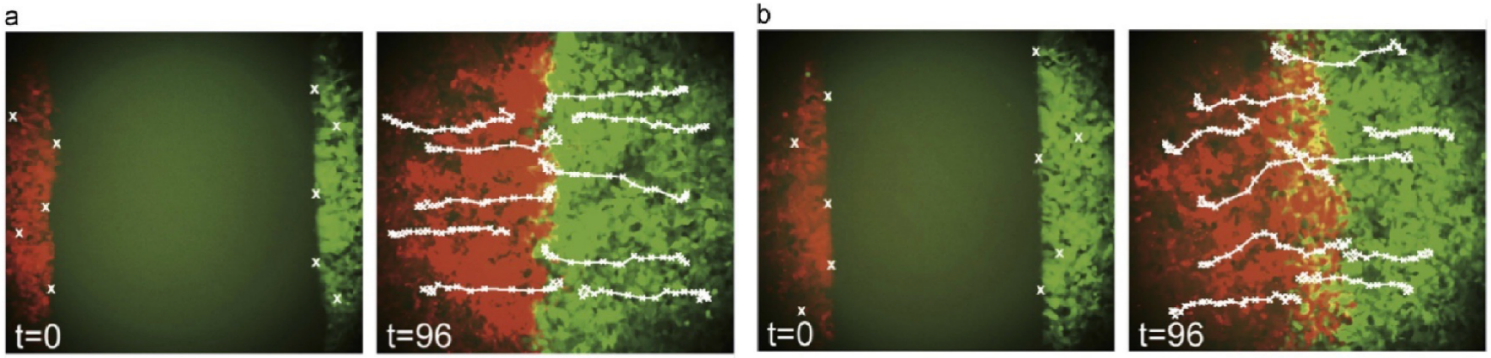


Mean of each (marginal) posterior distribution as a function of iteration number for 15000 interations of the MCMC algorithm (full lines) and true value (dashed lines).

What about real data?

Trajectory 1		Trajectory 2		Trajectory 3		Trajectory 4	
X	Y	X	Y	X	Y	X	Y
1.644727944	0	1.644727944	3.289455887	1.644727944	3.289455887	1.644727944	0
19.73673532	0	16.44727944	1.644727944	13.15782355	-13.15782355	4.934183831	-11.51309561
55.92075008	-16.44727944	34.53928682	-29.60510299	27.96037504	-29.60510299	11.51309561	-11.51309561

Snapshot of dataset from Murakawa and Togashi, 2015.



Experiments from Murakawa and Togashi, 2015, with marked trajectories.

... But we are working on it!

Key message: Dialogue is important, *before* conducting experiments if possible!

WARWICK
THE UNIVERSITY OF WARWICK

Discussion

Conclusions

Inverse problems when the model involves noise cannot be solved using standard techniques – I presented a *likelihood based* parameter estimation framework for SDEs which depend on their density.

- Two examples: pedestrian dynamics and cell population dynamics, where we also considered two-population models.
- Each model consists of SDEs for individual trajectories, coupled with one (or two) PDEs for the density of a crowd or of cell populations.
- Using this methodology, we can estimate some parameters, while others are not identifiable.
- Good estimates can be obtained both from generated data and trajectories from experiments.
- Future work involves using real data and modifying the methodology to estimate (if possible) parameters which were not identifiable so far.

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[12]
- Likelihood free inference, for when it is not possible to find likelihood functions. This is the basis of **Approximate Bayesian Computation (ABC)**[13]
- Sequential Monte Carlo techniques for when one has positions but not trajectories[14]
- Modelling and inference for pedestrians and traffic flow models[15]
- Work on the mean-field approximations for cell dynamics models[16]
- Several other interesting research avenues (happy to discuss later)

[12] see Sharrock et al, Stoch Proc Appl 2023, Pavliotis and Zanoni, SIAM J Appl Dyn Syst 2022, Amorino et al, Stoch Proc Appl, 2023.

[13] see Toni et al., Journal of the Royal Society Interface, 2009, or Lintusaari, et al., Systematic biology, 2017.

[14] see Cheng, Wen and Li, Roy Soc Open Sci, 2023.

[15] Wurth et al, Adv Comp Mat 2022, Corbetta and Toschi, Annual Review of Condensed Matter Physics, 2023, Gödel et al, Safety Science 2022.

[16] Morale, Capasso, Oelschläger, J Math Bio, 2005, Burger, Capasso, Morale, Nonlinear Analysis: Real World Applications, 2007.

Other activities

Working on these projects with clear applications opened a lot of other doors for me too...

Me in the Houses of Parliament in London in 2019 (left) for the STEM for Britain competition, and in 2021 (right) after being shortlisted for the L'Oreal-UNESCO For Women In Science Fellowship in 2020.

Thank you for your attention!

Susana N. Gomes

https://warwick.ac.uk/fac/sci/maths/people/staff/gomes

susana.gomes@warwick.ac.uk

WARWICK
THE UNIVERSITY OF WARWICK