


PAPER

Gradient-based parameter calibration of an anisotropic interaction model for pedestrian dynamics

Zhomart Turarov¹ and Claudia Totzeck² 

¹RPTU Kaiserslautern, Kaiserslautern 67663, Germany and ²BU Wuppertal, Wuppertal 42119, Germany

Corresponding author: Claudia Totzeck; Email: totzeck@uni-wuppertal.de

Received: 17 March 2022; **Revised:** 15 May 2023; **Accepted:** 17 May 2023; **First published online:** 18 July 2023

Keywords: 76A30; 82C22; 49K15; 49M05

2020 Mathematics Subject Classification: 76A30 (Primary); 49K15 (Secondary)

Abstract

We propose an extension of the anisotropic interaction model which allows for collision avoidance in pairwise interactions by a rotation of forces (Totzeck (2020) *Kinet. Relat. Models* **13**(6), 1219–1242.) by including the agents' body size. The influence of the body size on the self-organisation of the agents in channel and crossing scenarios as well as the fundamental diagram is studied. Since the model is stated as a coupled system of ordinary differential equations, we are able to give a rigorous well-posedness analysis. Then we state a parameter calibration problem that involves data from real experiments. We prove the existence of a minimiser and derive the corresponding first-order optimality conditions. With the help of these conditions, we propose a gradient descent algorithm based on mini-batches of the data set. We employ the proposed algorithm to fit the parameter of the collision avoidance and the strength parameters of the interaction forces to given real data from experiments. The results underpin the feasibility of the method.

1. Introduction

Mathematical modeling of pedestrian dynamics has a practical benefit in civil engineering [23, 26, 27] for example in the design of complex architectures, e.g., stadiums, city centers, and shopping malls, or for the management of large public events like festivals, concerts, pilgrimages, or manifestations [10]. Capturing both, the individual and collective behaviors in pedestrian dynamics, is rather complex [2, 7]. Many different approaches have been proposed in the literature: for example, models based on magnetic forces proposed by S. Okazaki and S. Matsushita in which pedestrians are modeled as magnetic charges in a magnetic field [33]; the gas-kinetic model which treats pedestrians as molecules in liquefied gas [22]; cellular automata [4, 5, 13]; models incorporating anticipative, rational behavior [1, 11, 12] and (smooth) sidestepping [15, 34]. Another group of pedestrian models has emerged from the pioneering work on social forces [20] and can be classified as force-based [9, 34] and the overview given in [8].

Most of these models share the property of reproducing collective features such as lane formation in counter-flow scenarios and traveling waves in crossing flows. Moreover, they can be used to study evacuation scenarios. On the other hand, they strongly differ in their description. Indeed, some models, for example, the class of forces-based models have a sound mathematical description and allow for a statement in terms of a closed system of ordinary or partial differential equations. Others have a rather algorithmic structure because they require the solution of optimization problems to estimate for example the time-to-collision in every iteration. For the latter, a rigorous mathematical study of well-posedness is difficult.

Naturally, the modeling process is followed by an optimization or calibration procedure. For pedestrian dynamics, the optimization of buildings, evacuation routes, and traffic safety or the minimization

of the occurrence of high densities are of special interest [28, 37]. Moreover, the collection of data from real-world experiments grew the interest in parameter fitting for the different pedestrian models [3, 17, 18].

This work is in a similar spirit. First, we extend the anisotropic model proposed in [34] by incorporating a body size for the interacting agents. This induces another dimension of volume exclusion in the model and makes the model more realistic. However, it is simple enough to derive the gradient used for the optimization explicitly. We emphasize that this is exceptional, as many other particle models for pedestrian dynamics include terms like the Heaviside function in the social force model or an optimization problem in their dynamics that prohibit the straight forward computation of adjoints and gradients. We study the influence of the body size on the formation of lanes and traveling waves as well as the fundamental diagram of the dynamics numerically and provide a rigorous study of the well-posedness of the interaction dynamics with and without body size employing standard ODE theory. The second part of the article is concerned with the rigorous derivation of a gradient-based parameter calibration algorithm. We begin with the derivation of the first-order optimality system and propose a mini-batch gradient-descent algorithm for the calibration problem.

In more detail, the article is organized as follows: the anisotropic model for pedestrian dynamics is extended by including the agents' body size in Section 2. Moreover, we study the influence of body size on collective behavior and the fundamental diagram in there. Section 3 is devoted to the parameter calibration problem. We begin with the statement of the problem, investigate its well-posedness, and derive the corresponding first-order optimality conditions. Finally, the iterative gradient descent algorithm based on mini-batches for the parameter calibration is proposed in Section 4, where we show results obtained with this algorithm. We conclude with a summary of the results and an outlook on future projects.

2. Microscopic model with body size

We include the body size in the anisotropic model proposed in [34] as follows: let us consider a second-order equation of motion with $N \in \mathbb{N}$ agents. Their positions and velocities are denoted by $x_i : [0, T] \rightarrow \mathbb{R}^2$ and $v_i : [0, T] \rightarrow \mathbb{R}^2$, $i = 1, \dots, N$. Moreover, the agents are assumed to have a body diameter $d > 0$. This leads to the following interaction dynamics

$$\frac{d}{dt}x_i = v_i, \tag{2.1a}$$

$$\frac{d}{dt}v_i = \tau (w_i - v_i) - \frac{1}{N} \sum_{j \neq i} M(v_i, v_j) K(d, x_i, x_j, v_i, v_j), \tag{2.1b}$$

$$x_i(0) = x_0^i, \quad v_i(0) = v_0^i, \quad i = 1, \dots, N, \tag{2.1c}$$

where $K(d, x_i, x_j, v_i, v_j) : \mathbb{R}^2 \times \mathbb{R}^2 \times \mathbb{R}^2 \times \mathbb{R}^2 \rightarrow \mathbb{R}^2$ is a pairwise interaction force between the agents i and j . The rotation matrix $M(v_i, v_j)$ changes the direction of the interaction force. It reads

$$M(v_i, v_j) = \begin{pmatrix} \cos \alpha_{ij} & -\sin \alpha_{ij} \\ \sin \alpha_{ij} & \cos \alpha_{ij} \end{pmatrix}, \alpha_{ij} = \begin{cases} \lambda \arccos \frac{v_i \cdot v_j}{\|v_i\| \|v_j\|}, & \text{if } v_i \neq 0, v_j \neq 0, \\ 0, & \text{else.} \end{cases} \tag{2.2}$$

In addition, the model includes a relaxation parameter $\tau > 0$ which controls the adaption of the current velocity v_i towards the given desired velocity $w_i \in C([0, T], \mathbb{R}^2)$. In general, the desired velocity w can be time dependent. For example, in evacuation scenarios, we may obtain the desired velocity with the help of the Eikonal equation. Here, we focus on simple scenarios to understand the basic properties and parameters of the model, we therefore set the desired velocity to a constant value for each specific scenario. The rotation of the force vectors induced by the matrix M models a collision avoidance behavior of the agents. The direction of the collision avoidance is controlled by the sign of the parameter λ . For

$\lambda > 0$ agents move to the right, to avoid a collision, for $\lambda < 0$ the movement is directed to the left. See [34] for further details.

For notational convenience, the solution of the system is expressed by the vectors $\mathbf{x}(t) = (x_1(t), \dots, x_N(t)) \in \mathbb{R}^{2N}$ and $\mathbf{v}(t) = (v_1(t), \dots, v_N(t)) \in \mathbb{R}^{2N}$ for $t \in [0, T]$.

Remark 1. We can easily include obstacles or walls in the model, by describing them as artificial agents with fixed positions and fixed velocities and adding an additional interaction term similar to the interaction of the agents in (2.1b).

2.1. Well-posedness

In this section, we study the well-posedness of the dynamics given in (2.1). We make the following assumptions on the interaction force $K(d, x_i, x_j, v_i, v_j)$ with $i, j \in \{1, \dots, N\}$.

Assumption 1. The interaction forces $K(d, x_i, x_j, v_i, v_j)$ are locally Lipschitz and globally bounded with respect to d , the positions x_i, x_j and the velocities v_i, v_j .

Assumption 2. The gradients of interaction forces $\nabla K(d, x_i, x_j, v_i, v_j)$ exist, are locally Lipschitz and globally bounded with respect to the positions x_i, x_j and the velocities v_i, v_j .

Remark 2. Note that the first assumption is necessary to show the well-posedness of (2.1) while we need the second assumption later on to obtain the well-posedness of the calibration problem.

A key step to derive the well-posedness of the system is to establish the Lipschitz property of the right-hand side. In particular, the rotation of the force vector is of interest.

Lemma 1. Let Assumption 1 hold. For the rotation of the force term with $v_i, v_j, v_k, v_\ell \in \mathbb{R}^2$ and $d \geq 0$ it holds

$$|M(v_i, v_j)K(d, x_i, x_j, v_i, v_j) - M(v_k, v_\ell)K(d, x_k, x_\ell, v_k, v_\ell)| \leq L_1|(v_i, v_j) - (v_k, v_\ell)| + L_2|(x_i, x_j) - (x_k, x_\ell)|$$

for some Lipschitz constants $L_1, L_2 > 0$.

Proof. We introduce the short hand notation

$$v^1 = (v_i, v_j), \quad v^2 = (v_k, v_\ell), \quad x^1 = (x_i, x_j), \quad x^2 = (x_k, x_\ell).$$

In the following, we omit the dependence on d . Hence we rewrite the left-hand side of the Lipschitz condition as

$$|M(v^1)K(x^1, v^1) - M(v^2)K(x^2, v^2)| \leq |K(x^1, v^1)|_\infty |M(v^1) - M(v^2)| + |M(v^2)|_\infty |K(x^1, v^1) - K(x^2, v^2)|. \tag{2.3}$$

We estimate the first term on the right-hand side of (2.3) as

$$|M(v^1) - M(v^2)| \leq \left| \int_0^1 \nabla M(v^2 + s(v^1 - v^2)) ds \right| |v^1 - v^2|, \tag{2.4}$$

where

$$\nabla M(v^1) = \left(\frac{dM}{dv_i}(v^1), \frac{dM}{dv_j}(v^1) \right),$$

with

$$\frac{dM}{dv_i} = \begin{pmatrix} -\sin \alpha & -\cos \alpha \\ \cos \alpha & -\sin \alpha \end{pmatrix} \cdot \frac{d\alpha}{dv_i},$$

$$\frac{d\alpha}{dv_i} = \begin{cases} -\lambda \frac{1}{\sqrt{(\|v_i\| \|v_j\|)^2 - (v_i, v_j)^2}} \left(v_j - \langle v_i, v_j \rangle \frac{v_i}{\|v_i\|^2} \right), & \text{if } v_i, v_j \neq 0 \\ 0, & \text{else} \end{cases}.$$

Analogously, we derive $\frac{dM}{dv_j}$. Each element of $\nabla M(v^2 + s(v^1 - v^2))$ is bounded, for a detailed proof of the boundedness see Appendix A. Note that K is globally bounded by Assumption 1. Moreover, K is locally Lipschitz by Assumption 1 which allows us to estimate the second term on the right-hand side of (2.3). Altogether, this proves the existence of the Lipschitz constants L_1 and L_2 as desired. \square

Lemma 2 (Existence and Uniqueness). *Let Assumption 1 hold. Further we assume $w_i \in C([0, T], \mathbb{R}^2)$, $i = 1, \dots, N$ and $\lambda \in [-1, 1]$.*

Then system (2.1) admits a unique solution $\mathbf{x} \in C^1([0, T], \mathbb{R}^{2N})$, $\mathbf{v} \in C^1([0, T], \mathbb{R}^{2N})$.

Proof. On the basis of Assumption 1 and Lemma 1, the result can be directly obtained with the help of the Picard-Lindelöf theorem. \square

2.1.1. Body size

In the previous discussion, the body size is an abstract parameter. To give more details, we consider a variation of the Morse potential [14] leading to the interaction potential

$$U(d, \|x_i - x_j\|) = R \cdot e^{-\frac{d - \|x_i - x_j\|}{r}} - A \cdot e^{-\frac{d - \|x_i - x_j\|}{a}}$$

leading to the forces K given by

$$K(d, \|x_i - x_j\|) = \left(\frac{A}{a} \cdot e^{-\frac{d - \|x_i - x_j\|}{a}} - \frac{R}{r} \cdot e^{-\frac{d - \|x_i - x_j\|}{r}} \right) \cdot \frac{x_i - x_j}{\|x_i - x_j\|}. \tag{2.5}$$

Remark 3. *Note that the forces depending on the body size as given in (2.5) with the standard regularisation satisfy the assumptions of Lemma 2. Hence, we have the existence and uniqueness of solutions for the case with body size $d > 0$ as well.*

2.2. Numerical studies

For the numerical studies of the model, we draw random initial positions with uniform distribution in the domain and set initial velocities with respect to the desired direction of motion. We set the initial velocity to the desired velocity. Then we solve (2.1) with a variant of the leap-frog scheme [34]. Indeed, the relaxation terms are solved implicitly and the interaction is solved explicitly as given by

$$x_i^{k'} = x_i^k + \frac{\Delta t}{2} v_i^k, \quad v_i^{k'} = (v_i^k + \Delta t \cdot \tau \cdot w_i) / (1 + \Delta t \cdot \tau)$$

$$v_i^{k+1} = v_i^{k'} + \Delta t \cdot \frac{1}{N} \sum_{j \neq i} M(v_i^{k'}, v_j^{k'}) \cdot K(x_i^{k'}, x_j^{k'}), \quad x_i^{k+1} = x_i^{k'} + \frac{\Delta t}{2} v_i^{k+1}, \quad i = 1, \dots, N, \tag{2.6}$$

where Δt denotes the step size of the time discretization.

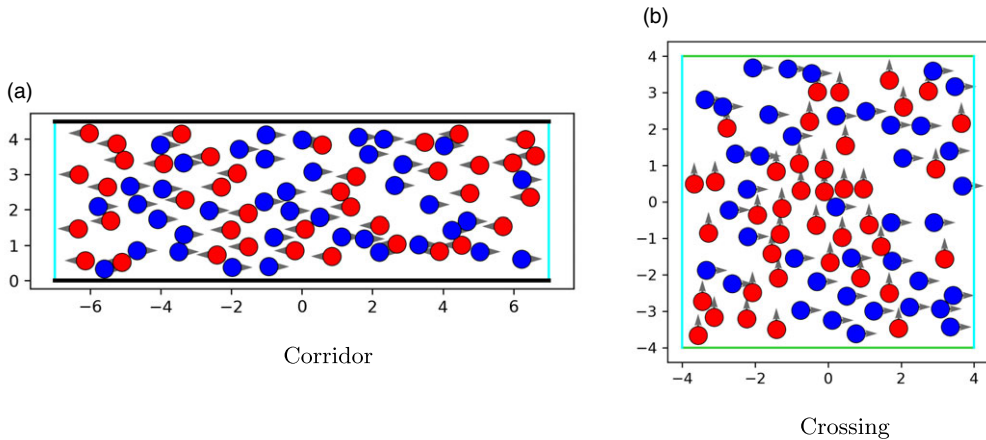


Figure 1. Initial positions and initial velocity vectors for the two scenarios with parameters $N = 80$, $N_{blue} = 40$, $N_{red} = 40$, $d = 0.2$.

2.2.1. Influence of the body size

In the following, we provide some numerical results showing the influence of the body size on lane formation in the corridor and crossing scenario, respectively. The first experiment simulates the movement of two oncoming streams of pedestrians along a spacious corridor. The group of blue agents moves from left to right with desired velocity $w_{blue} = (0.7, 0)^T$, whereas the red group of agents moves from right to left with desired velocity $w_{red} = (-0.7, 0)^T$. We consider N_{blue} blue and N_{red} red agents. Hence, the total number of pedestrians in the corridor is $N = N_{blue} + N_{red}$. The initial positions of the pedestrians $x(0) = \mathbf{x}_0$ and their initial velocities $v(0) = \mathbf{v}_0$ are illustrated in Figure 1a.

To assure that the pedestrians do not leave the scenario, we add reflective and periodic boundary conditions. In the corridor case the black lines (top and bottom) in Figure 1a show reflective boundaries. We model the avoidance of wall contact, by reflecting the velocity vector of an agent that would step outside of the domain in the next time step. The behavior of reflection from the wall is the same as in [34]. The light blue lines illustrate periodic boundaries. Blue agents leaving the domain at the boundary on the right, enter again from the left. Analogously for the red agents. With the periodic boundary condition, the number of agents in the system remains constant.

In the second scenario, we consider two groups of pedestrians at a crossing. Here, the blue group of pedestrians moves from left to right with desired velocity $w_{blue} = (0.7, 0)^T$, and the red group of pedestrians move from bottom to top with desired velocity $w_{red} = (0, 0.7)^T$. In total, there are $N = N_{blue} + N_{red}$ pedestrians. The initial positions and initial velocity vectors are presented in Figure 1b. There, light blue and green lines represent periodic boundaries for blue and red agents, respectively. Altogether, this results in Algorithm 1 for the state system (2.1).

Algorithm 1. state system

Input: initial positions \mathbf{x}_0 and velocities \mathbf{v}_0 of pedestrians, model parameters
while $t < T$ **do**
 update the positions and velocities of the agents using (2.6)
 check reflective boundary condition: if an agent is about to leave through a reflective border, its velocity is changed in the direction pointing inwards
 check periodic boundary condition: if an agent leaves through a periodic boundary, its position is adjusted due to the periodic boundary conditions
 $t = t + \Delta t$
end
Result: trajectories and velocities at the time steps

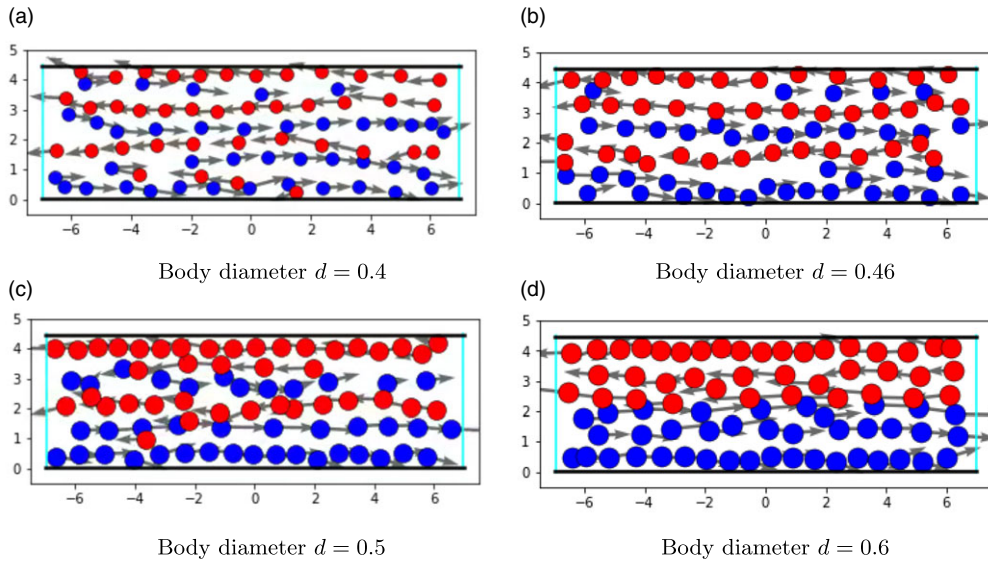


Figure 2. Simulation results in the corridor by different body sizes of pedestrians at time $T = 35$. In each simulation, we fix parameters: $A = 5$, $R = 20$, $a = 2$, $r = 0.5$, $\lambda = 0.25$. Desired velocities for red and blue agents are $w_{red} = (-0.7, 0)^T$ and $w_{blue} = (0.7, 0)^T$, respectively. The time step in the Leap-Frog Scheme is $\Delta t = 0.00625$.

2.2.2. Study for different body sizes

To analyze the simulation results for different body sizes, we fix values for the force parameters and desired velocities of each pedestrian. The parameters are chosen to satisfy the stability ranges of the interaction force discussed in [14]. In fact, in the range, $R/A > 1$ and $r/a < 1$ the interaction force K is repulsive in the short range and attractive in the long range. This allows the distance between pedestrians to be maintained. Even if we include body size into the interaction force, it remains repulsive in a short-range and attractive in a long range. The strength of attraction and repulsion forces between agents ensures that they remain at a comfortable distance from each other and avoid collisions, regardless of their color. However, the overall direction of movement of the agents is primarily determined by their desired velocity w , which varies depending on the type of directed motion, such as bi-directional movement, crossing motion, crossing at an angle, and so on.

Figure 2 shows the simulation results of the corridor scenario for different body sizes. The results indicate a relation between the body size and the number of lanes formed. The smaller the body size, the more lanes are obtained. The parameters used for the simulation are reported in the caption of the figure. Similar results are found for the crossing scenario in Figure 3. Again, the smaller the body size, the more lanes are formed.

In all simulations, we see the formation of so-called traffic lanes. This formation seems to be independent of the choice of the random initial positions and velocities. It is interesting to note that even though every pedestrian is guided by simple rules for movement and interaction, phenomena arise that go beyond the behavior of single pedestrians. Such phenomena of self-organization are manifested in many multi-agent systems [19]. They were reported in many articles concerning the movement of pedestrian flows [31, 36], which speaks in favor of the proposed model. Moreover, we want to emphasize that not only the body size can influence the number of lanes. In fact, the choice of the width of the corridor, the number of agents, and attraction and repulsion force parameters can change the formation of lanes as well. For the force parameters, this is shown exemplarily in Figure 4.

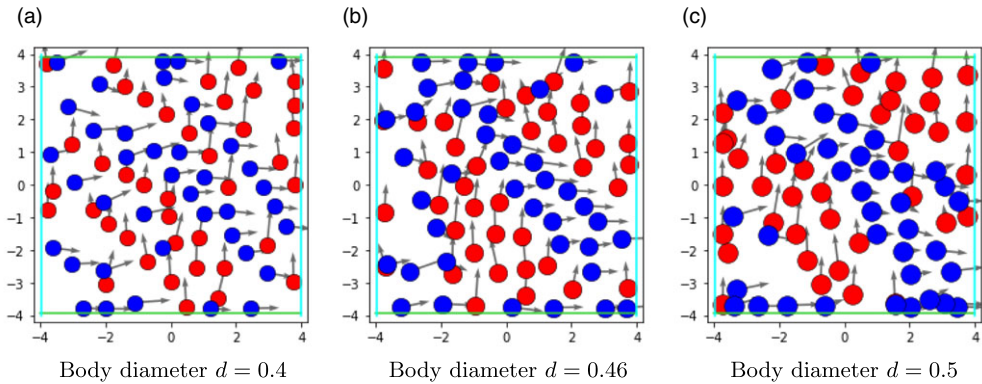


Figure 3. Simulation results at the crossing by different body sizes of pedestrians at time $T = 35$. In each simulation, we fix parameters: $A = 5, R = 20, a = 2, r = 0.5, \lambda = 0.25$. Desired velocities for red and blue agents are $\vec{w}_{red} = (0, 0.7)^T$ and $\vec{w}_{blue} = (0.7, 0)^T$ respectively. The time step in the leap-frog scheme is $\Delta t = 0.00625$.

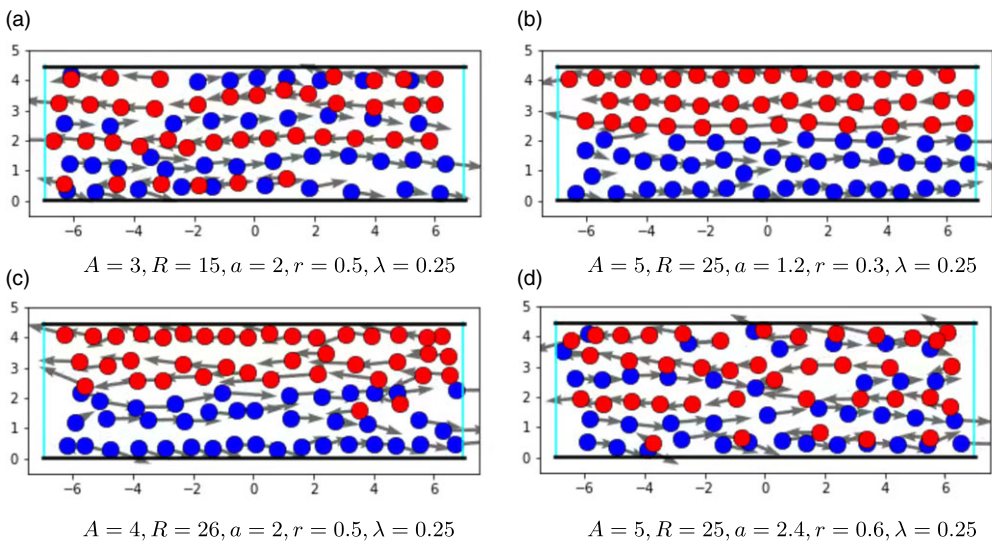


Figure 4. Simulation results in the corridor by different force parameters at time $T = 35$. In each simulation, the body diameter of the agents is fixed: $d = 0.5$. The time step in the leap-frog scheme is $\Delta t = 0.00625$.

As the body size, the ratio of repulsion and attraction amplitudes, and the size of the corridor have similar effects in terms of volume exclusion, we suspect from these studies that volume exclusion is the main driver of the lane formation process.

2.2.3. Fundamental diagram

Often fundamental diagrams are employed to analyze crowd motion models [27, 36]. Main objective is the relationship of speed and density [29, 32, 36] which we study for the bi-directional and crossing flow simulated with the model described in Section 2. For the illustration, we take the corridor with 17m length and 4m width, see Figure 5 for the bi-directional flow and the intersection of corridors with

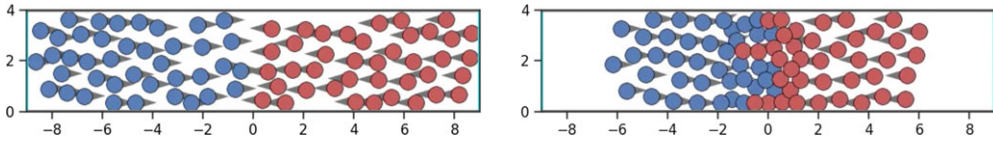


Figure 5. Bidirectional pedestrian flow at time $T = 0$ and $T = 5$ s.

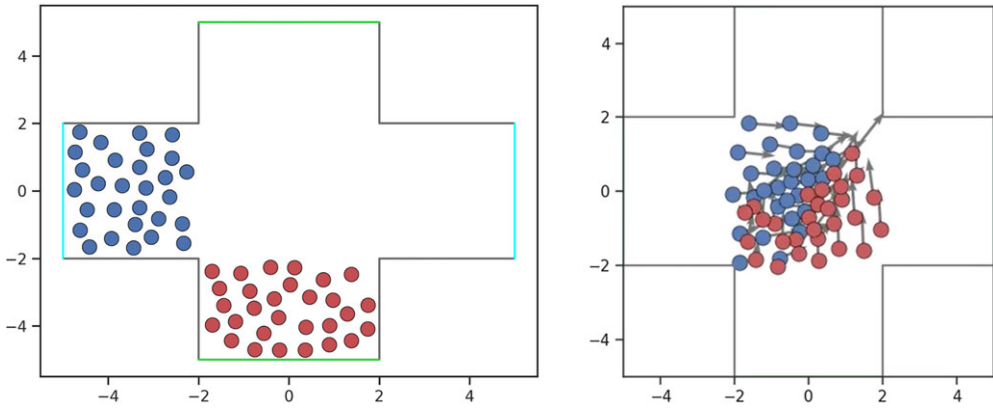


Figure 6. Crossing pedestrian flow at time $T = 0$ and $T = 4$ s.

a width of $4m$ and a length of $10m$ for the crossing flow, see Figure 6. The density approximation is realized with the help of Voronoi diagrams as proposed in [6, 32]. Initially, agents move with their desired walking speed until they slow down (or speed up) due to interaction forces. Most interactions take place in the center of the domain, it is therefore the focus of our interest.

To approximate the density we construct Voronoi cells based on the positions of agents. The Voronoi diagram allows separating the area into computational grids (polygons) based on the triangulation of the computational area, in particular, the Delaunay triangulation [30]. Every point on the Voronoi diagram has a region that is closer to it than any other [32]. In Figure 7a we see Voronoi cells of all agents. Figure 7b shows the same cells, but only the ones in the region of interest $\Omega = [-2, 2] \times [0, 4]$. For the crossing flow, the region of interest is $\Omega = [-2, 2] \times [-2, 2]$. Outside of the domain Ω agents are involved in fewer interactions and walk approximately with their desired velocities. The model parameters in these simulation are set $\lambda = -0.07$, $A = 6$, $R = 33$, $a = 1$, $r = 0.3$, $d = 0.46$. In the preceding studies we focussed on the effect of the body size on the model dynamics. The magnitude of the parameter values was incidental. Now, we investigate whether the model is able to reproduce the fundamental diagram, which was found in many experimental studies. We therefore already use here the parameters obtained from the calibration in Section 3.

We employ the method proposed in [6, 32] to compute the density as follows

$$\rho_{xy}(t) = \begin{cases} 1/A_i(t) & \text{if } (x, y) \in A_i(t), \\ 0 & \text{else} \end{cases}$$

where $A_i(t)$ gives the Voronoi cell area for agent i , and ρ_{xy} is the density distribution of the space. Voronoi cells are computed using the Python, Scipy module, with the Voronoi and ConvexHull methods. Note that the velocities are given by the integration of (2.1).

In Figure 9 we illustrate the relationship between density and speed at different time points. In the beginning, the agents move with the desired speed in regions with lower density leading to weak interaction forces. When the two groups meet, we see that as the density increases and at the same time the speed decreases as expected. Body size seems to have a minor influence on this effect.

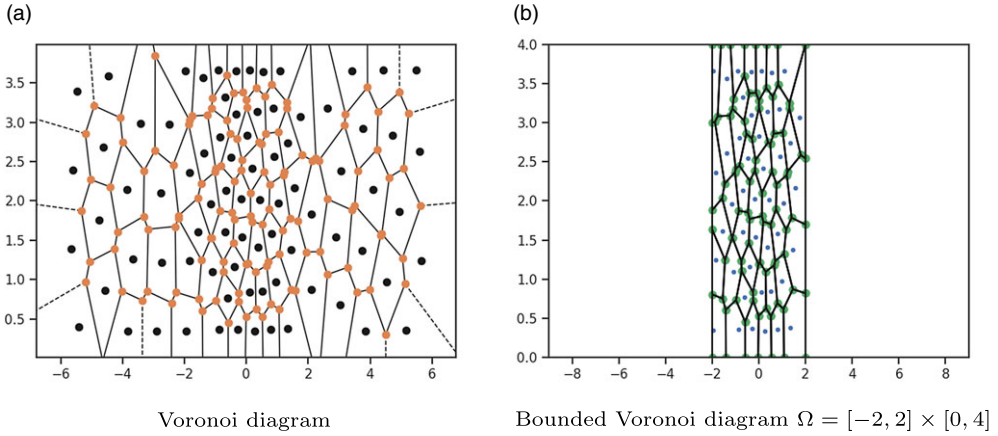


Figure 7. Voronoi diagrams of bidirectional pedestrian flow at $T = 5$ s.

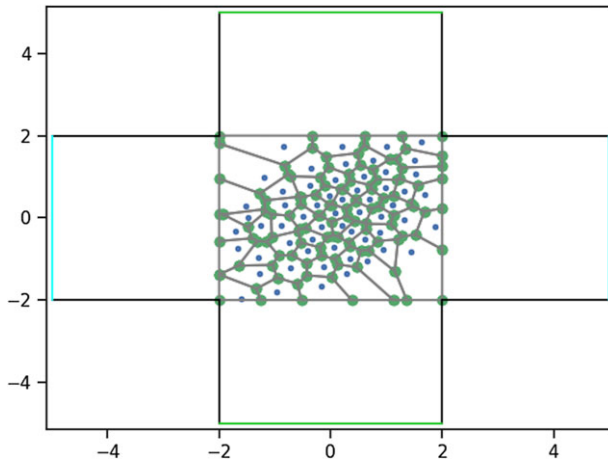


Figure 8. Bounded Voronoi diagram on $\Omega = [-2, 2] \times [-2, 2]$ of crossing pedestrian flow at $T = 5$ s.

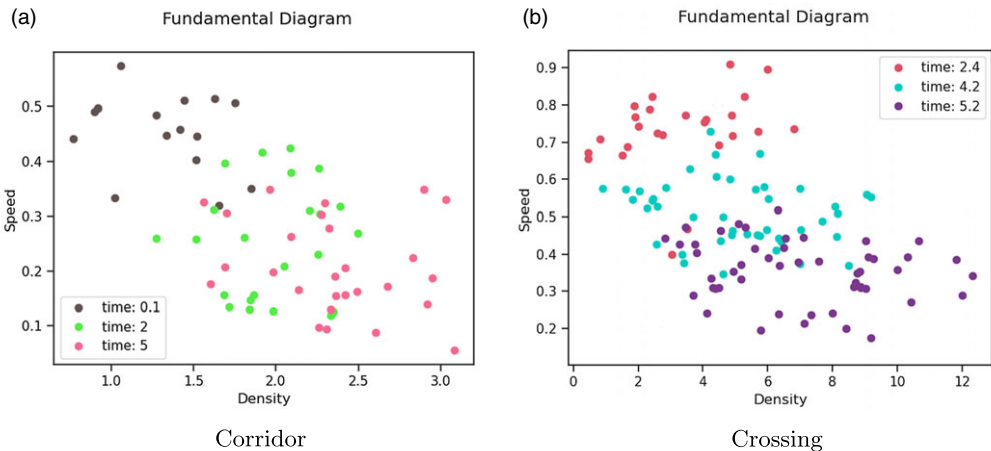


Figure 9. Fundamental diagrams of pedestrian flow in the corridor and at the crossing scenario.

Table 1. Model parameters

λ	Scaling factor (collision angle)	τ	Relaxation parameter (desired velocity)
A	Attractive force amplitude	R	Repulsive force amplitude
a	Attractive force range	r	Repulsive force range
w	Desired velocity	d	Body size of pedestrian

This ends our analysis and numerical study of the model. In the following sections, we are concerned with its parameter calibration based on real data [25]. We begin with the statement of the calibration problem and analyze its well-posedness.

3. Analysis of the parameter calibration problem

In this section, we state the parameter calibration problem and analyze the existence of minimizers. Then we lay the theoretical ground for the formulation of a gradient-descent method by deriving the first-order optimality conditions, the existence of adjoint states, and finally the identification of the gradient of the reduced cost functional. This prepares the formulation of a gradient descent-based calibration algorithm that will be employed in the numerical section to fit the parameters to real data from the BaSiGo experiment carried out in Düsseldorf in 2013 [6, 25] in the next section.

Let us begin with the statement of the parameter calibration problem. Table 1 shows all model parameters. For the calibration, we focus on the scaling factor λ involved in the collision avoidance process and the force strengths A and R and body size of agents d . We do not include the parameters a and r in the calibration problem, as the forces are driven by the ratios A/a as well as R/r , hence increasing A, R or decreasing a, r has similar effects. Numerical tests that included r, a in the calibration led to very unstable gradient behaviour, which we understood as a sign of overparameterization. We therefore fix r and a for the following considerations.

For fixed $1 \gg \epsilon > 0$, we define the set of admissible parameters

$$U_{\text{ad}} = \left\{ (\lambda, A, R, d) \in [-1 + \epsilon, 1 - \epsilon] \times [0, A_{\text{max}}] \times [0, R_{\text{max}}] \times [0, d_{\text{max}}] \right\}$$

and we want to find $u := (\lambda, A, R, d) \in U_{\text{ad}}$ such that the model trajectories fit the real trajectories best. We thus consider the cost functional

$$J(x, u) := \int_0^T \frac{\sigma_1}{2N} \sum_{i=1}^N \|x_i(t) - x_i^{\text{data}}(t)\|_2^2 dt + \frac{\sigma_2}{2} \|u - u_{\text{ref}}\|_2^2$$

with x^{data} given trajectory data from experiments. The first term measures the distance of the trajectories resulting from the model to the real trajectories from the data. The second term of the cost functional penalizes the distance of the parameters to some given reference parameters u_{ref} . In case there are no reference values available, we set $\sigma_2 = 0$.

To study well-posedness, we use the following notion of optimality:

Definition 3.1. We call $u \in U_{\text{ad}}$ optimal, if it is a solution to the optimization problem

$$\min_{(x,u) \in H^1([0,T], \mathbb{R}^{2N}) \times U_{\text{ad}}} J(x, u) \quad \text{subject to (2.1)}. \tag{P}$$

Note that the calibration problem is constrained by the ODE system without boundary conditions. The boundary conditions are only incorporated in the numerical simulations to reflect the domain of the experiments appropriately.

In the following, we consider the spaces Y and U given by

$$Y = [H^1([0, T], \mathbb{R}^{2N}) \times H^1([0, T], \mathbb{R}^{2N})], \quad U = [-1, 1] \times \mathbb{R} \times \mathbb{R} \times \mathbb{R}.$$

Note that both, Y and U are Hilbert spaces, and $U_{ad} \subset U$ is closed. For notational convenience we define the state vector $y = (x, v) \in Y$ and the state operator

$$e : Y \times U \rightarrow Z^*, \quad e(y, u) = \begin{pmatrix} \frac{d}{dt}y - F(y, u) \\ y(0) - y_0 \end{pmatrix}, \tag{3.1}$$

where $F(y, u)$ is the vector containing the right-hand side of (2.1a) and (2.1b), respectively.

3.1. Well-posedness of the parameter calibration problem

The proof of the existence of an optimal parameter set for the calibration problem will be based on the following Lemmata which are concerned with the boundedness of the states with respect to the control parameters and the weak continuity of the state operator e .

Lemma 3 (Boundedness). *Let $w_i \in C([0, T], \mathbb{R}^2)$ for all $i = 1, \dots, N$ and Assumption 1 hold and suppose the interaction forces are given by (2.5). For given $u \in U_{ad}$ there exists a constant $C > 0$ depending only on the body size d and*

$$\bar{w} := \max_i \sup_{t \in [0, T]} \|w_i(t)\|,$$

such that the solution $y \in Y$ with $e(y, u) = 0$ satisfies

$$\|y\|_Y \leq C(1 + \|u\|).$$

Proof. We begin with the estimate for $\|y\|_{L^2([0, T], \mathbb{R}^2)}$. It holds

$$\|x(t)\|^2 \leq 2 \|x_0\|^2 + 2 \int_0^t \|v(s)\|^2 ds, \quad \|v(t)\|^2 \leq 2 \|v_0\|^2 + 2 \int_0^t \|w(s) - v(s)\|^2 + Ce^d \|u\|^2 ds.$$

Hence, using Gronwall Lemma we obtain $\|y(t)\|^2 \leq \|u\|^2 e^{Ct}$ and integration over $[0, T]$ yields $\|y\|_{L^2(0, T, \mathbb{R}^2)} \leq C_1 \|u\|$. For the time derivatives, we find

$$\left\| \frac{d}{dt}x(t) \right\|^2 \leq \|v(s)\|^2, \quad \left\| \frac{d}{dt}v(t) \right\|^2 \leq 2 \|w(s) - v(s)\|^2 + 2Ce^d \|u\|^2.$$

Integration over $[0, T]$ leads to $\left\| \frac{d}{dt}y \right\|_{L^2(0, T, \mathbb{R}^2)} \leq C(1 + \|u\|)$. The two estimates together give the result. □

Lemma 4. *Let Assumption 1 hold. The state operator*

$$e : Y \times U \rightarrow Z^*, \quad e(y, u) = \begin{pmatrix} \frac{d}{dt}y - F(y, u) \\ y(0) - y_0 \end{pmatrix}$$

is weakly continuous.

Proof. Let $(y_k, u_k) \rightharpoonup (\hat{y}, \hat{u})$ as $k \rightarrow \infty$. We need to show that $e(y_k, u_k) \rightharpoonup e(\hat{y}, \hat{u})$ as $k \rightarrow \infty$, which can be reformulated as follows.

For any given test function $\varphi \in C_c^1(Z)$, we need to obtain the following convergence property

$$\lim_{k \rightarrow \infty} \int \langle e(y_k, u_k), \varphi \rangle dt \rightarrow \int \langle e(\hat{y}, \hat{u}), \varphi \rangle dt.$$

We estimate

$$\begin{aligned} & \lim_{k \rightarrow \infty} \int_0^t \left[\frac{d}{dt} y_k - \frac{d}{dt} \hat{y} - F(y_k, u_k) + F(\hat{y}, \hat{u}) \right] \varphi dt \\ & = : \lim_{k \rightarrow \infty} \int_0^t \left[\frac{d}{dt} y_k - \frac{d}{dt} \hat{y} \right] \varphi dt \\ & \quad + \lim_{k \rightarrow \infty} \int_0^t \left[\tau(-\hat{y} + v_k) - M_{\hat{\lambda}}(\hat{y})K_{\hat{A}, \hat{R}}(\hat{d}, \hat{x}) + M_{\lambda_k}(v_k)K_{A_k, R_k}(d_k, x_k) \right] \varphi dt \\ & = \lim_{k \rightarrow \infty} \int_0^t \left[\frac{d}{dt} y_k - \frac{d}{dt} \hat{y} \right] \varphi dt + \lim_{k \rightarrow \infty} \int_0^t \left[\begin{matrix} -(v_k - \hat{y}) \\ \tau(v_k - \hat{y}) \end{matrix} \right] \varphi dt \\ & \quad + \lim_{k \rightarrow \infty} \int_0^t \left[\begin{matrix} 0 \\ (M_{\lambda_k}(v_k) - M_{\hat{\lambda}}(\hat{y}))K_{A_k, R_k}(d_k, x_k) + M_{\hat{\lambda}}(\hat{y})(K_{A_k, R_k}(d_k, x_k) - K_{\hat{A}, \hat{R}}(\hat{d}, \hat{x})) \end{matrix} \right] \varphi dt \end{aligned}$$

Clearly, the first and second integral tends to zero for $k \rightarrow \infty$ by the weak convergence of y_k to $y \in Y$.

Let us consider the third integral separately. We can estimate the first summand as

$$\begin{aligned} & \|M_{\lambda_k}(v_k) - M_{\hat{\lambda}}(\hat{y})K_{A_k, R_k}(d_k, x_k)\| \\ & \leq \|M_{\lambda_k}(v_k) - M_{\lambda_k}(\hat{y}) + M_{\lambda_k}(\hat{y}) - M_{\hat{\lambda}}(\hat{y})\| \cdot \|K_{A_k, R_k}(d_k, x_k)\| \\ & \leq \|M_{\lambda_k}(v_k) - M_{\lambda_k}(\hat{y})\| \cdot \|K_{A_k, R_k}(d_k, x_k)\| + \|M_{\lambda_k}(\hat{y}) - M_{\hat{\lambda}}(\hat{y})\| \cdot \|K_{A_k, R_k}(d_k, x_k)\| \\ & \leq L_{\lambda_k} \|v_k - \hat{y}\| \cdot \|K_{A_k, R_k}(d_k, x_k)\| + \|M_{\lambda_k}(\hat{y}) - M_{\hat{\lambda}}(\hat{y})\| \cdot \|K_{A_k, R_k}(d_k, x_k)\|. \end{aligned}$$

From Lemma 1 we have L_{λ_k} . We recall that K is Lipschitz continuous by Assumption 1, and that v_k is in $H^1([0, T], \mathbb{R}^D) \hookrightarrow L^2([0, T], \mathbb{R}^D)$. By weak convergence of $v_k \rightharpoonup \hat{v}$ and by this compact embedding, the first norm tends to zero as $k \rightarrow \infty$. The $\|M_{\lambda_k}(\hat{y}) - M_{\hat{\lambda}}(\hat{y})\|$ tends to zero by continuity of the map $\lambda \mapsto M_{\lambda}(v)$.

Analogously we obtain the convergence of the term $\|M_{\hat{\lambda}}(\hat{y})(K_{A_k, R_k}(d_k, x_k) - K_{\hat{A}, \hat{R}}(\hat{d}, \hat{x}))\|$. Then, using continuity of the interaction force with respect to A, R and d , and weak convergence of $x_k \rightharpoonup \hat{x}$ and compact embedding x_k in $H^1([0, T], \mathbb{R}^D) \hookrightarrow L^2([0, T], \mathbb{R}^D)$, we obtain the desired result. \square

Theorem 3.2. *There exists at least one solution $(y^*, u^*) \in Y \times U_{ad}$ to (P).*

Proof. The cost functional J is bounded from below and the state system is well-posed, so there exists

$$m = \inf_{(y, u) \in Y \times U_{ad}} J(y, u).$$

Let $(u_k) \in U_{ad}$ be a minimizing sequence. The sequence $(u_k) \subset U_{ad}$ is bounded, and by the reflexivity of U it has a weakly convergent subsequence (not relabeled) with limit \hat{u} . By Lemma 3 we obtain the boundedness of (y_k) and, again by reflexivity, the existence of \hat{y} such that

$$u_k \rightharpoonup \hat{u} \text{ in } U_{ad} \quad \text{and} \quad y_k \rightharpoonup \hat{y} \text{ in } Y \text{ as } k \rightarrow \infty. \tag{3.2}$$

The weak continuity of $e(y, u)$ shown in Lemma 4 implies

$$\|e(\hat{y}, \hat{u})\| \leq \liminf_{k \rightarrow \infty} \|e(y_k, u_k)\| = 0.$$

Hence \hat{y} is the solution to the state equation with parameters \hat{u} . By the weak lower semi-continuity of the norm, we obtain

$$J(\hat{y}, \hat{u}) \leq \liminf_{k \rightarrow \infty} J(y_k, u_k) = m,$$

which allows us to conclude that (\hat{y}, \hat{u}) is a minimizer of (P). \square

Remark 4. Because of the non-linearity of the state equation, we cannot expect the uniqueness of the optimal control.

Having the existence of an optimal solution, we proceed with the derivation of the first-order necessary conditions, which will form the basis of the identification algorithm.

3.2. First-order necessary conditions

We introduce the dual pairing

$$\langle e(y, u), (\xi, \eta) \rangle_{Z, Z^*} = \int_0^T \left(\frac{d}{dt}y - F(y, u) \right) \cdot \xi(t)dt + (y(0) - y_0) \cdot \eta, \tag{3.3}$$

where ξ, η are the Lagrange multipliers in the Banach space

$$Z = [L^2([0, T], \mathbb{R}^{2N}) \times L^2([0, T], \mathbb{R}^{2N}) \times \mathbb{R}^{2N} \times \mathbb{R}^{2N}]$$

that represent the space of the adjoint states. Here, $\xi = (\xi_x, \xi_v), \eta = (\eta_x, \eta_v)$ and $(\xi, \eta) \in Z$. The space of states Y and the set of controls U are defined at the beginning of Section 3.

We formally derive the first-order optimality system with the help of the Lagrangian corresponding to the constrained parameter calibration problem given by

$$\mathcal{L} : Y \times U \times Z \rightarrow \mathbb{R}, \quad \mathcal{L}(y, u, \xi, \eta) = J(y, u) + \langle e(y, u), (\xi, \eta) \rangle_{Z, Z^*}. \tag{3.4}$$

3.2.1. Derivation of the adjoint system and the optimality condition

Solving $d\mathcal{L}(y, u, \xi, \eta) = 0$ yields the first-order optimality condition [21]. We begin with the computation of the directional derivatives of (3.4) with respect to the state y . For notational convenience we consider x and v separately.

First, we obtain Gâteaux derivatives of the objective function

$$d_{x_i}J(y, u)[h_{x_i}] = \frac{\sigma_1}{N} \int_0^T (x_i(t) - x_i^{data}(t))h_{x_i}(t)dt, \quad d_{v_i}J(y, u)[h_{v_i}] = 0.$$

Now we derive the directional derivatives with respect to the positions of agents for the second part of the Lagrange functional

$$d_{x_i} \langle e(y, u), (\xi, \eta) \rangle [h_{x_i}] = d_{x_i} \langle e_i(y, u), (\xi, \eta)_i \rangle [h_{x_i}] + \sum_{\substack{j=1 \\ j \neq i}}^N d_{x_i} \langle e_j(y, u), (\xi, \eta)_j \rangle [h_{x_i}].$$

Using integration by parts, we obtain

$$\begin{aligned} d_{x_i} \langle e(y, u), (\xi, \eta) \rangle [h_{x_i}] &= \int_0^T h'_{x_i} \xi_{x_i}(t) + \frac{1}{N} h_{x_i} \sum_{\substack{j=1 \\ j \neq i}}^N \left(M(v_i, v_j) d_{x_i} K(x_i, x_j) \right)^T \xi_{v_i}(t) dt \\ &+ \int_0^T h_{x_i} \frac{1}{N} \sum_{\substack{j=1 \\ j \neq i}}^N \left[M(v_j, v_i) d_{x_i} K(x_j, x_i) \right]^T \xi_{v_j}(t) dt + h_{x_i}(0) \eta_{x_i}. \end{aligned} \tag{3.5}$$

Since $M(v_i, v_j) = M(v_j, v_i)$ is symmetric and $d_{x_i}K(x_i, x_j) = -d_{x_j}K(x_j, x_i)$ by the radial symmetry of K , we obtain

$$\begin{aligned} d_{x_i} \langle e(y, u), (\xi, \eta) \rangle [h_{x_i}] &= h_{x_i}(0) \eta_{x_i} \\ &+ \int_0^T \left[h'_{x_i}(t) \xi_{x_i}(t) + h_{x_i}(t) \frac{1}{N} \sum_{\substack{j=1 \\ j \neq i}}^N \left(M(v_i, v_j) d_{x_i} K(x_i, x_j) \right)^T (\xi_{v_i}(t) - \xi_{v_j}(t)) \right] dt. \end{aligned}$$

Similarly, we obtain the derivative in direction h_{v_i} . Indeed, by using integration by parts we find

$$d_{v_i} \langle e(y, u), (\xi, \eta) \rangle [h_{v_i}] = d_{v_i} \langle e_i(y, u), (\xi, \eta)_i \rangle [h_{v_i}] + \sum_{\substack{j=1 \\ j \neq i}}^N d_{v_i} \langle e_j(y, u), (\xi, \eta)_j \rangle [h_{v_i}]. \tag{3.6}$$

To simplify the notation we introduce the operator $d_{v_i} M^*(v_i, v_j)$ resulting from matrix reformulations, see Appendix B for more details. We get

$$\begin{aligned} d_{v_i} \langle e(y, u), (\xi, \eta) \rangle [h_{v_i}] &= h_{v_i}(0) \eta_{v_i} + \int_0^T h'_{v_i} \xi_{v_i}(t) - h_{v_i} \xi_{x_i}(t) + \tau h_{v_i} \xi_{v_i}(t) dt \\ &+ \int_0^T h_{v_i} \left[\frac{1}{N} \sum_{\substack{j=1 \\ j \neq i}}^N d_{v_i} M^*(v_i, v_j) K(x_i, x_j) \xi_{v_i}(t) \right. \\ &\left. + \frac{1}{N} \sum_{\substack{j=1 \\ j \neq i}}^N d_{v_i} M^*(v_j, v_i) K(x_j, x_i) \xi_{v_j}(t) \right] dt. \end{aligned}$$

Moreover, the directional derivatives with respect to the control are given by

$$d_u J(y, u)[h_u] = \sigma_2(u - u_{ref})h_u, \tag{3.7}$$

$$d_u \langle e(y, u), (\xi, \eta) \rangle [h_u] = - \int_0^T h_u \sum_i [d_u F(y_i, u)]^T \xi_{v_i}(t) dt. \tag{3.8}$$

Remark 5. For the specific choice (2.5), as we choose the interaction forces for the numerical results, the derivatives read

$$\frac{dF(y_i, u)}{d\lambda} = \begin{cases} -\frac{1}{N} \sum_{j \neq i} \begin{pmatrix} -\sin \alpha_{ij} & -\cos \alpha_{ij} \\ \cos \alpha_{ij} & -\sin \alpha_{ij} \end{pmatrix} \arccos \frac{v_i \cdot v_j}{\|v_i\| \|v_j\|} \cdot K(x_i, x_j), & \text{for } v_i, v_j \neq 0, \\ \vec{0}_{2 \times 1}, & \text{else} \end{cases},$$

$$\frac{dF(y_i, u)}{dA} = -\frac{1}{a \cdot N} \sum_{j \neq i} M(v_i, v_j) \cdot e^{-\frac{d-\|x_i-x_j\|}{a}} \cdot \frac{x_i - x_j}{\|x_i - x_j\|},$$

$$\frac{dF(y_i, u)}{dR} = \frac{1}{r \cdot N} \sum_{j \neq i} M(v_i, v_j) \cdot e^{-\frac{d-\|x_i-x_j\|}{r}} \cdot \frac{x_i - x_j}{\|x_i - x_j\|}.$$

$$\frac{dF(y_i, u)}{dd} = -\frac{1}{N} \sum_{j \neq i} M(v_i, v_j) \cdot \left(\frac{A}{a^2} e^{-\frac{d-\|x_i-x_j\|}{a}} - \frac{R}{r^2} e^{-\frac{d-\|x_i-x_j\|}{r}} \right) \cdot \frac{x_i - x_j}{\|x_i - x_j\|}.$$

3.2.2. Existence of adjoint states

The proof of the existence of adjoint states is based on Corollary 1.3 in [21], which we give in Appendix C for completeness.

Theorem 3.3. Let $w_i \in C([0, T], \mathbb{R}^2), i = 1, \dots, N$ be given, the Assumption 1 and 2 hold and $u^* \in U_{ad}$ be an optimal solution of Problem (P) and let $y^* \in Y$ such that $e(y^*, u^*) = 0$. Then there exist an adjoint state $p^* = (\xi^*, \eta^*) \in Z^*$ such that the following optimality conditions hold

$$\begin{aligned} \langle e(y^*, u^*), p \rangle_{Z, Z^*} &= 0 & \forall p \in Z^*, \\ \langle L_y(y^*, u^*, p^*), h \rangle_{Y^*, Y} &= 0 & \forall h \in Y^*, \\ u^* \in U_{ad}, \langle L_u(y^*, u^*, p^*), u - u^* \rangle_{U^*, U} &\geq 0, & \forall u \in U_{ad}. \end{aligned}$$

Proof. We check requirements given in Assumption 3 (see Appendix C):

- (A1) $U_{ad} = [-1 + \epsilon, 1 - \epsilon] \times [0, A_{max}] \times [0, R_{max}] \times [0, d_{max}] \subset U$ is nonempty, convex and closed.
- (A2) We first note that U, Y, Z are Banach spaces. Further, J is of tracking type and therefore Fréchet differentiable [35]. We are left to show Fréchet differentiability of the state operator $e(y, u) : Y \times U \rightarrow Z$. In Section 3.2.1 we computed the first variations of $e(y, u)$ with respect to y and u by

$$d_y \langle e(y, u)[(h_y)], (\xi, \eta) \rangle = \lim_{\epsilon \rightarrow 0} \frac{1}{\epsilon} \langle e(y + \epsilon h_y, u) - e(y, u), (\xi, \eta) \rangle,$$

and obtained (3.5) and (3.6). There, the continuity of linear terms follows directly from the definition. We show the continuity for nonlinear terms:

Let, the sequence $y_k := (x_k, v_k) \subset Y$ have the limit y_k such that $x_k \rightarrow x$ and $v_k \rightarrow v$ as $k \rightarrow \infty$. Using Assumption 2 we show the continuity of nonlinear terms in (3.5). Indeed, for the i -th component, it holds

$$\begin{aligned} & \int_0^T \frac{1}{N} h_{x_i} \sum_{\substack{j=1 \\ j \neq i}}^N \left(M(v_{k_i}, v_{k_j}) d_{x_i} K(x_{k_i}, x_{k_j}) - M(v_i, v_j) d_{x_i} K(x_i, x_j) \right)^T (\xi_{v_i} - \xi_{v_j}) dt \\ & \leq L_1 \cdot L_2 \int_0^T \frac{1}{N} \sum_{\substack{j=1 \\ j \neq i}}^N \left(\|v_{k_i} - v_i\| + \|v_{k_j} - v_j\| + \|x_{k_i} - x_i\| \right. \\ & \quad \left. + \|x_{k_j} - x_j\| \right) \|\xi_{v_i} - \xi_{v_j}\| dt \leq L_1 \cdot L_2 \int_0^T \|y_{k_i} - y_i\| \|\xi_{v_i} - \xi_{v_j}\| dt, \end{aligned}$$

where L_1, L_2 are Lipschitz constants. Analogously, we show the continuity for the nonlinear terms of (3.6) using Assumption 1 and Appendix A. These, yield the continuity of the state operator $e(y, u)$ by y . The continuity with respect to u , can be concluded from the continuity of the $\lambda \mapsto M$, and the linearity of the force term with respect to A and R . Altogether, this proves the continuity of e as desired.

- (A3) By Lemma 2 the state system $e(y, u) = 0$ has a unique solution $y = y(u) \in Y$ for all $u \in V \subset U$ a neighborhood of U_{ad} .
- (A4) We have to show that $e_y(y(u), u) \in \mathcal{L}(Y, Z)$ has a bounded inverse for all $u \in V \supset U_{ad}$. We can write $d_y e(y, u)[h]$ in general form

$$d_y e(y, u)[h] = \frac{d}{dt} h(t) + d_y F(y, u)h(t),$$

where $d_y F(y, u)$ is integrable in t over every finite interval $I \subset [0, T]$ thanks to Assumption 2. We consider $d_y e(y, u)[h] = r$ for arbitrary $r \in Z^*$. By Carathéodory's existence theorem, we get for every $r \in Z^*$ a unique solution [16], namely $h = d_y e(y, u)^{-1}[r]$. Using $d_y e(y, u)$, we obtain

$$\|h(t)\| \leq \|h(0)\| + \int_0^t \|r(s)\| ds + C \exp \left(\int_0^t \|h(s)\| ds \right), \quad t \in [0, T]$$

and with the help of Gronwall's Lemma, we get the boundedness of the inverse of $d_y e(y(u), u)$.

Altogether, the requirements of Assumption 3 are satisfied, and thus Proposition 1 (see Appendix C) yields the existence of an adjoint state. □

Remark 6. Note that assuming more regularity of the adjoint states, for instance, $Z = Y$, we may establish the strong formulation of the adjoint system given by

$$\begin{aligned} \xi'_{x_i}(t) &= \frac{\sigma_1}{N}(x_i(t) - x_i^{data}(t)) + \frac{1}{N} \sum_{\substack{j=1 \\ j \neq i}}^N \left(M(v_i, v_j) d_{x_i} K(x_i, x_j) \right)^T (\xi_{v_i}(t) - \xi_{v_j}(t)), \\ \xi'_{v_i}(t) &= -\xi_{x_i}(t) + \tau \xi_{v_i}(t) - \frac{1}{N} \sum_{\substack{j=1 \\ j \neq i}}^N d_{v_i} M^*(v_i, v_j) K(x_i, x_j) \xi_{v_j}(t) \\ &\quad - \frac{1}{N} \sum_{\substack{j=1 \\ j \neq i}}^N d_{v_i} M^*(v_j(t), v_i(t)) K(x_j(t), x_i(t)) \xi_{v_j}(t) \end{aligned} \tag{3.9}$$

supplemented with the terminal conditions $\xi_{x_i}(T) = 0, \xi_{v_i}(T) = 0$. The strong form will be employed for the numerical results in Section 4.

3.2.3. Gradient of the reduced cost functional

To minimize the objective function we aim to apply a gradient descent algorithm. In order to determine the gradient, we define the control-to-state operator $\mathcal{F} : U \rightarrow Y$, and introduce the reduced cost functional as $\hat{J}(u) := J(\mathcal{F}(u), u)$. Using $e(y, u) = 0$, we obtain

$$0 = d_y e(\mathcal{F}(u), u)[d\mathcal{F}(u)] + d_u e(\mathcal{F}(u), u).$$

Taking the derivative of the Lagrangian with respect to the state y , we get

$$d_y e(y, u)^* \xi = -d_y J(y, u).$$

With these, we can compute the Gâteaux derivative of the reduced cost functional in the direction $h_u \in U$, and obtain

$$\begin{aligned} d\hat{J}(u)[h_u] &= \langle d_y J(y, u), d\mathcal{F}(u)[h_u] \rangle + \langle d_u J(y, u), h_u \rangle \\ &= \langle d_u e(y, u)^* \xi, h_u \rangle + \langle d_u J(y, u), h_u \rangle = d_u \mathcal{L}(y, u, \xi)[h_u]. \end{aligned}$$

Note that we have already computed $d_u J(y, u)$ in Section 3.2.1. This allows us to identify the gradient of the reduced cost functional as

$$\nabla \hat{J}(u) = \sigma_2(u - u_{ref}) - \int_0^T \sum_i [d_u F(y_i, u)]^T \xi_{v_i}(t) dt. \tag{3.10}$$

4. Calibration algorithm and results

To calibrate the control parameters we use real data from the BaSiGo experiment carried out in Düsseldorf in 2013 [6, 25]. In the following, we denote the trajectories from experimental data by $x_i^{data} : [0, T] \rightarrow \mathbb{R}^2, i = 1 \dots N$. For the corridor case, we take the data from file "bi_corr_400_a_02.txt" in [25]. These show the positions of the pedestrians in the domain $\Omega = [-6, 6] \times [0, 4.2]$, over time $t \in [0, 150]$ seconds. For the crossing case, we take the data from file "CROSSING_90_E_2.txt" in [25]. This file provides the positions of the pedestrians in the crossing corridors $\Omega = ([-5, 5] \times [-1.5, 2]) \cap ([-1.2, 2] \times [-5, 5])$, over time $t \in [0, 283]$ seconds. However, in the calibration algorithm, we use only 8-second intervals for both scenarios. In this way we can extract several batches from one movie and are computationally efficient. Moreover, once the model deviates from the data, cost will accumulate over time. By working with small time intervals this error is reduced. However, we also tested longer simulations times which led to similar results.

Algorithm 2. Steepest Descent Algorithm

Input: Initial data $(\mathbf{x}_0, \mathbf{v}_0)$ of the pedestrians and all parameter values, initial guess for the parameters u_0

```

while  $\epsilon_{rel} > 10^{-2}$  do
    Solve state system (2.1);
    Solve adjoint system (3.9);
    Calculate gradient of the reduced cost functional (3.10);
    Update control parameters (4.1);
    Update learning rate  $\beta$  using Armijo rule;
end
    
```

end
Result: calibrated controls \bar{u} , trajectories (\bar{x}, \bar{v}) and their cost $J(\bar{y}, \bar{u})$.

4.1. Numerical schemes and steepest descent algorithm

In general, the nonlinearities make it rather difficult to solve the optimality system all at once. We, therefore, opt for an iterative approach to compute the gradient of the calibration problem. Indeed, we first solve the state system (2.1) as considered in Section 2.1.1. Then, we integrate the adjoint system (3.9) with the help of a second-order Runge-Kutta method backward in time. Here, we use the same time steps as for the state problem and transform the time via $s = T - t$ to recover an initial value problem. With the state solution and the adjoint solution, we calculate the gradient using (3.10), where the integral is approximated with the trapezoidal rule.

We apply a steepest descent algorithm to update the control parameters in every iteration

$$u^{k+1} = u^k - \beta_k \cdot \nabla \hat{J}(u) \tag{4.1}$$

where u^k denotes the control on current time step, and $\nabla \hat{J}(u)$ denotes the descent direction and $\beta_k \in \mathbb{R}^4$ is a positive scaling vector. The complete optimization procedure is summarized in Algorithm 2.

Since we cannot assume that the data and the model are a perfect match, we employ a stochastic gradient descent approach using mini-batches [24]. To obtain the mini-batches from the data trajectories which are given on the interval $[0, T]$, we split the interval into M mini-batches, each of size $b_i = T/M$, $i = 1 \dots M$. Then we randomly select $m < M$ mini-batches for each of the gradient steps.

In more detail, at each iteration we compute the gradients of the m batches and approximate the gradient using the average

$$\nabla \hat{J}(u) = \frac{1}{m} \sum_{i=1}^m \nabla \hat{J}_{b_i}(u)$$

to update the control via (4.1). The stopping criterion of the calibration algorithm is based on the relative error between the previous and current cost function value denoted by ϵ_{rel} .

4.2. Numerical results

In this section, we discuss numerical results generated with Algorithm 2 using experimental data from the Pedestrian Dynamics Data Archive [25]. In particular, we retrieve the trajectories from video recordings showing bi-directional and cross-directional flows.

We fix the velocity scaling $\tau = 1$, attractive potential range $a = 1$, repulsive potential range $r = 0.3$. The total number of pedestrians presented in the data is $N = 84$ in the corridor case and $N = 71$ in the crossing scenario. The desired velocities of agents that go from right to left and from left to right are respectively $w_{red} = (-0.7, 0)^T$ and $w_{blue} = (0.7, 0)^T$. The value 0.7 is the average velocity that we extracted from the experimental data. The desired velocities for the crossing scenario are again computed from the experimental data and set to $w_{red} = (0, 1.2)^T$ and $w_{blue} = (1.2, 0)^T$. It is interesting that in the crossing

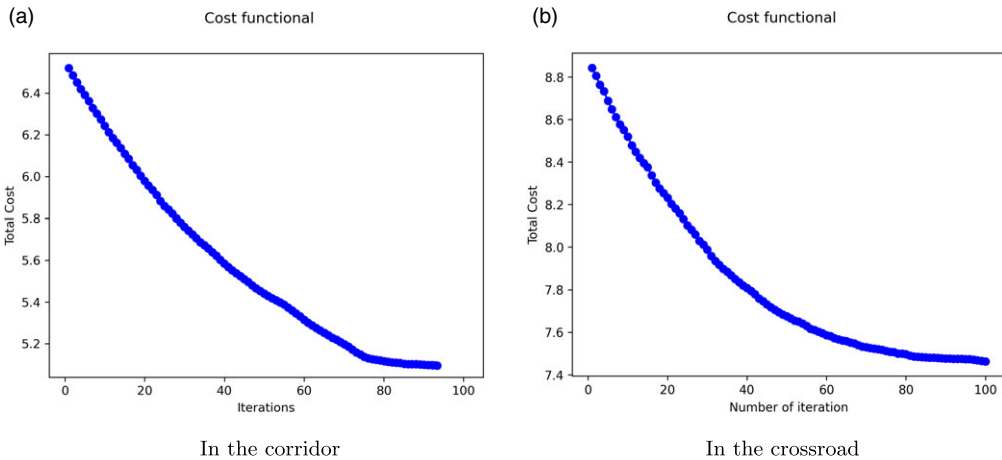


Figure 10. Cost functionals of the corridor and crossing scenarios.

scenario agents move faster than agents in the corridor. However, from the video data of the crossing scenario, we see that in the interacting part of the domain are fewer people. Also, in the corridor case, agents try to move alongside the whole corridor avoiding collisions, while in the crossing case, agents interact just on the crossroad.

The time step in the Leap Frog scheme and second order Runge-Kutta method is set to $\Delta t = 0.00625$. Simulations are done in the time interval $t \in [0, 8]$. The gradient is calculated with $m = 50$ mini-batches of length $|b_i| = 10 \cdot \Delta t$. The short mini-batch length can lead to a non-smooth decrease in the cost function.

The initial positions of the agents \mathbf{x}_0 coincide with the initial positions of the experimental data, which are distributed in the domain $[-6, 6] \times [0, 4.2]$. As the initial velocity of the pedestrians, we set the average velocity from the experimental data. We probably induce some error here, as we do not have the exact values from the data.

As an initial guess of the control parameters, we take $u_0 = (0, 0, 40, 0.6)$ for both experiments. We set the regularization parameters in the cost functional to $\sigma_1 = 1$ and $\sigma_2 = 0$. We, therefore, do not need to choose reference values for the control. In this setting, we repeat Algorithm 2 while the given stopping criterion is fulfilled. As the interval length of the mini-batches is rather short, the magnitude of the gradient components is rather small. We therefore multiply with a step size parameter β which accounts for the different ranges of the control values. This can be seen as a kind of preconditioning. The initial step size parameter $\beta_0 = (20, 4000, 4000, 20)$ was found by trial and error. Further, we change β_k using the Armijo rule. The Armijo rule ensures a sufficient decrease in the objective function.

Figure 10a illustrates the decrease of the cost functional for the corridor case with initial body diameter $d = 0.6$. It starts from approximately 6.53 in the first iteration and terminates at around 5.14 in the last iteration. In Figure 10b we see the evolution of the cost for the crossing scenario from 8.83 to 7.46. We note that the cost values of the corridor are smaller than the ones of the crossing case. This indicates that the proposed model captures the effects of counter-flow better than the effects of crossing flow.

In Table 2 the calibration results with different body size are presented. The parameters of the first calibration procedure reached optimal values $\lambda = -0.072$, $A = 6.14$, $R = 33.29$, and $d = 0.46$ for the simulation in the corridor. For the simulation at the crossing scenario, optimal values reached $\lambda = -0.068$, $A = 8.66$, $R = 32.91$, and $d = 0.456$. Interestingly, the optimal λ is negative and in the same range in both scenarios. This indicates that the pedestrians in the data set have a tendency to move to the left to avoid collisions. Moreover, the repulsion strengths are similar for both scenarios, but the values for the attraction strengths differ. The difference in the attraction strengths may arise from the post-interaction behavior of the model. Simulations, as reported in [34], show that two interacting agents

Table 2. Calibration results for the corridor and the crossing cases by different initial guesses for body sizes. The attraction and repulsion ranges are set $a = 1$, $r = 0.3$ in both scenarios

	Corridor					Crossing			
	Initial	$d = 0.6$	$d = 0.5$	$d = 0.35$		Initial	$d = 0.6$	$d = 0.5$	$d = 0.35$
λ	0	-0.072	-0.07	-0.065	λ	0	-0.068	-0.065	-0.071
A	0	6.14	6.12	6.23	A	0	8.66	8.97	9.28
R	40	33.29	30.18	31.47	R	40	32.91	29.39	31.62
d_{opt}		0.46	0.45	0.42	d_{opt}		0.456	0.44	0.43
J_{min}		5.14	5.27	5.19	J_{min}		7.46	7.53	7.34

move on with slightly shifted positions after a collision avoiding interaction. We suspect that this has more impact on the results in the crossing than the corridor scenario. So far, however, this is only a conjecture and needs to be proven or discarded by further investigations.

5. Conclusion and outlook

We extended the anisotropic interaction model proposed in [34] by including body size and thus additional volume exclusion effects. Numerical studies indicate that body size has an influence on the lane formation process. In fact, it seems that a smaller body size leads to a higher number of lanes formed and *vice versa*. Moreover, we investigated the fundamental diagram of the dynamics and found that higher densities lead to lower velocities. This observation was expected from other experiments and underlines the feasibility of the approach.

Due to the ODE formulation of the model, we were able to analyze the model in terms of well-posedness and further rigorously derive a gradient-based descent algorithm for a calibration problem using real data. The optimal scaling parameters for the collision avoidance and the repulsion strength turn out to match very well for the two tested scenarios. For the attraction parameter, we find different values for the two scenarios. We suspect that this is related to the post-collision behavior of the model.

To prove or discard this conjecture is interesting future work. Moreover, the rigorous analysis of stationary states like the lanes in the corridor case and the traveling waves in the crossing case is planned for future investigations.

Acknowledgement. ZT acknowledges funding by the German Academic Exchange Service programme “Mathematics in Industry and Commerce, MIC”.

Competing interests. The authors hereby declare that there are no competing interests.

References

- [1] Bailo, R., Carrillo, J. A. & Degond, P. (2018) *Pedestrian Models Based on Rational Behaviour*, Springer International Publishing, Cham, pp. 259–292.
- [2] Bellomo, N., Gibelli, L., Quaini, A. & Reali, A. (2022) Towards a mathematical theory of behavioral human crowds. *Math. Models Methods Appl. Sci.* **32** (02), 1–38.
- [3] Bode, N. W. F. & Ronchi, E. (2019) Statistical model fitting and model selection in pedestrian dynamics research. *Collective Dyn.* **4**, 1–32.
- [4] Burger, M., Hittmeir, S., Ranetbauer, H. & Wolfram, M.-T. (2016) Lane formation by side stepping. *SIAM Math. Anal.* **48** (2), 981–1005.
- [5] Burstedde, C., Klauck, K., Schadschneider, A. & Zittartz, J. (2001) Simulation of pedestrian dynamics using a two-dimensional cellular automaton. *Phys. A Stat. Mech. Appl.* **295** (3–4), 507–525.
- [6] Cao, S., Seyfried, A., Zhang, J., Holl, S. & Song, W. (2017) Fundamental diagrams for multidirectional pedestrian flows. *J. Stat. Mech. Theory Exp.* **3** (3), 033404.
- [7] Castellano, C., Fortunato, S. & Loreto, V. (2009) Statistical physics of social dynamics. *Rev. Mod. Phys.* **81** (2), 591–646.

- [8] Chen, X., Treiber, M., Kanagaraj, V. & Li, H. (2008) Social force models for pedestrian traffic-state of the art. *Transp. Rev.* **38** (5), 625–653.
- [9] Chraïbi, M., Kemloh, U. & Seyfried, A. (2011) Force-based models of pedestrian dynamics. *Netw. Heterog. Media* **6** (3), 425–442.
- [10] Chraïbi, M., Tordeux, A., Schadschneider, A. & Seyfried, A. (2018). Modelling of pedestrian and evacuation dynamics. **In: Encyclopedia of Complexity and Systems Science**, Springer, Berlin, Heidelberg, pp. 1–23.
- [11] Christiani, E., Priuli, F. S. & Tosin, A. (2015) Modeling rationality to control self-organization of crowds: an environmental approach. *SIAM J. Appl. Math.* **75** (2), 605–629.
- [12] Degond, P., Appert-Rolland, C., Moussaid, M., Petre, J. & Theraulaz, G. (2013) A hierarchy of heuristic-based models of crowd dynamics. *J. Stat. Phys.* **152** (6), 1033–1068.
- [13] Dijkstra, J., Jessurun, J. & Timmermans, H. J. (2001) A multi-agent cellular automata model of pedestrian movement. *Pedestrian Evacuation Dyn.* **173**, 173–180.
- [14] D’Orsogna, R., Chuang, L., Bertozzi, L. & Chayes, S. (2006) Self-propelled particles with soft-core interactions: patterns, stability, and collapse. *Phys. Rev. Lett.* **96** (10), 104302.
- [15] Festa, A., Tosin, A. & Wolfram, M.-T. (2018) Kinetic description of collision avoidance in pedestrian crowds by sidestepping. *Kinet. Relat. Models* **11** (3), 491–520.
- [16] Fornasier, M. & Solombrino, F. (2014) Mean-field optimal control. *ESAIM Control Optim. Calc. Var.* **20** (4), 1123–1152.
- [17] Gomes, S. N., Stuart, A. M. & Wolfram, M.-T. (2019) Parameter estimation for macroscopic pedestrian dynamics models from microscopic data. *SIAM Appl. Math.* **79** (4), 475–1500.
- [18] Göttlich, S. & Totzeck, C. (2021) Parameter calibration with stochastic gradient descent for interacting particle systems driven by neural networks. *Math. Control Signals Syst.*, 1–30.
- [19] Helbing, D., Farkas, I. J., Molnar, P. & Vicsek, T. (2002) Simulation of pedestrian crowds in normal and evacuation situations. *Pedestrian Evacuation Dyn.* **21** (2), 21–58.
- [20] Helbing, D. & Molnar, P. (1995) Social force model for pedestrian dynamics. *Phys. Rev. E* **51** (5), 4282–4286.
- [21] Hinze, M., Pinnau, R., Ulbrich, M. & Ulbrich, S. (2009). *Optimization with PDE Constraints*, Springer.
- [22] Hoogendoorn, S. P. & Bovy, P. H. (2001) Generic gas-kinetic traffic systems modeling with applications to vehicular traffic flow. *Transp. Res. Part B Methodol.* **35** (4), 317–336.
- [23] Khelfa, B., Korbmacher, R., Schadschneider, A. & Tordeux, A. (2021) Heterogeneity-induced lane and band formation in self-driven particle systems. Preprint on arXiv: 2110.05874.
- [24] Li, M., Zhang, T., Chen, Y. & Smola, A. J. (2014) Efficient mini-batch training for stochastic optimization. **In: Proceedings of the 20th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining**, pp. 661–670.
- [25] Pedestrian Dynamics Data Archive. <https://ped.fz-juelich.de/da/doku.php>.
- [26] Schadschneider, A., Chraïbi, M., Seyfried, A., Tordeux, A. & Zhang, J. (2018) Pedestrian dynamics: from empirical results to modeling. **In: Crowd Dynamics, Volume 1, Modeling and Simulation in Science, Engineering and Technology**, Birkhäuser, Cham, pp. 1–29.
- [27] Schadschneider, A., Klüpfel, H., Kretz, T., Rogsch, C. & Seyfried, A. (2009) Fundamentals of pedestrian and evacuation dynamics. **In: Multi-Agent Systems for Traffic and Transportation Engineering**, IGI Global, pp. 124–154.
- [28] Seitz, M. J. (2016) Simulating pedestrian dynamics (Doctoral dissertation, Technische Universität München).
- [29] Seyfried, A., Steffen, B., Klingsch, W. & Boltes, M. (2005) The fundamental diagram of pedestrian movement revisited. *J. Stat. Mech. Theory Exp.* **2005** (10), P10002–P10002.
- [30] Shewchuk, J. R. (2002) Delaunay refinement algorithms for triangular mesh generation. *Comput. Geom.* **22** (1-3), 21–74.
- [31] Sieben, A., Schumann, J. & Seyfried, A. (2017) Collective phenomena in crowds - where pedestrian dynamics need social psychology. *PLoS One* **12** (6), e0177328.
- [32] Steffen, B. & Seyfried, A. (2010) Methods for measuring pedestrian density, flow, speed and direction with minimal scatter. *Phys. A Stat. Mech. Appl.* **389** (9), 1902–1910.
- [33] Teknomo, K., Takeyama, Y. & Inamura, H. (2016) Review on microscopic pedestrian simulation model. arXiv preprint arXiv:1609.01808.
- [34] Totzeck, C. (2020) An anisotropic interaction model for pedestrian dynamics with collision avoidance. *Kinet. Relat. Models* **13** (6), 1219–1242.
- [35] Tröltzsch, F. (2005) *Optimal Control of Partial Differential Equations. Theory, Methods and Applications*, American Mathematical Society, Graduate Studies in Mathematics.
- [36] Zhang, D., Zhu, H., Hostikka, S. & Qiu, S. (2019) Pedestrian dynamics in a heterogeneous bidirectional flow: overtaking behaviour and lane formation. *Phys. A Stat. Mech. Appl.* **525**, 72–84.
- [37] Zhou, M., Dong, H., Ioannou, P. A., Zhao, Y. & Wang, F. Y. (2019) Guided crowd evacuation: approaches and challenges. *IEEE/CAA J. Autom. Sin.* **6**, 1081–1094.

Appendix A. Boundedness of $\nabla M(\cdot)$

We show the boundedness of $\nabla M(v_i, v_j)$, where

$$\nabla M(v_i, v_j) = \left(\frac{dM(v_i, v_j)}{dv_i}, \frac{dM(v_i, v_j)}{dv_j} \right),$$

with

$$\frac{dM}{dv_i} = \begin{pmatrix} -\sin \alpha & -\cos \alpha \\ \cos \alpha & -\sin \alpha \end{pmatrix} \cdot \frac{d\alpha}{dv_i},$$

$$\frac{d\alpha}{dv_i} = \begin{cases} -\lambda \frac{1}{\sqrt{(\|v_i\| \|v_j\|)^2 - \langle v_i, v_j \rangle^2}} \left(v_j - \langle v_i, v_j \rangle \frac{v_i}{\|v_i\|^2} \right), & \text{if } v_i, v_j \neq 0 \\ 0, & \text{else} \end{cases} \tag{A1}$$

We, therefore, transform the system to polar coordinates and introduce the notations $v_i = \begin{pmatrix} r_1 \cos \phi_1 \\ r_1 \sin \phi_1 \end{pmatrix}$, and $v_j = \begin{pmatrix} r_2 \cos \phi_2 \\ r_2 \sin \phi_2 \end{pmatrix}$. Here, r_1 and r_2 are positive scalars. It is clear that $\left| \begin{pmatrix} -\sin \alpha & -\cos \alpha \\ \cos \alpha & -\sin \alpha \end{pmatrix} \right| < \infty$. Then, we need to show $\left| \frac{d\alpha}{dv_i} \right| = C < \infty$.

For further usage, we define

$$\|v_i\|_2 = r_1, \quad \|v_j\|_2 = r_2 \text{ and } \langle v_i, v_j \rangle = r_1 r_2 (\cos \phi_1 \cos \phi_2 + \sin \phi_1 \sin \phi_2).$$

Substituting the velocity vectors in (A1), we get

$$\begin{aligned} \left| \frac{d\alpha}{dv_i} \right|_2 &= \left| \frac{-\lambda}{\sqrt{r_1^2 r_2^2 - (r_1 r_2 \cos(\phi_1 - \phi_2))^2}} \right| \left| \begin{pmatrix} r_2 \cos \phi_2 \\ r_2 \sin \phi_2 \end{pmatrix} - r_1 r_2 \cos(\phi_1 - \phi_2) \frac{1}{r_1^2} \begin{pmatrix} r_1 \cos \phi_1 \\ r_1 \sin \phi_1 \end{pmatrix} \right|_2 \\ &= \frac{|\lambda|}{r_1 r_2 |\sin(\phi_1 - \phi_2)|} \left| \begin{pmatrix} r_2 \cos \phi_2 - r_2 \cos \phi_1 \cos(\phi_1 - \phi_2) \\ r_2 \sin \phi_2 - r_2 \sin \phi_1 \cos(\phi_1 - \phi_2) \end{pmatrix} \right|_2 \\ &= \frac{|\lambda|}{r_1 r_2 |\sin(\phi_1 - \phi_2)|} \left[r_2^2 (\cos \phi_2 - \cos \phi_1 \cos(\phi_1 - \phi_2))^2 \right. \\ &\quad \left. + r_2^2 (\sin \phi_2 - \sin \phi_1 \cos(\phi_1 - \phi_2))^2 \right]^{\frac{1}{2}} \\ &= \frac{|\lambda|}{r_1 r_2 |\sin(\phi_1 - \phi_2)|} r_2 \sqrt{1 - \cos^2(\phi_1 - \phi_2)} = \frac{|\lambda|}{r_1} < \infty. \end{aligned}$$

Analogously, we can show that $\left| \frac{d\alpha}{dv_j} \right|_2 = \frac{|\lambda|}{r_2} < \infty$. This proves the boundedness of $\nabla M(v_i, v_j)$ as desired.

Appendix B. Tensor $d_{v_i} M^*(v_i, v_j)$

Here, we present the details of the mathematical manipulations made to tensor $d_{v_i} M(v_i, v_j)$ in the derivation of $d_{v_i} M^*(v_i, v_j)$. Note that $d_{v_i} M(v_i, v_j)$ is a tensor with dimension $2 \times 2 \times 2$ given by:

$$d_{v_i} M(v_i, v_j) = \begin{pmatrix} -\sin \alpha_{ij} \cdot d_{v_i} \alpha_{ij} & -\cos \alpha_{ij} \cdot d_{v_i} \alpha_{ij} \\ \cos \alpha_{ij} \cdot d_{v_i} \alpha_{ij} & -\sin \alpha_{ij} \cdot d_{v_i} \alpha_{ij} \end{pmatrix}_{2 \times 2 \times 2}, \tag{B1}$$

where

$$d_{v_i} \alpha_{ij} = \begin{cases} -\lambda \cdot \frac{1}{\sqrt{(\|v_i\| \|v_j\|)^2 - \langle v_i, v_j \rangle^2}} \cdot \left(v_j - \langle v_i, v_j \rangle \frac{v_i}{\|v_i\|^2} \right), & \text{for } v_i \neq 0, v_j \neq 0, \\ \vec{0}, & \text{else} \end{cases}$$

The elements of the tensor $d_{v_i}M(v_i, v_j)$ in (B1) are denoted as:

$$d_{v_i}M(v_i, v_j) = \begin{pmatrix} \begin{pmatrix} m_{11} \\ m_{12} \\ m_{31} \\ m_{32} \end{pmatrix} & \begin{pmatrix} m_{21} \\ m_{22} \\ m_{41} \\ m_{42} \end{pmatrix} \end{pmatrix}, \quad \text{with } m_{ij} \in \mathbb{R}, i = 1, \dots, 4, j = 1, 2,$$

then its dual $d_{v_i}M^*(v_i, v_j)$ in (3.6) is the transposed tensor after swapping its axes, as given below:

$$d_{v_i}M^*(v_i, v_j) = \begin{pmatrix} \begin{pmatrix} m_{11} \\ m_{21} \\ m_{12} \\ m_{22} \end{pmatrix} & \begin{pmatrix} m_{31} \\ m_{41} \\ m_{32} \\ m_{42} \end{pmatrix} \end{pmatrix}.$$

Appendix C. Existence of adjoint states

For completeness, we give the assumptions and the statement of Corollary 1.3 of [21] which we use to prove the existence of adjoint states in Section 3.2.2.

Assumption 3. *Let the following assumptions hold:*

- (A1) $U_{\text{ad}} \subset U$ is nonempty, convex, and closed.
- (A2) $J: Y \times U \rightarrow \mathbb{R}$ and $e: Y \times U \rightarrow Z$ are continuously Fréchet differentiable and U, Y, Z are Banach spaces.
- (A3) For all $u \in V$ in a neighbourhood $V \subset U$ of U_{ad} , the state equation $e(y, u) = 0$ has a unique solution $y = y(u) \in Y$.
- (A4) $e_y(y(u), u) \in \mathcal{L}(Y, Z)$ has a bounded inverse for all $u \in V$.

Proposition 1 (Existence of adjoint states [21]). *Let (\bar{y}, \bar{u}) be an optimal solution of $\min J(y, u)$ subject to $e(y, u) = 0$*

and let Assumption 3 hold.

Then there exists an adjoint state (or Lagrange multiplier) $\bar{p} \in Z^$ such that the following optimality conditions hold*

$$\begin{aligned} \langle e(\bar{y}, \bar{u}), p \rangle_{Z, Z^*} &= 0 & \forall p \in Z^*, \\ \langle L_y(\bar{y}, \bar{u}, \bar{p}), v \rangle_{Y^*, Y} &= 0 & \forall v \in Y^*, \\ \bar{u} \in U_{\text{ad}}, \langle L_u(\bar{y}, \bar{u}, \bar{p}), u - \bar{u} \rangle_{U^*, U} &\geq 0, & \forall u \in U_{\text{ad}}. \end{aligned}$$