RESEARCH ARTICLE

Space mapping-based optimization with the macroscopic limit of interacting particle systems

Jennifer Weißen[1] · Simone Göttlich[1] · Claudia Totzeck[1]

Abstract

We propose a space mapping-based optimization algorithm for microscopic interacting particle dynamics which are infeasible for direct optimization. This is of relevance for example in applications with bounded domains for which the microscopic optimization is difficult. The space mapping algorithm exploits the relationship of the microscopic description of the interacting particle system and a corresponding macroscopic description as partial differential equation in the "many particle limit". We validate the approach with the help of a toy problem that allows for direct optimization. Then we study the performance of the algorithm in two applications. A pedestrian flow is considered and the transportation of goods on a conveyor belt is optimized. The numerical results underline the feasibility of the proposed algorithm.

1 Introduction

In the recent decades interacting particle systems attracted a lot of attention from researchers of various fields such as swarming, pedestrian dynamics and opinion formation (cf. Albi and Pareschi 2013; Helbing and Molnár 1995; Toscani 2006; Totzeck 2020 and the references therein). In particular, a model hierarchy was

✉ Simone Göttlich
goettlich@uni-mannheim.de

Jennifer Weißen
jennifer.weissen@uni-mannheim.de

Claudia Totzeck
totzeck@uni-mannheim.de

[1] University of Mannheim, Department of Mathematics, 68131 Mannheim, Germany

established Carrillo et al. 2010; Golse 2003. The main idea of the hierarchy is to model the same dynamics with different accuracies, each having its own advantages and disadvantages. The model with the highest accuracy is the microscopic one. It describes the positions and velocities of each particle explicitly. For applications with many particles involved, this microscopic modelling leads to a huge amount of computational effort and storage needed. Especially, when it comes to the optimization of problems with many particles Burger et al. 2020 2020.

There is also an intermediate level of accuracy given by the mesoscopic description, see Albi and Pareschi 2013; Carrillo et al. 2010; Totzeck 2020. We do not want to give its details here, instead, we directly pass to the macroscopic level, where the velocities are averaged and a position-dependent density describes the probability of finding a particle of the dynamics at a given position. Of course, we lose the explicit information of each particle, but have the advantage of saving a lot of storage in the simulation of the dynamics. Despite the lower accuracy many studies Albi and Pareschi 2013; Burger et al. 2020; Mahato et al. 2018 indicate that the evolution of the density yields a good approximation of the original particle system, see also Weissen et al. (2021), which proposed a limiting procedure that is considered in more detail below.

Moreover, the macroscopic description naturally involves boundary conditions which play a crucial role in many engineering experiments with interacting particles, see for example the pedestrian dynamics data archive by the Civil Safety Research of the Forschungszentrum Jülich Seyfried and Boltes (2021).

This observation motivates us to exploit the aforementioned relationship of microscopic and macroscopic models and propose a space mapping-based optimization scheme for interacting particle dynamics which are inappropriate for direct optimization. This might be the case for particle dynamics that involve a huge number of particles for which traditional optimization is expensive in terms of storage, computational effort and time. Another example is the optimization of particle dynamics in bounded domains, where the movement is restricted by obstacles or walls. In fact, systems based on ordinary differential equations (ODEs) do not have a natural prescription of zero-flux or Neumann boundary data, but those conditions are often critical in applications.

An exemplary application is the movement of large crowds Göttlich and Pfirsching 2018; Helbing et al. 2000; Helbing and Molnár 1995 which includes the combination of individual behavior with a herding instinct leading to collective motion. Systematic studies analyze the behavior of human crowds in crowded buildings or at event venues with a high number of participants such as in concerts Johnson 1987 or religious gatherings Haase et al. 2019. Dangerous overcrowding and jamming might occur at bottlenecks Helbing et al. 2000. The herding instinct entails that all individuals move in the same direction of probably blocked pathways. Their movement becomes uncoordinated and since paths are clogged, jams build up. If people get too close together they start to push and interact with each other. The interaction between two pedestrians is modeled using an interaction force which pushes the pedestrians apart if they touch each other. It becomes the dominant force in this case and the physical interaction within these crowds might even become dangerous Helbing et al. 2000. The interactions with walls is also modeled with a similar

interaction force Helbing et al. 2000 and the geometry of the domain strongly influences the behavior of the crowd. Interesting scenarios are the evacuation of people from a hotel or an office building Okazaki and Matsushita 1993, football stadia Elliott and Smith 1993, passenger ships Klüpfel et al. 2001 or aircrafts Miyoshi et al. 2012. All of the described settings naturally involve bounded domains and the pedestrian flows depend on the geometry of the domain Helbing et al. 2002. Simulations show the regions where congestion and stagnation in the movement occur. This information can help designers to examine and verify evacuation routes. Another useful application is the transport of material flow on conveyor belts Göttlich et al. 2014; Göttlich and Pfirsching 2018. Conveyor belts are often used in factories to transport large amounts of goods, e.g. in bottling plants Festa et al. 2019. The goods are transported on the belts and redirected by obstacles such as diverter equipment Prims et al. 2019. When parts are too close together they exhibit repelling forces on each other. The same holds true for the interaction with obstacles or walls through which parts experience forces that allow to alter their transport direction.

While boundary interaction can be included in the ODE dynamics, their treatment is not self-evident. The same holds for optimization algorithms based on these simulations. To the authors' knowledge, alternatives to black box optimization algorithms that allow for convenient optimization of ODE systems within bounded domains do not exist. In contrast to the microscopic ones, models based on partial differential equations (PDEs) require boundary conditions. Often zero-flux or Neumann type boundary conditions are prescribed. Further, numerical schemes naturally integrate these boundary conditions and optimization algorithms can be directly formulated. The approach discussed in the following allows us to approximate the optimizer of microscopic dynamics with additional boundary behavior while only optimizing the macroscopic model with an adjoint based scheme. It is therefore a sophisticated alternative to black box optimization approaches.

1.1 Modeling equations and general optimization problem

We begin with the general framework and propose the space mapping technique to approximate an optimal solution of the interacting particle system. In general, the interacting particle dynamics for $N \in \mathbb{N}$ particles in the microscopic setting is given by the ODE system

$$\frac{\mathrm{d}x_i}{\mathrm{d}t} = v_i,$$
$$m\frac{\mathrm{d}v_i}{\mathrm{d}t} = G(x_i, v_i) + A\sum_{j \neq i} F(x_i - x_j), \qquad i = 1, \ldots N \qquad (1)$$
$$x_i(0) = x_i^0, v_i(0) = v_i^0,$$

where $x_i \in \mathbb{R}^2, v_i \in \mathbb{R}^2$ are the position and the velocity of particle i supplemented with initial condition $x_i(0) = x_i^0, v_i(0) = v_i^0$ for $i = 1, \ldots, N$. Here, F denotes an interaction kernel which is often given as a gradient of a potential D'Orsogna et al.

2006a. For notational convenience, we define the state vector $y = (x_i, v_i)_{i=1,...,N}$ which contains the position and velocity information of all particles.

Remark 1 Note that there are models that include boundary dynamics with the help of soft core interactions, see for example Helbing and Molnár 1995. In general, these models allow for direct optimization. Nevertheless, for $N \gg 1$ the curse of dimensionality applies and the approach discussed here may still be useful.

Sending $N \to \infty$ and averaging the velocity, we formally obtain a macroscopic approximation of the ODE dynamics given by the PDE

$$\partial_t \rho + \nabla \cdot \left(\rho \bar{v}(x) - k(\rho)\nabla\rho \right) = 0, \qquad (x,t) \in \Omega \times [0,T]$$
$$\rho(x,0) = \rho^0(x), \qquad x \in \Omega \tag{2}$$

where $\rho = \rho(x,t)$ denotes the particle density in the domain $\Omega \subseteq \mathbb{R}^2$. The velocity $\bar{v}$ is the averaged velocity depending on the position and $k(\rho)$ is the diffusion coefficient.

We consider constrained optimization problems of the form

$$\min_{u \in \mathcal{U}_{ad}} J(u,y)$$

$$\text{subject to } E(u,y) = 0,$$

where J is the cost functional, $\mathcal{U}_{ad}$ is the set of admissible controls and y are the state variables with $E(u,y) = 0$. In the following, for a given control $u \in \mathcal{U}_{ad}$, the constraint $E(u,y)$ contains the modeling equations for systems of ODEs or PDEs. With the additional assumption that for a given control u, the model equations have a unique solution, we can express $y = y(u)$ and consider the reduced problem

$$\min_{u \in \mathcal{U}_{ad}} J(u, y(u)). \tag{3}$$

This is a nonlinear optimization problem, which we intend to solve for an ODE constraint $E(u, y(u))$. To do this, one might follow a standard approach Hinze et al. 2009 and apply a gradient descent method based on adjoints Tröltzsch 2010 to solve the microscopic reduced problem iteratively. In contrast, the space mapping technique employs a cheaper, substitute model (coarse model) for the optimization of the fine model optimization problem. Under the assumption that the optimization of the microscopic system is difficult and the optimization of the macroscopic system can be computed efficiently, we propose space mapping-based optimization. The main objective is to iteratively approximate an optimal control for the microscopic dynamics. To get there, we solve a related optimal control problem on the macroscopic level in each iteration.

1.2 Literature review and outline

Space mapping was originally introduced in the context of electromagnetic optimization Bandler et al. 1994. The original formulation has been subject to improvements

and changes Bandler et al. 2004 and enhanced by classical methods for nonlinear optimization. The use of Broyden's method to construct a linear approximation of the space mapping function, so-called aggressive space mapping (ASM) was introduced by Bandler et al. Bandler et al. (1995). We refer to Bakr et al. (2000); Bandler et al. (2004) for an overview of space mapping methods.

More recently, space mapping has been successfully used in PDE based optimization problems. Banda and Herty Banda and Herty (2011) presented an approach for dynamic compressor optimization in gas networks. Göttlich and Teuber Göttlich and Teuber (2018) use space mapping based optimization to control the inflow in transmission lines. In both cases, the fine model is given by hyperbolic PDEs on networks and the main difficulty arises from the nonlinear dynamics induced by the PDE. These dynamics limit the possibility to efficiently solve the optimization problems. In their model hierarchy, a simpler PDE serves as the coarse model and computational results demonstrate that such a space mapping approach enables to efficiently compute accurate results.

Pinnau and Totzeck Totzeck and Pinnau (2020) used space mapping for the optimization of a stochastic interacting particle system. In their approach the deterministic state model was used as coarse model and lead to satisfying results.

Here, we employ a mixed hyperbolic-parabolic PDE as the coarse model in the space mapping technique to solve a control problem on the ODE level. Our optimization approach therefore combines different hierarchy levels. As discussed, the difficulty on the ODE level can arise due to boundaries in the underlying spatial domain or due to a large number of interacting particles. In contrast, the macroscopic equation naturally involves boundary conditions and its computational effort is independent of the particle number.

The outline of the paper is as follows: We introduce the space mapping technique in Sect 2 together with the fine and coarse model description in the Sects. 2.1 and 2.2. Particular attention is paid to the solution approach for the discretized coarse model in Sect. 2.2.2, which is an essential step in the space mapping algorithm. The discretized fine model optimal control problem is presented in Sect. 3. We give an example where we can compare and validate our approach to a standard optimization approach for the fine model. We provide numerical examples in bounded domains in Sect. 4. In the Sects. 4.1 and 4.2, the microscopic optimization approach cannot be applied due to the additional boundary interaction. However, the space mapping algorithm can still be applied and properly includes the boundary interaction. Various controls such as the source of an eikonal field in evacuation dynamics, cf. Sect. 4.1, and the conveyor belt velocity in a material flow setting, cf. Sect. 4.2, demonstrate the diversity of the proposed space mapping approach. In the conclusion in Sect. 5 our insights are summarized.

2 Space mapping technique

Space mapping considers a model hierarchy consisting of a coarse and a fine model. Let $\mathcal{G}^c : \mathcal{U}_{ad}^c \to \mathbb{R}^{n_c}, \mathcal{G}^f : \mathcal{U}_{ad}^f \to \mathbb{R}^{n_f}$ denote the operators mapping a given control u to a specified observable $\mathcal{G}^c(u)$ in the coarse and $\mathcal{G}^f(u)$ in the fine model, respectively.

The idea of space mapping is to find the optimal control $u_*^f \in \mathcal{U}_{ad}^f$ of the complicated (fine) model control problem with the help of a coarse model, that is simple to optimize.

We assume that the optimal control of the fine model

$$u_*^f = \operatorname{argmin}_{u \in \mathcal{U}_{ad}^f} \|\mathcal{G}^f(u) - \omega_*\|,$$

where $\omega_* \in \mathbb{R}^n$ is a given target state, is inappropriate for optimization. In contrast, we assume the optimal control $u_*^c \in \mathcal{U}_{ad}^c$ of the coarse model control problem

$$u_*^c = \operatorname{argmin}_{u \in \mathcal{U}_{ad}^c} \|\mathcal{G}^c(u) - \omega_*\|,$$

can be obtained with standard optimization techniques. While it is computationally cheaper to solve the coarse model, it helps to acquire information about the optimal control variables of the fine model. By exploiting the relationship of the models, space mapping combines the simplicity of the coarse model and the accuracy of the more detailed, fine model very efficiently Bakr et al. 2001; Echeverría and Hemker 2005.

Definition 2.1 The space mapping function $\mathcal{T}: \mathcal{U}_{ad}^f \to \mathcal{U}_{ad}^c$ is defined by

$$\mathcal{T}(u^f) = \operatorname{argmin}_{u \in \mathcal{U}_{ad}^c} \|\mathcal{G}^c(u) - \mathcal{G}^f(u^f)\|.$$

The process of determining $\mathcal{T}(u^f)$, the solution to the minimization problem in Definition 2.1, is called parameter extraction. It requires a single evaluation of the fine model $\mathcal{G}^f(u^f)$ and a minimization in the coarse model to obtain $\mathcal{T}(u^f) \in \mathcal{U}_{ad}^c$. Uniqueness of the solution to the optimization problem is desirable but in general not ensured since it strongly depends on the two models and the admissible sets of controls $\mathcal{U}_{ad}^f, \mathcal{U}_{ad}^c$, see Echeverría and Hemker 2005 for more details.

The basic idea of space mapping is that either the target state is reachable, i.e., $\mathcal{G}^f(u_*^f) \approx \omega_*$ or both models are relatively similar in the neighborhood of their optima, i.e., $\mathcal{G}^f(u_*^f) \approx \mathcal{G}^c(u_*^c)$. Then we have

$$\mathcal{T}(u_*^f) = \operatorname{argmin}_{u \in \mathcal{U}_{ad}^c} \|\mathcal{G}^c(u) - \mathcal{G}^f(u_*^f)\| \approx \operatorname{argmin}_{u \in \mathcal{U}_{ad}^c} \|\mathcal{G}^c(u) - \omega_*\| = u_*^c,$$

compare Echeverría and Hemker 2005. In general, it is very difficult to establish the whole mapping $\mathcal{T}$, we therefore only use evaluations. In fact, the space mapping algorithms allow us to shift most of the model evaluations in an optimization process to the faster, coarse model. In particular, no gradient information of the fine model is needed to approximate the optimal fine model control Bakr et al. 2001. Figure 1 illustrates the main steps of the space mapping algorithm, see also Bandler et al. (2004); Koziel et al. (2011).

In the literature, many variants of the space mapping idea can be found Bandler et al. 2004. We will use the ASM algorithm, see algorithm 1 in "Appendix A" or the references Bandler et al. 1995; Göttlich and Teuber 2018 for algorithmic details. Starting from the iterate $u_1 = u_*^c$, the descent direction d_k is updated in each iteration k using the space mapping evaluation $\mathcal{T}(u_k)$. The algorithm terminates when the

Fig. 1 Schematic representation of a space mapping algorithm
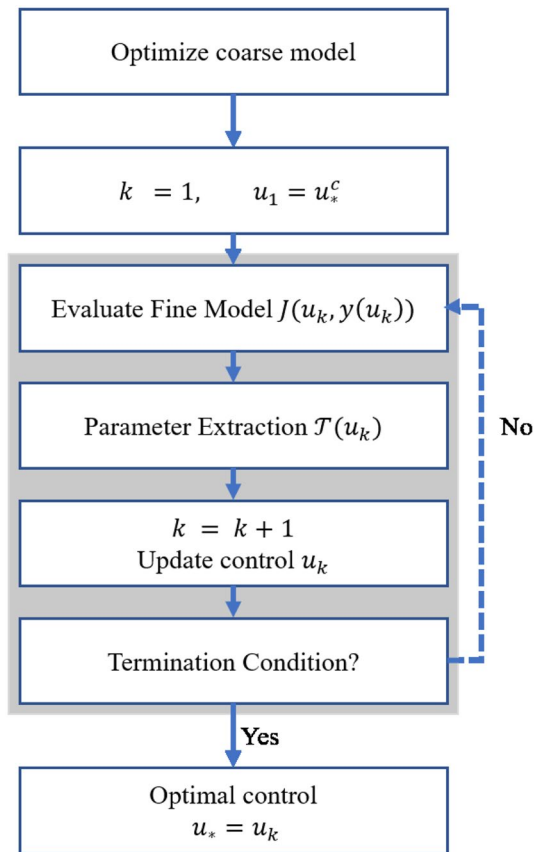
parameter extraction maps the current iterate u_k (approximately) to the coarse model optimum u_*^c, such that $\|\mathcal{T}(u_k) - u_*^c\|$ is smaller than a given tolerance in an appropriate norm $\|\cdot\|$. The solutions u_*^c and $\mathcal{T}(u_k)$ are computed using adjoints here and will be explained in Sect. 2.2.2.

2.1 Fine model

We seek to control a general microscopic model for the movement of N particles with dynamics given by (1). We choose the velocity selection mechanism

$$G(x, v) = -\frac{(v - \bar{v}(x))}{\tau},$$

which describes the correction of the particle velocities towards an equilibrium velocity $\bar{v}(x)$ with relaxation time τ. Such systems describe the movements of biological ensembles such as school of fish, flocks of birds Armbruster et al. 2017; Chuang et al. 2007; D'Orsogna et al. 2006b, ants Boi et al. 2000 or bacterial colonies Koch and White 1998 as well as pedestrian crowds Göttlich et al. 2018;

Helbing and Molnár 1995 and transport of material Göttlich et al. 2014, 2015. In general, the force F occuring in (1) is a pairwise interaction force between particle i and particle j. We choose to activate it whenever two particles overlap and therefore $\|x_i - x_j\|_2 < 2R$. For $\|x_i - x_j\|_2 \geq 2R$, the interaction force is assumed to be zero. In the following, we restrict ourselves to forces described by

$$F(x_i - x_j) = \begin{cases} b_F \big(\|x_i - x_j\|_2 - 2R \big)^2 \frac{x_i - x_j}{\|x_i - x_j\|_2} & \text{if } \|x_i - x_j\|_2 \leq 2R, \\ 0 & \text{otherwise.} \end{cases} \quad (4)$$

where $b_F > 0$.

We consider the optimization problem (3) and set $E(u, y) = 0$ if and only if the microscopic model equations (1) are satisfied to investigate various controls u. For example, u being the local equilibrium velocity $\bar{v}(x)$ of the velocity selection mechanism or u being the factor A scaling the interaction force between the particles. The objective function under consideration in each of the scenarios is the squared deviation of the performance evaluation $j(u, y(u))$ from the target value $\omega_* \in \mathbb{R}$, that is

$$J(u, y(u)) = \frac{1}{2} \big(j(u, y(u)) - \omega_* \big)^2. \quad (5)$$

In the following, we discuss the macroscopic approximation which is used as a coarse model for the space mapping.

2.2 Coarse model

Reference Weissen et al. 2021 shows that in the many particle limit, $N \to \infty$, the microscopic system (1) can be approximated by the advection-diffusion Eq. (2) with diffusion coefficient

$$k(\rho) = C\rho H(\rho - \rho_{crit}), \qquad C > 0.$$

The density $\rho_{crit} = 1$ is a density threshold, above which diffusion in the macroscopic model is activated. H denotes the Heaviside function

$$H(x) = \begin{cases} 0 & \text{if } x < 0, \\ 1 & \text{otherwise .} \end{cases}$$

At the boundary, we apply zero-flux boundary conditions for the advective and diffusive flux

$$\begin{aligned} (\bar{v}\rho) \cdot \vec{n} &= 0, \quad x \in \partial\Omega, \\ (k(\rho)\nabla\rho) \cdot \vec{n} &= 0, \quad x \in \partial\Omega, \end{aligned} \quad (6)$$

where $\vec{n} = (n^{(1)}, n^{(2)})^T$ is the outer normal vector at the boundary $\partial\Omega$.

The advection-diffusion Eq. (2) serves as the coarse model in the space mapping technique. To solve optimization problems in the coarse model, we pursue a first-discretize-then-optimize approach. In the following, we discretize the macroscopic

model and derive the first order optimality system for the discretized macroscopic system.

Remark 2 We recommend to choose the optimization approach depending on the structure of the macroscopic equation. Here, the PDE is hyperbolic whenever no particles overlap, we therefore choose first-discretize-then-optimize. If the macroscopic equation would be purely diffusive, one might employ a first-optimize-then-discretize approach instead.

2.2.1 Discretization of the macroscopic model

We discretize a rectangular spatial domain $(\Omega \cup \partial\Omega) \subset \mathbb{R}^2$ with grid points $x_{ij} = (x_{ij}^{(1)}, x_{ij}^{(2)})^T$, $(i,j) \in \mathcal{I}_\Omega = \{1, \dots, N_x\} \times \{1, \dots N_x\}$ on a grid with step size Δx in both coordinate directions. The boundary $\partial\Omega$ is described with the set of indices $\mathcal{I}_{\partial\Omega} \subset \mathcal{I}_\Omega$. The time discretization of the coarse model is Δt^c and the grid constant is $\lambda = \Delta t^c / \Delta x$. We compute the approximate solution to the advection-diffusion Eq. (2) as follows

$$\rho(x,t) = \rho_{ij}^s \text{ for } \begin{cases} x \in \mathcal{C}_{ij}, \\ t \in [t^s, t^{s+1}), \end{cases}$$

where

$$\mathcal{C}_{ij} = \left[x_{i-\frac{1}{2},j}^{(1)}, x_{i+\frac{1}{2},j}^{(1)} \right) \times \left[x_{i,j-\frac{1}{2}}^{(2)}, x_{i,j+\frac{1}{2}}^{(2)} \right),$$

$$t^s = s\Delta t^c \quad \text{for } s = 1, \dots, N_t^c.$$

The discretization of the initial density in (2) is obtained from the microscopic initial positions using a kernel density approach, see e.g. Fan and Seibold (2013); Parzen (1962); Rosenblatt (1956). This means that the initial density is constructed from the microscopic positions which are smoothed with a Gaussian filter such that the initial density reads

$$\rho^0(x) = \frac{\pi R^2}{2\pi} \sum_{i=1}^{N} \exp\left(-\frac{1}{2} \|x - x_i^0\|_2^2 \right). \tag{7}$$

To compute $\rho_{ij}^s, s > 0$, we solve the advection part with the Upwind scheme and apply dimensional splitting. The diffusion part is solved implicitly

$$\begin{aligned} \tilde{\rho}_{ij}^s &= \rho_{ij}^s - \frac{\Delta t^c}{\Delta x}\left(\mathcal{F}_{ij}^{(1),s,+} - \mathcal{F}_{ij}^{(1),s,-} \right), \\ \overline{\rho}_{ij}^s &= \tilde{\rho}_{ij}^s - \frac{\Delta t^c}{\Delta x}\left(\mathcal{F}_{ij}^{(2),s,+} - \mathcal{F}_{ij}^{(2),s,-} \right), \\ \rho_{ij}^{s+1} &= \overline{\rho}_{ij}^s + \frac{\Delta t^c}{(\Delta x)^2} B_{ij}^{s+1}, \end{aligned} \tag{8}$$

where the following short notation is used

$$\mathcal{F}_{ij}^{(1),s,+} = \mathcal{F}^{(1)}(\rho_{ij}^s, \rho_{i+1j}^s), \qquad \mathcal{F}_{ij}^{(1),s,-} = \mathcal{F}^{(1)}(\rho_{i-1j}^s, \rho_{ij}^s),$$

$$\mathcal{F}_{ij}^{(2),s,+} = \mathcal{F}^{(2)}(\tilde{\rho}_{ij}^s, \tilde{\rho}_{ij+1}^s), \qquad \mathcal{F}_{ij}^{(2),s,-} = \mathcal{F}^{(2)}(\tilde{\rho}_{ij-1}^s, \tilde{\rho}_{ij}^s),$$

$$B_{ij}^{s+1} = B\Big(\rho_{i-1j}^{s+1}, \rho_{i+1j}^{s+1}, \rho_{ij}^{s+1}, \rho_{ij-1}^{s+1}, \rho_{ij+1}^{s+1}\Big).$$

Moreover, the fluxes $\mathcal{F}^{(1)}$, $\mathcal{F}^{(2)}$ and B are given by

$$\mathcal{F}^{(1)}(\rho_{ij}^s, \rho_{i+1j}^s) = \begin{cases} \rho_{ij}^s \overline{v}_{ij}^{(1)} & \text{if } \overline{v}_{ij}^{(1)} \geq 0, (i+1,j) \in \mathcal{I}_\Omega \setminus \mathcal{I}_{\partial\Omega}, \\ \rho_{i+1j}^s \overline{v}_{ij}^{(1)} & \text{if } \overline{v}_{ij}^{(1)} < 0, (i,j) \in \mathcal{I}_\Omega \setminus \mathcal{I}_{\partial\Omega}, \\ 0 & \text{otherwise}, \end{cases}$$

$$\mathcal{F}^{(2)}(\tilde{\rho}_{ij}^s, \tilde{\rho}_{ij+1}^s) = \begin{cases} \tilde{\rho}_{ij}^s \overline{v}_{ij}^{(2)} & \text{if } \overline{v}_{ij}^{(2)} \geq 0, (i,j+1) \in \mathcal{I}_\Omega \setminus \mathcal{I}_{\partial\Omega}, \\ \tilde{\rho}_{ij+1}^s \overline{v}_{ij}^{(2)} & \text{if } \overline{v}_{ij}^{(2)} < 0, (i,j) \in \mathcal{I}_\Omega \setminus \mathcal{I}_{\partial\Omega}, \\ 0 & \text{otherwise}, \end{cases}$$

$$B(\rho_{i-1j}^{s+1}, \rho_{i+1j}^{s+1}, \rho_{ij}^{s+1}, \rho_{ij-1}^{s+1}, \rho_{ij+1}^{s+1}) = b_{i-1j}^{s+1} + b_{i+1j}^{s+1} - 4b_{ij}^{s+1} + b_{ij-1}^{s+1} + b_{ij+1}^{s+1},$$

where $\quad \overline{v}(x_{ij}) = \overline{v}_{ij}, \qquad \overline{v}_{ij} = 0 \ \forall (i,j) \in \mathcal{I}_{\partial\Omega}$ and $\qquad b_{ij}^{s+1} = b(\rho_{ij}^{s+1}) \qquad$ with $b(\rho) = \int_0^\rho CzH(z - \rho_{crit})\, dz$. The Heaviside function H is smoothly approximated and the time step restriction for the numerical simulations is given by the CFL condition of the hyperbolic part

$$\Delta t^c \leq \min_{(i,j)} \frac{1}{\frac{|\overline{v}_{ij}^{(1)}|}{\Delta x} + \frac{|\overline{v}_{ij}^{(2)}|}{\Delta x}}, \tag{9}$$

compare Holden et al. 2000; Weissen et al. 2021. We denote the vector of density values $\boldsymbol{\rho} = (\rho_{ij}^s)_{(i,j,s) \in \mathcal{I}_\Omega \times \{0,\dots N_t^c\}}$. It is the discretized solution (8) of the macroscopic Eq. (2) which depends on a given control u. The vectors containing intermediate density values $\tilde{\boldsymbol{\rho}}, \overline{\boldsymbol{\rho}}$ and Lagrange parameters $\boldsymbol{\mu}, \tilde{\boldsymbol{\mu}}, \overline{\boldsymbol{\mu}}$ used below are defined analogously.

2.2.2 Solving the coarse model optimization problem

Next, we turn to the solution of the coarse-scale optimization problem. The construction of a solution to this problem is paramount to the space mapping algorithm. We provide a short discussion on the adjoint method for the optimization problem (3) before we specify the macroscopic adjoints.

First order optimality system Let $J(u, y(u))$ be an objective function which depends on the given control u. We wish to solve the optimization problem (3) and apply a descent algorithm. In a descent algorithm, a current iterate u_k, is updated in the direction of descent of the objective function J until the first order optimality condition is satisfied. An efficient way to compute the first order optimality conditions is based on the adjoint, which we recall in the following. Let the Lagrangian function be defined as

$$L(u, y(u)) = J(u, y(u)) + \mu^T E(u, y(u)),$$

where μ is called the Lagrange multiplier.

Solving $\mathrm{d}L = 0$ yields the first order optimality system

(i) $E(u, y(u)) = 0,$
(ii) $(\partial_y E(u, y(u))^T)\mu = -(\partial_y J(u, y(u)))^T,$
(iii) $\frac{\mathrm{d}}{\mathrm{d}u} J(u, y(u)) = \partial_u J(u, y(u)) + \mu^T \partial_u E(u, y(u)) = 0.$

For nonlinear systems it is difficult to solve the coupled optimality system (i)–(iii) all at once. We therefore proceed iteratively: for the computation of the total derivative $\frac{\mathrm{d}}{\mathrm{d}u} J(u, y(u))$, the system $E(u, y(u)) = 0$ is solved forward in time. Then, the information of the forward solve is used to solve the adjoint system (*ii*) backwards in time. Lastly, the gradient is obtained from the adjoint state and the objective function derivative.

Nonlinear conjugate gradient method We use a nonlinear conjugate gradient method Dai and Yuan 1999; Fletcher and Reeves 1964 within our descent algorithm to update the iterate as follows

$$d_k = -\nabla J\big(u_k, y(u_k)\big) + \hat{\beta}_{k-1}\hat{d}_{k-1}, \qquad u_{k+1} = u_k + \sigma_k d_k. \tag{10}$$

The step size σ_k is chosen such that it satisfies the Armijo-Rule Hinze et al. (2009); Nocedal and Wright (2006)

$$J\big(u_k + \sigma_k d_k, y(u_k + \sigma_k d_k)\big) - J(u_k, y(u_k)) \le \sigma_k c_1 \nabla J(u_k, y(u_k))^T d_k, \tag{11}$$

and the standard Wolfe condition Nocedal and Wright 2006

$$\nabla J(u_k + \sigma_k d_k, y(u_k + \sigma_k d_k))^T d_k \ge c_2 \nabla J(u_k, y(u_k))^T d_k, \tag{12}$$

with $0 < c_1 < c_2 < 1$. We start from $\sigma_k = 1$ and cut the step size in half until (11)–(12) are satisfied. The parameter $\hat{\beta}_k$ is given by

$$\hat{\beta}_k = \frac{\|\nabla J(u_{k+1}, y(u_{k+1}))\|^2}{d_k^T \hat{d}_k} \text{ with } \hat{d}_k = \nabla J(u_{k+1}, y(u_{k+1})) - \nabla J(u_k, y(u_k)),$$

which together with conditions (11)–(12) ensures convergence to a minimizer Dai and Yuan 1999. We refer to this method as adjoint method (AC). In the following we apply this general strategy to our macroscopic equation.

Macroscopic Lagrangian We consider objective functions depending on the density, i.e., $J^c(u, \rho)$. The discrete Lagrangian $L = L(u, \rho, \tilde{\rho}, \overline{\rho}, \mu, \tilde{\mu}, \overline{\mu})$ is given by

$$L = J^c(u, \rho) + \sum_{s=0}^{N_t^c} \sum_{i=1}^{N_x} \sum_{j=1}^{N_x} \mu_{ij}^s \left(\frac{\tilde{\rho}_{ij}^s - \rho_{ij}^s}{\Delta t^c} + \frac{\mathcal{F}_{ij}^{(1),s,+} - \mathcal{F}_{ij}^{(1),s,-}}{\Delta x} \right)$$

$$+ \sum_{s=0}^{N_t^c} \sum_{i=1}^{N_x} \sum_{j=1}^{N_x} \tilde{\mu}_{ij}^s \left(\frac{\overline{\rho}_{ij}^s - \tilde{\rho}_{ij}^s}{\Delta t^c} + \frac{\mathcal{F}_{ij}^{(2),s,+} - \mathcal{F}_{ij}^{(2),s,-}}{\Delta x} \right) \quad (13)$$

$$+ \sum_{s=0}^{N_t^c} \sum_{i=1}^{N_x} \sum_{j=1}^{N_x} \bar{\mu}_{ij}^s \left(\frac{\rho_{ij}^{s+1} - \overline{\rho}_{ij}^s}{\Delta t^c} - \frac{B_{ij}^{s+1}}{(\Delta x)^2} \right).$$

We differentiate the Lagrangian with respect to ρ_{ij}^s

$$\partial_{\rho_{ij}^s} L = \partial_{\rho_{ij}^s} J^c(u, \rho)$$

$$- \mu_{ij}^s \left(\frac{1}{\Delta t^c} - \frac{\partial_{\rho_{ij}^s} \mathcal{F}_{ij}^{(1),s,+}}{\Delta x} + \frac{\partial_{\rho_{ij}^s} \mathcal{F}_{ij}^{(1),s,-}}{\Delta x} \right)$$

$$+ \mu_{i-1j}^s \frac{\partial_{\rho_{ij}^s} \mathcal{F}_{i-1j}^{(1),s,+}}{\Delta x} - \mu_{i+1j}^s \frac{\partial_{\rho_{ij}^s} \mathcal{F}_{i+1j}^{(1),s,-}}{\Delta x}$$

$$+ \bar{\mu}_{ij}^{s-1} \left(\frac{1}{\Delta t^c} - \frac{\partial_{\rho_{ij}^s} B_{ij}^s}{(\Delta x)^2} \right) - \bar{\mu}_{i-1j}^{s-1} \frac{\partial_{\rho_{ij}^s} B_{i-1j}^s}{(\Delta x)^2}$$

$$- \bar{\mu}_{i+1j}^{s-1} \frac{\partial_{\rho_{ij}^s} B_{i+1j}^s}{(\Delta x)^2} - \bar{\mu}_{ij-1}^{s-1} \frac{\partial_{\rho_{ij}^s} B_{ij-1}^s}{(\Delta x)^2} - \bar{\mu}_{ij+1}^{s-1} \frac{\partial_{\rho_{ij}^s} B_{ij+1}^s}{(\Delta x)^2}$$

$$= 0.$$

Rearranging terms yields

$$T^{i,j}(\overline{\mu}^{s-1}) = \overline{\mu}_{ij}^{s-1} - \frac{\Delta t^c}{(\Delta x)^2} \left(\overline{\mu}_{i-1j}^{s-1} \partial_{\rho_{ij}^s} B_{i-1j}^s + \overline{\mu}_{i+1j}^{s-1} \partial_{\rho_{ij}^s} B_{i+1j}^s \right.$$

$$\left. + \overline{\mu}_{ij}^{s-1} \partial_{\rho_{ij}^s} B_{ij}^s + \overline{\mu}_{ij-1}^{s-1} \partial_{\rho_{ij}^s} B_{ij-1}^s + \overline{\mu}_{ij+1}^{s-1} \partial_{\rho_{ij}^s} B_{ij+1}^s \right)$$

$$= -\Delta t^c \partial_{\rho_{ij}^s} J^c(u, \rho) + \mu_{ij}^s \left(1 - \lambda \partial_{\rho_{ij}^s} \mathcal{F}_{ij}^{(1),s,+} + \lambda \partial_{\rho_{ij}^s} \mathcal{F}_{ij}^{(1),s,-} \right)$$

$$- \mu_{i-1j}^s \lambda \partial_{\rho_{ij}^s} \mathcal{F}_{i-1j}^{(1),s,+} + \mu_{i+1j}^s \lambda \partial_{\rho_{ij}^s} \mathcal{F}_{i+1j}^{(1),s,-}.$$

Using $\partial_{\rho_{ij}^s} B_{i-1j}^s = \partial_{\rho_{ij}^s} B_{i+1j}^s = \partial_{\rho_{ij}^s} B_{ij-1}^s = \partial_{\rho_{ij}^s} B_{ij+1}^s = k(\rho_{ij}^s)$ and $\partial_{\rho_{ij}^s} B_{ij}^s = -4k(\rho_{ij}^s)$ on the left-hand side and (17)–(18), see "Appendix B", on the right-hand side, leads to

$$T^{i,j}(\overline{\mu}^{s-1}) = \overline{\mu}_{ij}^{s-1} - \frac{\Delta t^c}{(\Delta x)^2} k(\rho_{ij}^s) \left(\overline{\mu}_{i-1j}^{s-1} + \overline{\mu}_{i+1j}^{s-1} - 4\overline{\mu}_{ij}^{s-1} + \overline{\mu}_{ij-1}^{s-1} + \overline{\mu}_{ij+1}^{s-1} \right)$$

$$\overset{(17),(18)}{=} -\Delta t^c \partial_{\rho_{ij}^s} J^c(u, \rho) + \mu_{ij}^s$$

$$- \lambda \left(\left(\mu_{ij}^s - \mu_{i+1j}^s \right) \partial_{\rho_{ij}^s} \mathcal{F}_{ij}^{(1),s,+} - \left(\mu_{ij}^s - \mu_{i-1j}^s \right) \partial_{\rho_{ij}^s} \mathcal{F}_{ij}^{(1),s,-} \right).$$

This is solved backward in time to obtain the Lagrange parameter $(\mu_{ij}^{s-1})_{(i,j) \in \mathcal{I}_\Omega}$. Note that the above expression $T(\overline{\mu}^{s-1}) = \left(T^{i,j}(\overline{\mu}^{s-1}) \right)_{(i,j) \in \mathcal{I}_\Omega}$ defines a coupled system for the Lagrange parameter of time step $s - 1$ in space and has to be solved in each time step. This system arises from the implicit treatment of the diffusion term in the forward system (8). It is the main difference to adjoints for purely hyperbolic equations where the Lagrange parameters in step $s - 1$ in the backward system are simply obtained as a convex combination of those from step s, see Erbrich et al. 2018. Proceeding further, we differentiate the Lagrangian with respect to $\tilde{\rho}_{ij}^s$ to get

$$\partial_{\tilde{\rho}_{ij}^s} L = \frac{\mu_{ij}^s}{\Delta t^c} - \tilde{\mu}_{ij}^s \left(\frac{1}{\Delta t^c} - \frac{\partial_{\tilde{\rho}_{ij}^s} \mathcal{F}_{ij}^{(2),s,+}}{\Delta x} + \frac{\partial_{\tilde{\rho}_{ij}^s} \mathcal{F}_{ij}^{(2),s,-}}{\Delta x} \right)$$

$$+ \tilde{\mu}_{ij-1}^s \frac{\partial_{\tilde{\rho}_{ij}^s} \mathcal{F}_{ij-1}^{(2),s,+}}{\Delta x} - \tilde{\mu}_{ij+1}^s \frac{\partial_{\tilde{\rho}_{ij}^s} \mathcal{F}_{ij+1}^{(2),s,-}}{\Delta x}$$

$$= 0.$$

Again, rearranging terms yields

$$\mu_{ij}^s = \tilde{\mu}_{ij}^s \left(1 - \lambda \partial_{\tilde{\rho}_{ij}^s} \mathcal{F}_{ij}^{(2),s,+} + \lambda \partial_{\tilde{\rho}_{ij}^s} \mathcal{F}_{ij}^{(2),s,-} \right)$$

$$- \tilde{\mu}_{ij-1}^s \lambda \partial_{\tilde{\rho}_{ij}^s} \mathcal{F}_{ij-1}^{(2),s,+} + \tilde{\mu}_{ij+1}^s \lambda \partial_{\tilde{\rho}_{ij}^s} \mathcal{F}_{ij+1}^{(2),s,-}$$

$$\overset{(19),(20)}{=} \tilde{\mu}_{ij}^s - \lambda \left(\left(\tilde{\mu}_{ij}^s - \tilde{\mu}_{ij+1}^s \right) \partial_{\tilde{\rho}_{ij}^s} \mathcal{F}_{ij}^{(2),s,+} - \left(\tilde{\mu}_{ij}^s - \tilde{\mu}_{ij-1}^s \right) \partial_{\tilde{\rho}_{ij}^s} \mathcal{F}_{ij}^{(2),s,-} \right).$$

Finally, we differentiate the Lagrangian with respect to $\overline{\rho}_{ij}^s$ to obtain

$$\partial_{\overline{\rho}_{ij}^s} L = \frac{\tilde{\mu}_{ij}^s}{\Delta t^c} - \frac{\overline{\mu}_{ij}^s}{\Delta t^c} = 0 \qquad \Rightarrow \qquad \tilde{\mu}_{ij}^s = \overline{\mu}_{ij}^s.$$

The equality of the Lagrange parameters $\tilde{\mu}, \overline{\mu}$ stems from the fact that the diffusion is solved implicitly in the forward system (8)[1]. In the next section, we consider the diffusion coefficient C as control for the macroscopic system, $u = C$. In this case, the derivative of the Lagrangian with respect to the control reads

---

[1] Note that these parameters would be different if the diffusion was solved explicitly using values of $\overline{\rho}^s$ instead of ρ^{s+1} in the diffusion operator B of (8).

$$\partial_C L = \sum_{s=0}^{N_t^c} \sum_{i=1}^{N_x} \sum_{j=1}^{N_x} -\frac{1}{C} \frac{\bar{\mu}_{ij}^s}{(\Delta x)^2} \left(b_{i-1j}^{s+1} + b_{i+1j}^{s+1} - 4b_{ij}^{s+1} + b_{ij-1}^{s+1} + b_{ij+1}^{s+1} \right).$$

3 Validation of the approach

To validate the proposed approach, we consider a toy problem and compare the results of the space mapping method to optimal solutions computed directly on the microscopic level. For the toy problem, we control the potential strength A of the microscopic model. The macroscopic analogue is the diffusion coefficient C.

3.1 Discrete microscopic adjoint

Let $N_t^f \in \mathbb{N}$ and $\Delta t^f \in \mathbb{R}$ be the number of time steps and the time step size, respectively. For simplicity, we normalize the particle mass in our example, i.e. $m = 1$. We discretize the fine, microscopic model (1) in time to obtain

$$x_i^{s+1} = x_i^s + \Delta t^f v_i^s, \qquad v_i^{s+1} = v_i^s + \Delta t^f \left(G(x_i^{s+1}, v_i^{s+1}) + A \sum_{j \neq i} F(x_i^{s+1}, x_j^{s+1}) \right)$$

for $s = 1, \ldots, N_t^f$. We denote

$$\boldsymbol{x} = (x_i^s)_{(i,s)\in\{1,\ldots,N\}\times\{0,\ldots,N_t^f\}} \quad \text{and} \quad \boldsymbol{v} = (v_i^s)_{(i,s)\in\{1,\ldots,N\}\times\{0,\ldots,N_t^f\}}.$$

Furthermore, let $J^f(u, \boldsymbol{x})$ be the microscopic objective function. The microscopic Lagrange function $L(u, \boldsymbol{x}, \boldsymbol{v}, \boldsymbol{\mu}, \tilde{\boldsymbol{\mu}}, \overline{\boldsymbol{\mu}}, \hat{\boldsymbol{\mu}})$ is then given by

$$L = J^f(u, \boldsymbol{x}) + \sum_{s=0}^{N_t^f} \sum_{i=1}^{N} \mu_i^s \left(\frac{x_i^{(1),s+1} - x_i^{(1),s}}{\Delta t^f} - v_i^{(1),s} \right)$$

$$+ \sum_{s=0}^{N_t^f} \sum_{i=1}^{N} \tilde{\mu}_i^s \left(\frac{x_i^{(2),s+1} - x_i^{(2),s}}{\Delta t^f} - v_i^{(2),s} \right)$$

$$+ \sum_{s=0}^{N_t^f} \sum_{i=1}^{N} \overline{\mu}_i^s \left(\frac{v_i^{(1),s+1} - v_i^{(1),s}}{\Delta t^f} - G_i^{(1),s+1} - A \sum_{j \neq i} F_{ij}^{(1),s+1} \right) \qquad (14)$$

$$+ \sum_{s=0}^{N_t^f} \sum_{i=1}^{N} \hat{\mu}_i^s \left(\frac{v_i^{(2),s+1} - v_i^{(2),s}}{\Delta t^f} - G_i^{(2),s+1} - A \sum_{j \neq i} F_{ij}^{(2),s+1} \right),$$

where

$$G_i^{(l),s}(x_i^s, v_i^s) = -\frac{v_i^{(l),s} - \bar{v}^{(l)}(x_i^s)}{\tau},$$

$$F_{ij}^{(l),s}(x_i^s, x_j^s) = \begin{cases} \frac{b_F\left(\|x_i^s - x_j^s\|_2 - 2R\right)^2}{\|x_i^s - x_j^s\|_2}\left(x_i^{(l),s} - x_j^{(l),s}\right) & \text{if } \|x_i^s - x_j^s\|_2 < 2R, \\ 0 & \text{otherwise,} \end{cases}$$

for $l = 1, 2$. The details of the derivatives of the force terms and the computation of the adjoint state can be found in "Appendix C". Moreover, the derivative of the Lagrangian with respect to the control $u = A$ reads

$$\partial_A L = -\sum_{s=0}^{N_t^f}\sum_{i=1}^{N}\sum_{j\neq i}\left(\bar{\mu}_i^s F_{ij}^{(1),s+1} + \hat{\mu}_i^s F_{ij}^{(2),s+1}\right).$$

3.2 Comparison of space mapping to direct optimization

We apply ASM and the direct optimization approach AC to the optimization problem (3). In each iteration k of the adjoint method for the fine model, a computation of the gradient ∇J^f for the stopping criterion as well as several objective function and gradient evaluations for the computation of the step size σ_k are required. These evaluations are (mostly) shifted to the coarse model in ASM. Let $\Omega = [-5, 5]^2$ be the domain and $\bar{v}(x) = -x$ the velocity field of our toy example. We investigate whether the macroscopic model is an appropriate coarse model in the space mapping technique. For the microscopic interactions, we use the force term (4) with $b_F = 1/R^5$. Without interaction forces, $A = 0$, all particles are transported to the center of the domain $(x^{(1)}, x^{(2)}) = (0, 0)$ within finite time. Certainly, they overlap after some time. With increasing interaction parameter, i.e., increasing A, particles encounter stronger forces as they collide. Therefore, scattering occurs and the spatial spread increases. We penalize the spatial spread of the particle ensemble at $t = T$ in the microscopic model, leading to a cost

$$J^f(A, \boldsymbol{x}) = \frac{1}{N}\sum_i^N \langle x_i^{N_t^f}, x_i^{N_t^f}\rangle,$$

and the objective function derivative with respect to the state variables x_i is given by

$$\partial_{x_i^{(l),s}} J^f(A, \boldsymbol{x}) = \begin{cases} \left(\frac{1}{N}\sum_i \langle x_i^{N_t^f}, x_i^{N_t^f}\rangle - \omega_*\right)\frac{2x_i^{(l),s}}{N} & \text{if } s = N_t^f, \\ 0 & \text{otherwise.} \end{cases}$$

We choose A, the scaling parameter of the interaction force, as microscopic control. The coarse, macroscopic model is given by (2) and the spatial spread of the density at $t = T$ is given by

Table 1 Model parameters

T	R	N	Δx	Δt^c	Δt^f	m	b_F	τ
3	0.2	200	0.5	0.05	0.00125	1	$1/R^5$	$1/15$

$$j^c(C, \boldsymbol{\rho}) = \frac{1}{M} \sum_{(i,j)} \rho_{ij}^{N_t^c} \langle x_{ij}, x_{ij} \rangle,$$

$$\partial_{\rho_{ij}^s} J^c(C, \boldsymbol{\rho}) = \begin{cases} \frac{\langle x_{ij}, x_{ij} \rangle}{M} \left(\left(\frac{1}{M} \sum_{(i,j)} \rho_{ij}^{N_t^c} \langle x_{ij}, x_{ij} \rangle \right) - \omega_* \right) & \text{if } s = N_t^c, \\ 0 & \text{otherwise,} \end{cases}$$

where M is the total mass, i.e., $M = \sum_{(i,j)} \rho_{ij}^0 (\Delta x)^2$. We consider the parameters in Table 1. The macroscopic model requires spatial and time step sizes. Here, we choose a large spatial step size Δx such that the macroscopic model can be computed sufficiently fast. The time step size Δt^c satisfies the CFL condition (9). The simulation of the microscopic model is more involved. For the simulation of pedestrian dynamics, compare e.g. Helbing et al. 2000, the parameter τ is chosen small such that deviations of individual velocities from the velocity field $\bar{v}$ are corrected fast and the parameter b_F in the interaction force is chosen large such that the interaction force becomes dominant when pedestrians are close together. Since the particles move to the center of the domain and the interaction force gets dominant, when particles are close together, we have to choose a small time step size Δt^f. (Fig. 2)

Two particle collectives with $N/2 = 100$ particles are placed in the domain, see Fig. 3a. The macroscopic representation (7) of the particle groups is shown in Fig. 3b. We set box constraints on the controls $0 \leq A, C \leq 10$ and compare the number of iterations of the two approaches to obtain a given accuracy[2] of $\|J^f(u_k, \boldsymbol{x})\|_2 < 10^{-5}$. The step sizes σ_k for AC are chosen such that they satisfy the Armijo Rule and standard Wolfe condition (11)–(12) with $c_1 = 0.01, c_2 = 0.9$. If an iterate violates the box constraint, it is projected into the feasible set.

In the space mapping algorithm, the parameter extraction $\mathcal{T}(u_k)$ is the solution of an optimization problem in the coarse model space, see Definition 2.1. In each iteration, the parameter extraction identifies the coarse model control C which matches best the microscopic model with control A by solving an optimization in the coarse model. The optimization is solved via adjoint calculus with c_1, c_2 as chosen above and $u_{start} = \mathcal{T}(u_{k-1})$, which we expect to be close to $\mathcal{T}(u_k)$. Further, to determine the step size σ_k for the control update, we consider step sizes such that $u_{k+1} = u_k + \sigma_k d_k$ satisfies $\|\mathcal{T}(u_{k+1}) - u_*^c\|_2 < \|\mathcal{T}(u_k) - u_*^c\|_2$ and thus decreases the distance of the parameter extraction to the coarse model optimal control from one space mapping iteration to the next. The optimization results and computation times (obtained as average computation time of 20 runs on an Intel(R) Core(TM) i7-6700 CPU 3.40 GHz, 4 Cores) for target values $\omega_* \in \{1, 2, 3\}$ are

---

[2] To ensure comparability of the two optimization approaches, we use the same stopping criterion $\|J^f(u_k, \boldsymbol{x})\|_2 < tolerance = 10^{-5}$.
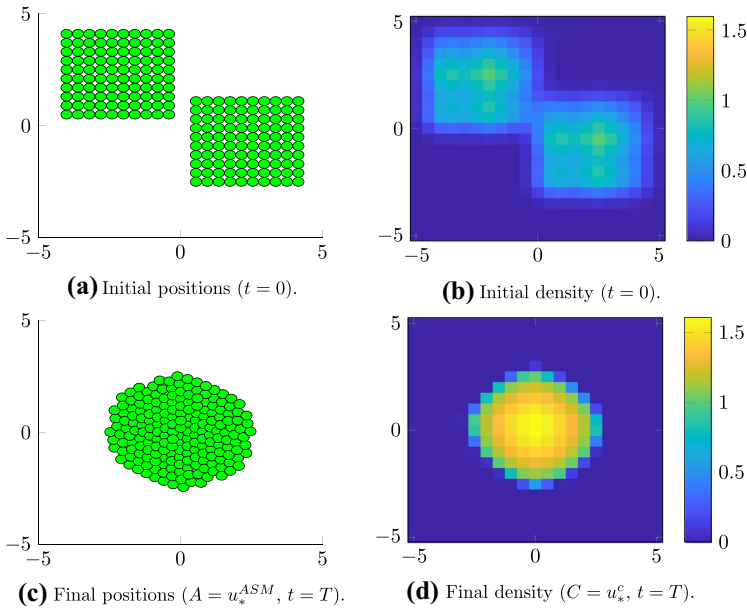
Table 2 Aggressive Space Mapping (ASM) vs. Adjoint Calculus (AC)

	$\omega_* = 1$	$\omega_* = 2$	$\omega_* = 3$
u_0	8	8	8
u_*^{AC}	0.1851	0.8884	3.6720
u_*^{ASM}	0.1829	0.8855	3.6262
$u_1^{ASM} = u_*^c$	0.1218	0.8740	3.8211
u_2^{ASM}	0.1767	0.8994	3.6262
u_3^{ASM}	0.1829	0.8855	-
$J^f\left(u_*^{AC}, \boldsymbol{x}\right)$	$8.31 \cdot 10^{-6}$	$5.00 \cdot 10^{-7}$	$4.40 \cdot 10^{-6}$
$J^c\left(u_*^c, \boldsymbol{\rho}\right)$	$3.78 \cdot 10^{-7}$	$7.51 \cdot 10^{-7}$	$1.12 \cdot 10^{-7}$
$J^f(u_*^{ASM}, \boldsymbol{x})$	$8.18 \cdot 10^{-7}$	$2.49 \cdot 10^{-6}$	$1.36 \cdot 10^{-6}$
$\bar{t}_A C$ [s]	93.27	103.18	101.21
$\bar{t}_A SM$ [s]	84.51	62.49	238.57
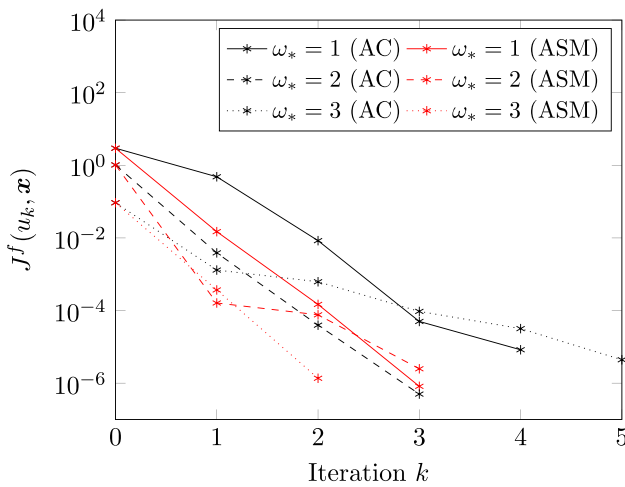
Fig. 2 Objective function value of iterates

compared in Table 2. Both optimization approaches start far from the optima at $u_0 = 8$. Optimal controls u_*^{AC} and u_*^{ASM} closely match. The objective function evaluations $J^f(u_*^{AC}, \boldsymbol{x})$, $J^c(u_*^c, \boldsymbol{\rho})$ describe the accuracy at which the fine and coarse model control problem are solved, respectively. $J^f(u_*^{ASM}, \boldsymbol{x})$ denotes the accuracy of the space mapping optimal control when the control is plugged into the fine model and the fine model objective function is evaluated. Note that the ASM approach in general does not ensure a descent in the microscopic objective function value $J^f(u_k, \boldsymbol{x})$ during the iterative process and purely relies on the idea to reduce the distance $\|\mathcal{T}(u_k) - u_*^c\|_2$. However, ASM also generates small target values $J^f(u_*^{ASM}, \boldsymbol{x})$ and therefore validates the proposed approach. Moreover, the model responses of the optimal controls illustrate the similarity of the fine and the coarse model, see Fig. 3c–d.

(a) Initial positions ($t = 0$).

(b) Initial density ($t = 0$).

(c) Final positions ($A = u_*^{ASM}$, $t = T$).

(d) Final density ($C = u_*^c$, $t = T$).

Fig. 3 Initial conditions and space mapping solution for $\omega_* = 3$

The space mapping iteration finishes within two to three iterations and therefore needs less iterations than the pure optimization on the microscopic level here, see Fig. 2. Note that each of the space mapping iterations involves the solution of the coarse optimal control problem. Hence, the comparison of the iterations may be misleading and we consider the computation times as additional feature. It turns out that the iteration times vary and therefore this data does not allow to prioritize one of the approaches based on computational time. Obviously, the times depend on the number of particles, the space and time discretizations. However, very few iterations within the space mapping approach are needed to obtain the results. We see that the space mapping technique is validated, because we can compare ASM solutions to solutions computed with the microscopic adjoints. Next, we apply ASM to scenarios in bounded domain, where solutions cannot be computed on the microscopic level anymore due to the boundary interactions.

4 Space mapping in bounded domains

In the following, we consider problems with dynamics restricted to a spatial domain with boundaries. For the microscopic simulations we add artificial boundary behaviour, tailored for each application, to the ODEs.

4.1 Crowd dynamics

Interesting dynamics evolve in the modeling of pedestrian dynamics, when pedestrian groups cross at intersections, crowds pass through corridors or a bottleneck, or they try to reach the location of a staircase, an elevator or escalators in case of an emergency. People trying to move can be injured when they are pushed into obstacles in their way Helbing et al. 2002. To apply the space mapping technique to pedestrian dynamics, we consider a scenario similar to the evacuation of N individuals from a domain with obstacles. The goal is to gather as many individuals as possible at an assembly point $x_s \in \Omega \subset \mathbb{R}^2$ up to the time T. The control is the assembly point $x_s = (x_s^{(1)}, x_s^{(2)})$. We model this task with the help of the following cost functions

$$j^f(x_s, \boldsymbol{x}) = \frac{1}{N} \sum_{(i)} \langle x_i^{N_t^f} - x_s, x_i^{N_t^f} - x_s \rangle,$$

$$j^c(x_s, \boldsymbol{\rho}) = \frac{1}{M} \sum_{(i,j)} \rho_{ij}^{N_t^c} \langle x_{ij} - x_s, x_{ij} - x_s \rangle,$$

for the fine and coarse model, respectively. They measure the spread of the crowd at time $t = T$ with respect to the location of the source.

The velocity $\bar{v}(x)$ is based on the solution to the eikonal equation with point source x_s. In more detail, we solve the eikonal equation

$$|\nabla T(x)| = \frac{1}{f(x)}, \qquad x \in \Omega, \qquad T(x_s) = 0,$$

where $T(x)$ is the minimal amount of time required to travel from x to x_s and $f(x)$ is the speed of travel. We choose $f(x) = 1$ and set the velocity field to

$$\bar{v}(x) = \frac{\nabla T(x)}{\|\nabla T(x)\|_2} \min\{\|x - x_s\|_2, 1\}. \tag{15}$$

In this way, the velocity vectors point into the direction of the gradient of the solution to the eikonal equation and the speed depends on the distance of the particle to x_s. The particles are expected to slow down when approaching x_s and the maximum velocity is bounded $\|\bar{v}(x)\|_2 \leq 1$. The solution to the eikonal equation on the 2-D cartesian grid is computed using the fast marching algorithm implemented in C with Matlab interface[3]. The travel time isoclines of the eikonal equation and the corresponding velocity field are illustrated in Fig. 4. Note that we have to set the travel time inside the boundary to a finite value to obtain a smooth velocity field.

The derivative of the macroscopic Lagrangian (13) with respect to the location of the point source, $u = x_s$, is given by

---

[3] http://www-m3.ma.tum.de/Software/FMWebHome#Fast_Eikonal_Solver_in_2D_and_3D__40with_MATLAB_interface_41 by Volkmar Bornemann and Christian Ludwig.
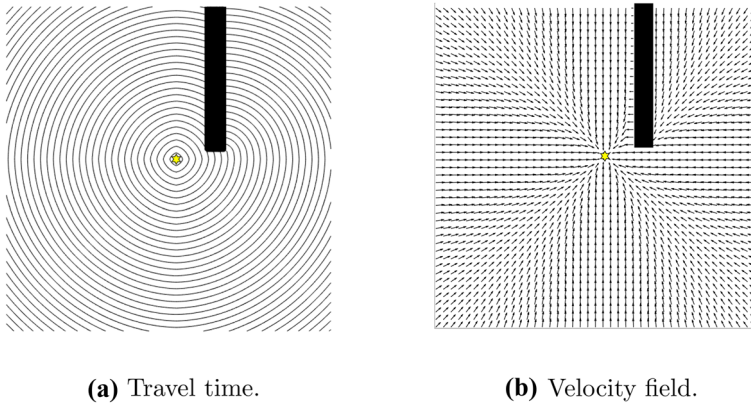
(a) Travel time.

(b) Velocity field.

Fig. 4 Solution of the eikonal equation in a bounded domain

$$
\partial_{x_s^{(l)}} L = \sum_{s=0}^{N_t^c} \sum_{i=1}^{N_x} \sum_{j=1}^{N_x} \frac{\mu_{ij}^s}{\Delta x} \left(\partial_{x_s^{(l)}} \mathcal{F}_{ij}^{(1),s,+} - \partial_{x_s^{(l)}} \mathcal{F}_{ij}^{(1),s,-} \right)
$$
$$
+ \sum_{s=0}^{N_t^c} \sum_{i=1}^{N_x} \sum_{j=1}^{N_x} \frac{\tilde{\mu}_{ij}^s}{\Delta x} \left(\partial_{x_s^{(l)}} \mathcal{F}_{ij}^{(2),s,+} - \partial_{x_s^{(l)}} \mathcal{F}_{ij}^{(2),s,-} \right),
$$

where

$$
\partial_{x_s^{(l)}} \mathcal{F}_{ij}^{(1),s,+} = \begin{cases} \rho_{ij}^s \partial_{x_s^{(l)}} \overline{v}_{ij}^{(1)} & \text{if } \overline{v}_{ij}^{(1)} \geq 0, (i+1,j) \in \mathcal{I}_\Omega \setminus \mathcal{I}_{\partial\Omega}, \\ \rho_{i+1j}^s \partial_{x_s^{(l)}} \overline{v}_{ij}^{(1)} & \text{if } \overline{v}_{ij}^{(1)} < 0, (i,j) \in \mathcal{I}_\Omega \setminus \mathcal{I}_{\partial\Omega}, \qquad l = 1, 2, \\ 0 & \text{otherwise}, \end{cases}
$$

$$
\partial_{x_s^{(l)}} \mathcal{F}_{ij}^{(1),s,-} = \begin{cases} \rho_{i-1j}^s \partial_{x_s^{(l)}} \overline{v}_{i-1j}^{(1)} & \text{if } \overline{v}_{i-1j}^{(1)} \geq 0, (i,j) \in \mathcal{I}_\Omega \setminus \mathcal{I}_{\partial\Omega}, \\ \rho_{ij}^s \partial_{x_s^{(l)}} \overline{v}_{i-1j}^{(1)} & \text{if } \overline{v}_{i-1j}^{(1)} < 0, (i-1,j) \in \mathcal{I}_\Omega \setminus \mathcal{I}_{\partial\Omega}, \qquad l = 1, 2, \\ 0 & \text{otherwise}, \end{cases}
$$

and $\partial_{x_s^{(l)}} \mathcal{F}_{ij}^{(2),s,+}, \partial_{x_s^{(l)}} \mathcal{F}_{ij}^{(2),s,-}$ are defined analogously.

To obtain the partial derivatives $\partial_{x_s^{(l)}} \overline{v}_{ij}^{(k)}$, the travel-time source derivative of the eikonal equation is required. It is approximated numerically with finite differences

$$
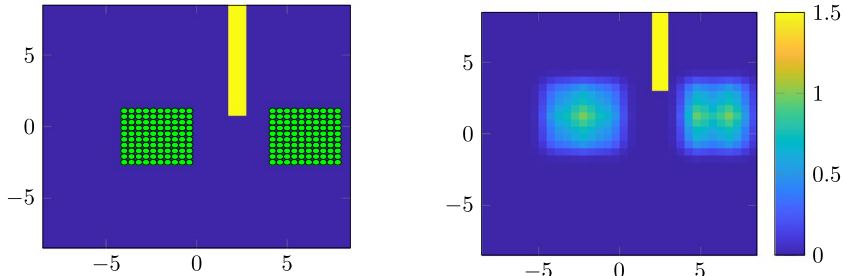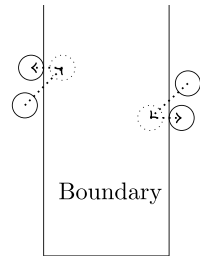\partial_{x_s^{(l)}} \overline{v}_{ij}^{(k)} \approx \frac{\overline{v}_{ij}^{(k)}(x_s + \Delta x e^{(l)}) - \overline{v}_{ij}^{(k)}(x_s - \Delta x e^{(l)})}{2\Delta x}, \qquad k = 1, 2,
$$

where $e^{(1)} = (1, 0)^T, e^{(2)} = (0, 1)^T$ denote the unit vectors.

4.1.1 Discussion of the numerical results

To investigate the robustness of the space mapping algorithm, we consider different obstacles in the microscopic and macroscopic setting. Let $\Omega = [-8, 8]^2$

Fig. 5 Reflection at the boundary

(a) Microscopic domain and initial positions x^0. (b) Macroscopic domain and initial density ρ^0.

Fig. 6 Initial conditions with $gap = 2$

be the domain. For the microscopic model we define an internal boundary $2 \leq x^{(1)} \leq 3, 1 \leq x^{(2)} \leq 8$, see Fig. 6a. For the macroscopic setting the obstacle is shifted by $gap \geq 0$ in the $x^{(2)}$-coordinate. Additionally, we shift the initial density with the same gap, see Fig. 6b. It is interesting to see whether the space mapping technique is able to recognize the linear shift between the microscopic and the macroscopic model. This is not trivial due to the non-linearities in the models and the additional non-linearities induced by the boundary interactions. Macroscopically, we use the zero flux conditions (6) at the boundary. Microscopically, a boundary correction is applied, that means, a particle which would hypothetically enter the boundary is reflected into the domain, see Fig. 5.

For computational simplicity, we restrict the admissible set of the controls

$$\mathcal{U}_{ad}^f = \mathcal{U}_{ad}^c = [-8, 2] \times [-8, 8],$$

i.e., the point source is located to the left-hand side of the obstacle.

The velocity $\bar{v}(x)$, given by (15), is restricted to the grid with spatial step sizes $\Delta x = 0.5$ for the macroscopic model. To obtain the velocity field on the grid, the source location $x_s \in \mathcal{C}_{ij}$ is thereby projected to the cell center of the corresponding cell

$$P(x_s) = x_{ij}, \qquad x_s \in \mathcal{C}_{ij}. \tag{16}$$

The continuous velocity field of the microscopic model is approximated by the eikonal solution on a grid with smaller grid size. (Fig. 6)

Table 3 Iterates of ASM

gap	Iteration	u_k	$j^f(u_k, \boldsymbol{x})$	$\mathcal{T}(u_k)$	$j^c(\mathcal{T}(u_k), \boldsymbol{\rho})$
0	$k = 1$	[1.5, –0.5]	3.0652	[1.5, –0.5]	3.0218
1	$k = 1$	[1.5, 0.5]	3.5725	[1.5, 1.5]	3.7905
	$k = 2$	[1.5, –0.5]	3.0652	[1.5, 0.5]	3.0218
2	$k = 1$	[1.5, 1.5]	4.8059	[1.5, 3]	4.4370
	$k = 2$	[1.5, 0]	3.2800	[1.5, 2]	3.3058
	$k = 3$	[1.5, –0.5]	3.0652	[1.5, 1.5]	3.0218
3	$k = 1$	[1.5, 2.5]	7.1550	[1.5, 5.5]	8.2927
	$k = 2$	[1.5, –0.5]	3.0652	[1.5, 2.5]	3.0218

We choose the parameters from Sect. 3.2, Table 1 except for T which is set to $T = 5$. Moreover, we consider $A, C = 0.87$ for which the macroscopic and microscopic model behavior match well in the situation without boundary interactions, see Table 1 in Sect. 3.1.

We apply the space mapping method to the described scenario with $gap \in \{0, 1, 2, 3\}$. Due to the grid approximation, we formally move from continuous optimization problems to discrete ones which we approximately solve by applying ASM (and AC for the parameter extraction within ASM) for continuous optimization and project each iterate to the grid using (16). In general, due to the grid approximation we cannot ensure that arbitrarily small stepsizes $\sigma_k \geq 0$ exist for which the Armijo condition is satisfied in the parameter extraction with $c_1 > 0$. Therefore, we choose $c_1 = 0, c_2 = 0.9$ and formally lose the convergence of our descent algorithm to a minimizer. Nevertheless, it is still ensured that the distance to the coarse model optimum in ASM is nonincreasing since the step size is chosen such that

$$\|\mathcal{T}(u_k + \sigma_k d_k) - u_*^c\|_2 \leq \|\mathcal{T}(u_k) - u_*^c\|_2$$

holds.

As starting point for the parameter extraction, we choose $u_{start} = u_*^c$ and *tolerance* is set to 10^{-5}. We remark that the parameter extraction does not have a unique solution here, therefore, providing $u_{start} = u_*^c$ as starting value is used to stipulate the parameter extraction identifying a solution $\mathcal{T}(u_k)$ near u_*^c.

The macroscopic optimal solution with the corresponding *gap* is given by $u_*^c = [1.5, - - 0.5 + gap]$, compare Table 3. For $gap = 0$, we have $\mathcal{T}(u_*^c) = u_*^c$ and the space mapping is finished at $k = 1$ since the model optima coincide. For $gap > 0$, the parameter extraction identifies a shift between the modeling hierarchies since the coarse model optimum is not optimal for the fine model. Indeed, the application of the coarse model optimal control leads to collision of the particles with the boundary and therefore delays gathering of the particles around the source location u_*^c, see Fig. 7b. The particles are spread more widely, because the boundary as a physical obstacle prevents that the crowd gathers circularly shaped around the source location u_*^c as it is the case for the macroscopic model, compare Fig. 7a . Space mapping for $gap \in \{1, 3\}$ finishes within one iteration since the parameter extraction of

(a) Density ($u = u_*^c$).   **(b)** Positions ($u = u_*^c$).   **(c)** Positions ($u = u_*^{ASM}$).
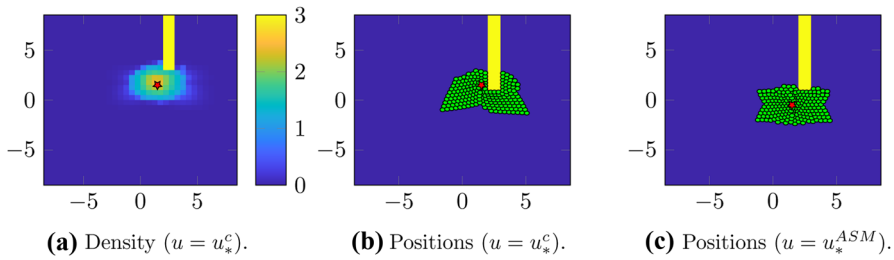
Fig. 7 Solutions of the space mapping iterates at $t = T$ with $gap = 2$

u_1 is given by $\mathcal{T}(u_1) = u_1 + [0, gap]$ and $\mathcal{T}(u_2) = u_*^c$. For $gap = 2$, the first parameter extraction underestimates the shift in $x^{(2)}$-direction and thus, two iterations are needed to obtain the optimal solution, see Table 3.

We investigated the need for additional iterations in more detail. It turned out that the behavior is caused by the discretization of the optimization problem on the macroscopic grid. We have $j^c([1.5, 3.0], \rho) = 4.4370$ and $j^c([1.5, 3.5], \rho) = 5.3451$, which indicates that the true (continuous) value $\mathcal{T}([1.5, 1.5])$ lies between the two grid values. However, the discrete optimization for the parameter extraction terminates with $\mathcal{T}([1.5, 1.5]) = [1.5, 3.0]$, because it is closer to the microscopic simulation result $j^f([1.5, 1.5], x)$. The microscopic optimal solution is shown in Fig. 7c. In comparison to the result with the control u_*^c shown in Fig. 7b, we observe in Fig. 7c that the crowd is gathered together more closely and has a smaller spread, i.e., $j^f(u_3, x) < j^f(u_1, x)$, compare Table 3.

4.2 Material flow

In the following, the control of a material flow system with a conveyor belt is considered. Similar control problems have been investigated in Erbrich et al. (2018). We use the microscopic model proposed in Göttlich et al. (2014) that describes the transport of homogeneous parts with mass m and radius R on a conveyor belt $\Omega \subset \mathbb{R}^2$ with velocity $v_T = (v_T^{(1)}, 0)^T \in \mathbb{R}^2$. The bottom friction

$$G(v) = -\gamma_b(v - v_T),$$

with bottom damping parameter $\gamma_b \geq 0$ corrects deviations of the parts' velocities from the conveyor belt velocity. The interaction force F modelling interparticle repulsion is given by

$$F(x) = \begin{cases} c_m(2R - \|x\|_2)\frac{x}{\|x\|_2} & \text{if } \|x\|_2 \leq 2R, \\ 0 & \text{otherwise,} \end{cases}$$

where $c_m > 0$ scales the interaction force and depends on the material of the parts.

We investigate the control of the material flow via the conveyor belt velocity $v_T^{(1)}$. The particles (goods) are redirected at a deflector to channel them. A common way to describe such boundary interactions is to apply obstacle forces which are modeled

	T	R	N	Δx	Δt^c	Δt^f	m	c_m	c_{obst}	γ_b
Table 4 Model parameters	1	0.012	100	0.02	$5 \cdot 10^{-3}$	$5 \cdot 10^{-3}$	0.01	200	1000	0.5

similar to the interaction force between particles Helbing and Molnár 1995. Here, we consider

$$F_{obst}(x) = \begin{cases} c_{obst}(R - \|x\|_2)\frac{x}{\|x\|_2} & \text{if } \|x\|_2 \leq R, \\ 0 & \text{otherwise,} \end{cases}$$

where x is the distance to the closest point of the boundary. Note that this is a slight variation of Helbing and Molnár (1995) as the interaction takes place with the closest boundary point only, see also Remark 3. Further note that the computation of adjoint states analogous to Sect. 3.1 can become very complicated for this boundary interaction. We therefore avoid the computation of the microscopic optimal solution u_*^f and use the proposed space mapping approach instead.

The performance evaluation used here is the number of goods in the domain Ω at time T given by

$$j^f(v_T^{(1)}, \boldsymbol{x}) = \sum_{i=1}^{N} \mathbb{1}_{\left(x_i^{N_t^f} \in \Omega\right)}.$$

The transport is modeled macroscopically with the advection-diffusion Eq. (2). The corresponding macroscopic performance evaluation is given by

$$j^c(v_T^{(1)}, \boldsymbol{\rho}) = \frac{N}{M} \sum_{(i,j):x_{ij} \in \Omega} \rho_{ij}^{N_t^c} (\Delta x)^2.$$

We apply zero-flux boundary conditions (6) for the advective and the diffusive flux at the deflector.

Remark 3 Note that if the boundary was discretized with stationary points and boundary interaction was modeled with the help of soft core interaction forces in the microscopic setting, as for example in Helbing and Molnár (1995), the model would allow for direct optimization. Nevertheless, many applications involve a huge number of (tiny) goods, for example the production of screws. The pairwise microscopic interactions would blow up the computational effort, hence it makes sense to consider a macroscopic approximation for optimization tasks.

4.2.1 Dependency on the diffusion coefficient

We investigate the robustness of the space mapping technique for different diffusion coefficients C and investigate whether variations in the diffusion coefficient affect the performance of the space mapping algorithm or the accuracy of the final result. We set $\Omega = [0, 0.64] \times [0, 0.4]$, $N = 100$, $\omega_* = 25$. The parameters R, m, c_m, c_{obst}, γ_b

Table 5 Space mapping with different diffusion coefficients C

C	0	0.1	0.5	1
u_*^c	0.6893	0.6604	0.6612	0.6615
u_*^{ASM}	0.6160	0.6163	0.6093	0.6185
$j^f\left(u_*^{ASM}, \boldsymbol{x}\right)$	25	25	25	25
Iterations	3	4	2	7

(a) Density.

(b) Positions.
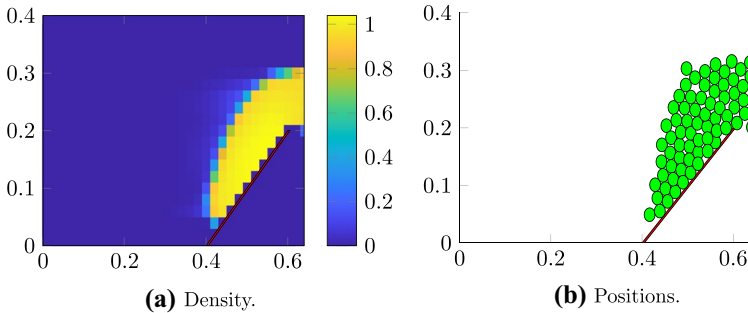
Fig. 8 Density with $u = 0.6612$ and particles with $u = 0.6093$ at $t = 0.7$

are given in Table 4 and are validated with a classical Runge-Kutta method of fourth order against real data Göttlich and Pfirsching 2018[Section 4.1]. We set $u_0 = 0.5$ and compute the space mapping solution with the ASM algorithm and stopping criterion $\|\mathcal{T}(u_k) - u_*^c\|_2 < 10^{-3}$. The diffusion constants $C \in \{0, 0.1, 0.5, 1\}$ are tested to see whether the space mapping approach is successful for all values of C. The results of our space mapping approach are summarized in Table 5. Each parameter extraction uses $u_{start} = \mathcal{T}(u_{k-1})$ and has an optimality tolerance of 10^{-5}.

For every diffusion coefficient, space mapping finishes in less than seven iterations and Table 5 shows that a microscopic optimal solution is obtained for $u_*^f \in [0.6093, 0.6185]$. In all cases, space mapping generates optimal solutions. Even for the case with $C = 0$, which is pure advection (without diffusion) in the macroscopic model, the ASM algorithm is able to identify a solution. This underlines the robustness of the space mapping algorithm and emphasizes that even a very rough depiction of the underlying process can serve as coarse model. However, the advection-diffusion equations with $C > 0$ clearly match the microscopic situation better and portray the spread of particles in front of the obstacle more realistically, see Fig. 8 for $C = 0.5$.

5 Conclusion

We proposed space mapping-based optimization algorithms for interacting particle systems. The coarse model of the space mapping is chosen to be the macroscopic approximation of the fine model that considers every single particle. The algorithm is validated

with the help of a toy problem that allows for the direct computation of optimal controls on the particle level. Further, the algorithm was tested in scenarios where the direct computation of microscopic gradients is infeasible due to boundary conditions that do not naturally appear in the particle system formulation. Numerical studies underline the feasibility of the approach and motivate to use it in further applications and more intricate domains.

Appendix A

The aggressive space mapping algorithm used to obtain the numerical results is given by

---

Algorithm 1 Aggressive space mapping

Require: $u_0, tolerance$
1: Compute u_*^c iteratively with adjoint calculus and starting value u_0
2: $k = 1$
3: $u_1 = u_*^c$
4: Compute $\mathcal{T}(u_1)$ with adjoint calculus and starting value u_{start}
5: **while** $\|\mathcal{T}(u_k) - u_*^c\| > tolerance$ **do**
6:     $d_k = -(\mathcal{T}(u_k) - u_*^c)$
7:     Choose step length σ_k (such that $\|\mathcal{T}(u_k + \sigma_k d_k) - u_*^c\| \le \|\mathcal{T}(u_k) - u_*^c\|$)
8:     $u_{k+1} = u_k + \sigma_k d_k$
9:     Compute update $\mathcal{T}(u_{k+1})$ with adjoint calculus and starting value u_{start}
10:    $k = k + 1$
11: **end while**
12: $u_*^{ASM} = u_k$

Appendix B

We provide more details on the derivatives in the macroscopic Lagrangian (13).

$$\partial_{\rho_{ij}^s} \mathcal{F}_{ij}^{(1),s,-} = \begin{cases} \overline{v}_{i-1j}^{(1)} & \text{if } \overline{v}_{i-1j}^{(1)} < 0, (i-1,j) \in \mathcal{I}_\Omega \setminus \mathcal{I}_{\partial\Omega}, \\ 0 & \text{otherwise,} \end{cases}$$
$$= \partial_{\rho_{ij}^s} \mathcal{F}_{i-1j}^{(1),s,+}, \tag{17}$$

$$\partial_{\rho_{ij}^s} \mathcal{F}_{ij}^{(1),s,+} = \begin{cases} \overline{v}_{ij}^{(1)} & \text{if } \overline{v}_{ij}^{(1)} \ge 0, (i+1,j) \in \mathcal{I}_\Omega \setminus \mathcal{I}_{\partial\Omega}, \\ 0 & \text{otherwise,} \end{cases}$$
$$= \partial_{\rho_{ij}^s} \mathcal{F}_{i+1j}^{(1),s,-}, \tag{18}$$

$$\partial_{\tilde{\rho}^s_{ij}} \mathcal{F}^{(2),s,-}_{ij} = \begin{cases} \overline{v}^{(2)}_{ij-1} & \text{if } \overline{v}^{(2)}_{ij-1} < 0, (i,j-1) \in \mathcal{I}_\Omega \setminus \mathcal{I}_{\partial\Omega}, \\ 0 & \text{otherwise}, \end{cases}$$

$$= \partial_{\tilde{\rho}^s_{ij}} \mathcal{F}^{(2),s,+}_{ij-1}, \tag{19}$$

$$\partial_{\tilde{\rho}^s_{ij}} \mathcal{F}^{(2),s,+}_{ij} = \begin{cases} \overline{v}^{(2)}_{ij} & \text{if } \overline{v}^{(2)}_{ij} \geq 0, (i,j+1) \in \mathcal{I}_\Omega \setminus \mathcal{I}_{\partial\Omega}, \\ 0 & \text{otherwise}. \end{cases}$$

$$= \partial_{\tilde{\rho}^s_{ij}} \mathcal{F}^{(2),s,-}_{ij+1}. \tag{20}$$

Appendix C

We provide more details on the derivatives of the microscopic Lagrangian (14). The derivatives of the terms G, F for $k, l \in \{1, 2\}$ are defined in the following. The derivatives of the velocity selection mechanism with respect to the state variables are

$$\partial_{x^{(l),s}_i} G^{(k)}_i = \frac{\partial_{x^{(l),s}_i} \overline{v}^{(k)}(x^s_i)}{\tau}, \tag{21}$$

$$\partial_{v^{(l),s}_i} G^{(k)}_i = \begin{cases} -\frac{1}{\tau} & \text{if } l = k, \\ 0 & \text{otherwise}. \end{cases} \tag{22}$$

The derivatives of the interaction force F are

$$\partial_{x^{(l),s}_i} F^{(k)}_{ij} = \begin{cases} \frac{b_F \left(\|x^s_i - x^s_j\|_2 - 2R \right)^2}{\|x^s_i - x^s_j\|_2} + (x^{(l),s}_i - x^{(l),s}_j) \partial_{x^{(l),s}_i} \frac{b_F \left(\|x^s_i - x^s_j\|_2 - 2R \right)^2}{\|x^s_i - x^s_j\|_2} & \text{if } \|x^s_i - x^s_j\| < 2R, l = k, \\ \partial_{x^{(l),s}_i} \frac{b_F \left(\|x^s_i - x^s_j\|_2 - 2R \right)^2}{\|x^s_i - x^s_j\|_2} \left(x^{(k),s}_i - x^{(k),s}_j \right) & \text{if } \|x^s_i - x^s_j\| < 2R, l \neq k, \\ 0 & \text{otherwise}, \end{cases}$$

and more specifically

$$\partial_{x^{(l),s}_i} \frac{b_F \left(\|x^s_i - x^s_j\|_2 - 2R \right)^2}{\|x^s_i - x^s_j\|_2}$$

$$= \left(\frac{2b_F \left(\|x^s_i - x^s_j\|_2 - 2R \right)}{\|x^s_i - x^s_j\|_2^2} - \frac{b_F \left(\|x^s_i - x^s_j\|_2 - 2R \right)^2}{\|x^s_i - x^s_j\|_2^3} \right) (x^{(l),s}_i - x^{(l),s}_j).$$

Now, we differentiate the Lagrangian with respect to the state variables. First, we differentiate with respect to $x^{(1),s}_i$ to obtain

$$\mu_i^{s-1} = -\Delta t^f \partial_{x_i^{(1),s}} J^f(u, \boldsymbol{x}) + \mu_i^s + \Delta t^f \left(\bar{\mu}_i^{s-1} \partial_{x_i^{(1),s}} G_i^{(1),s} + A \sum_{j \neq i} \partial_{x_i^{(1),s}} F_{ij}^{(1),s} \left(\bar{\mu}_i^{s-1} - \bar{\mu}_j^{s-1} \right) \right.$$
$$\left. + \hat{\mu}_i^{s-1} \partial_{x_i^{(1),s}} G_i^{(2),s} + A \sum_{j \neq i} \partial_{x_i^{(1),s}} F_{ij}^{(2),s} \left(\hat{\mu}_i^{s-1} - \hat{\mu}_j^{s-1} \right) \right).$$

Second, we differentiate with respect to $x_i^{(2),s}$ to obtain

$$\tilde{\mu}_i^{s-1} = -\Delta t^f \partial_{x_i^{(2),s}} J^f(u, \boldsymbol{x}) + \tilde{\mu}_i^s + \Delta t^f \left(\bar{\mu}_i^{s-1} \partial_{x_i^{(2),s}} G_i^{(1),s} + A \sum_{j \neq i} \partial_{x_i^{(2),s}} F_{ij}^{(1),s} \left(\bar{\mu}_i^{s-1} - \bar{\mu}_j^s \right) \right.$$
$$\left. + \hat{\mu}_i^{s-1} \partial_{x_i^{(2),s}} G_i^{(2),s} + A \sum_{j \neq i} \partial_{x_i^{(2),s}} F_{ij}^{(2),s} \left(\hat{\mu}_i^{s-1} - \hat{\mu}_j^{s-1} \right) \right).$$

Third, we differentiate with respect to $v_i^{(1),s}$ and obtain

$$\bar{\mu}_i^{s-1} = \Delta t^f \mu_i^s + \bar{\mu}_i^s + \Delta t^f \left(\partial_{v_i^{(1),s}} G_i^{(1),s} \bar{\mu}_i^{s-1} + \partial_{v_i^{(1),s}} G_i^{(2),s} \hat{\mu}_i^{s-1} \right).$$

Lastly, we differentiate with respect to $v_i^{(2),s}$ and obtain

$$\hat{\mu}_i^{s-1} = \Delta t^f \tilde{\mu}_i^s + \hat{\mu}_i^s + \Delta t^f \left(\partial_{v_i^{(2),s}} G_i^{(1),s} \bar{\mu}_i^{s-1} + \partial_{v_i^{(2),s}} G_i^{(2),s} \hat{\mu}_i^{s-1} \right).$$

Note that due to (22), the terms with $\hat{\mu}_i^{s-1}$ (resp. $\bar{\mu}_i^{s-1}$) vanish in the last two equations and we can directly compute $\bar{\mu}_i^{s-1}$ and $\hat{\mu}_i^{s-1}$.

References

Albi G, Pareschi L (2013) Modeling self-organized systems interacting with few individuals: from microscopic to macroscopic dynamics. Appl Math Lett 26:397–401

Armbruster D, Martin S, Thatcher A (2017) Elastic and inelastic collisions of swarms. Phys D: Nonlinear Phenom 344:45–57

Bakr M, Bandler J, Madsen K, Søndergaard J (2001) An introduction to the space mapping technique. Optim Eng 2:369–384

Bakr MH, Bandler JW, Madsen K, Søndergaard J (2000) Review of the space mapping approach to engineering optimization and modeling. Optim Eng 1:241–276

Banda MK, Herty M (2011) Towards a space mapping approach to dynamic compressor optimization of gas networks. Optim Control Appl Methods 32:253–269

Bandler JW, Biernacki RM, Chen SH, Grobelny PA, Hemmers RH (1994) Space mapping technique for electromagnetic optimization. IEEE Trans Microw Theor Tech 42:2536–2544

Bandler JW, Biernacki RM, Chen SH, Hemmers RH, Madsen K (1995) Electromagnetic optimization exploiting aggressive space mapping. IEEE Trans Microw Theor Tech 43:2874–2882

Bandler JW, Cheng QS, Dakroury SA, Mohamed AS, Bakr MH, Madsen K, Sondergaard J (2004) Space mapping: the state of the art. IEEE Trans Microw Theor Tech 52:337–361

Boi S, Capasso V, Morale D (2000) Modeling the aggregative behavior of ants of the species Polyergus rufescens. Nonlinear Anal: Real World Appl 1:163–176

Burger M, Pinnau R, Totzeck C, Tse O (2020). *Mean-field optimal control and optimality conditions in the space of probability measures*, accepted for publication in SCICON

Burger M, Pinnau R, Totzeck C, Tse O, Roth A (2020) Instantaneous control of interacting particle systems in the mean-field limit. J Comput Phys 405:109181

Carrillo J. A, Fornasier M, Toscani G, Vecil F (2010)*Particle, kinetic, and hydrodynamic models of swarming*, in Mathematical modeling of collective behavior in socio-economic and life sciences, Springer, pp. 297–336

Chuang Y-L, D'Orsogna MR, Marthaler D, Bertozzi AL, Chayes LS (2007) State transitions and the continuum limit for a 2D interacting, self-propelled particle system. Phys D: Nonlinear Phenom 232:33–47

Dai YH, Yuan Y (1999) A nonlinear conjugate gradient method with a strong global convergence property. SIAM J Optim 10:177–182

D'Orsogna MR, Chuang YL, Bertozzi AL, Chayes LS (2006) Self-propelled particles with soft-core interactions: patterns, stability, and collapse. Phys Rev Lett 96:104302

D'Orsogna MR, Chuang Y-L, Bertozzi AL, Chayes LS (2006) Self-propelled particles with soft-core interactions: patterns, stability, and collapse. Phys Rev Lett 96:104302

Echeverría D, Hemker PW (2005) Space mapping and defect correction. Comput Methods Appl Math 5:107–136

Elliott D, Smith D (1993) Football stadia disasters in the united kingdom: learning from tragedy? Indus Environ Crisis Q 7:205–229

Erbrich M, Göttlich S, Pfirsching M (2018) Optimal packing of material flow on conveyor belts. Optim Eng 19:71–96

Fan S, Seibold B (2013) Data-fitted first-order traffic models and their second-order generalizations: comparison by trajectory and sensor data. Transp Res Record 2391:32–43

Festa A, Göttlich S, Pfirsching M (2019) A model for a network of conveyor belts with discontinuous speed and capacity. Networks & Heterogeneous Media 14

Fletcher R, Reeves CM (1964) Function minimization by conjugate gradients. Computer J 7:149–154

Golse F (2003) The mean-field limit for the dynamics of large particle systems. Journées équations aux dérivées partielles 1–47

Göttlich S, Hoher S, Schindler P, Schleper V, Verl A (2014) Modeling, simulation and validation of material flow on conveyor belts. Appl Math Modell 38:3295–3313

Göttlich S, Klar A, Tiwari S (2015) Complex material flow problems: a multi-scale model hierarchy and particle methods. J Eng Math 92:15–29

Göttlich S, Knapp S, Schillen P (2018) A pedestrian flow model with stochastic velocities: microscopic and macroscopic approaches. Kinetic Related Models 11:1333–1358

Göttlich S, Pfirsching M (2018) A micro-macro hybrid model with application for material and pedestrian flow. Cogent Math Stat 5:1–20

Göttlich S, Teuber C (2018) Space mapping techniques for the optimal inflow control of transmission lines. Optim Methods Softw 33:120–139

Haase K, Kasper M, Koch M, Müller S (2019) A pilgrim scheduling approach to increase safety during the hajj. Op Res 67:376–406

Helbing D, Farkas I, Vicsek T (2000) Simulating dynamical features of escape panic. Nature 407:487–490

Helbing D, Farkas I. J, Molnar P, Vicsek T (2002)*Simulation of pedestrian crowds in normal and evacuation situations*, in Pedestrian and Evacuation Dynamics, Springer, Berlin, 1 ed., pp. 21–58

Helbing D, Molnár P (1995) Social force model for pedestrian dynamics. Phys Rev E 51:4282–4286

Hinze M, Pinnau R, Ulbrich M, Ulbrich S (2009) Optimization with PDE constraints. Springer, New York

Holden H, Karlsen KH, Lie KA (2000) Operator splitting methods for degenerate convection-diffusion equations II: numerical examples with emphasis on reservoir simulation and sedimentation. Comput Geosci 4:287–322

Johnson NR (1987) Panic at the who concert stampede: an empirical assessment. Social Probl 34:362–373

Klüpfel H, Meyer-König T, Wahle J, Schreckenberg M (2001) Microscopic simulation of evacuation processes on passenger ships, in theory and practical issues on cellular automata. Springer, London, pp 63–71

Koch AL, White D (1998) The social lifestyle of myxobacteria. BioEssays 20:1030–1038

Koziel S, Ciaurri DE, Leifsson L (2011) Surrogate-based methods, in computational optimization, methods and algorithms. Springer, Berlin, pp 33–59

Mahato NK, Klar A, Tiwari S (2018) Particle methods for multi-group pedestrian flow. Appl Math Modell 53:447–461

Miyoshi T, Nakayasu H, Ueno Y, Patterson P (2012) An emergency aircraft evacuation simulation considering passenger emotions. Computers Indus Eng 62:746–754

Nocedal J, Wright SJ (2006) Numerical Optimization, 2nd edn. Springer Series in Operations Research and Financial Engineering, Springer-Verlag, New York, New York, NY

Okazaki S, Matsushita S (1993) *A study of simulation model for pedestrian movement with evacuation and queuing*, in International Conference on Engineering for Crowd Safety, pp. 271–280

Parzen E (1962) On estimation of a probability density function and mode. Annal Math Stat 33:1065–1076

Prims D, Kötz J, Göttlich S, Katterfeld A(2019) *Validation of flow models as new simulation approach for parcel handling in bulk mode*, Logistics Journal : referierte Veröffentlichungen

Rosenblatt M (1956) Remarks on some nonparametric estimates of a density function. Annal Math Stat 27:832–837

Seyfried A, Boltes M *Pedestrian Dynamics Data Archive*. https://ped.fz-juelich.de/da/doku.php. Accessed: 2021-06-07

Toscani G (2006) Kinetic models of opinion formation. Commun Math Sci 4:481–496

Totzeck C (2020) An anisotropic interaction model with collision avoidance. Kinetic Related Models 13:1219–1242

Totzeck C, Pinnau R (2020) Space mapping-based receding horizon control for stochastic interacting particle systems: dogs herding sheep. J Math Indus 10:1–19

Tröltzsch F (2010) Optimal Control of Partial Differential Equations: Theory, Methods and Applications, vol 112, 1st edn. Graduate Studies in Mathematics. American Mathematical Society, Providence, Rhode Island

Weissen J, Göttlich S, Armbruster D (2021)*Density dependent diffusion models for the interaction of particle ensembles with boundaries*, http://arxiv.org/abs/2101.03950, pp. 1–25