Optimal Control of Hughes' Model for Pedestrian Flow via Local Attraction

Roland Herzog[1] · Jan-Frederik Pietschmann[2] · Max Winkler[3]

Abstract

We discuss the control of a human crowd whose dynamics is governed by a regularized version of Hughes' model, cf. Hughes (Transp Res Part B: Methodol 36(6):507–535, 2002. https://doi.org/10.1016/s0191-2615(01)00015-7). We assume that a finite number of agents act on the crowd and try to optimize their paths in a given time interval. The objective functional can be general and it can correspond, for instance, to the desire for fast evacuation or to maintain a single group of individuals. We provide an existence and regularity result for the coupled PDE-ODE forward model via an approximation argument, study differentiability properties of the control-to-state map, establish the existence of a globally optimal control and formulate optimality conditions.

Keywords Nonlinear PDEs · Eikonal equation · ODE-PDE coupling · Optimal control · Pedestrian motion

Mathematics Subject Classification 35Q93 · 49J20 · 35K20 · 92D50

✉ Max Winkler
max.winkler@mathematik.tu-chemnitz.de

Roland Herzog
roland.herzog@iwr.uni-heidelberg.de

Jan-Frederik Pietschmann
jan-f.pietschmann@uni-a.de

[1] Interdisciplinary Center for Scientific Computing, Universität Heidelberg, 69120 Heidelberg, Germany

[2] Chair for Inverse Problems, Universität Augsburg, 86159 Augsburg, Germany

[3] Faculty of Mathematics, Technische Universität Chemnitz, 09107 Chemnitz, Germany

1 Introduction

The starting point of this work is Hughes' model for the movement of a (large) crowd of pedestrians introduced in [37]. Its unknowns are the density $\rho = \rho(t, x)$ and the potential $\phi = \phi(t, x)$ functions for $x \in \Omega \subset \mathbb{R}^2$ and $t \in (0, T)$. With the space-time cylinder denoted by $Q_T := (0, T) \times \Omega$, the model reads

$$\partial_t \rho - \nabla \cdot (\rho \, f(\rho)^2 \, \nabla \phi) = 0 \quad \text{in } Q_T, \tag{1.1a}$$

$$|\nabla \phi| = \frac{1}{f(\rho)} \quad \text{in } Q_T. \tag{1.1b}$$

In the simplest case, the model is supplemented with homogeneous Dirichlet boundary conditions for ϕ and ρ. Due to the hyperbolic nature of the first equation, the boundary conditions for ρ have to be posed in a suitable sense (i. e., on a generalized inflow part involving the function f). While many models for pedestrian dynamics are microscopic in the sense that they provide constitutive laws for the motion of each pedestrian (e. g., systems of ODEs or cellular automata or the social force model, cf. [5, 14, 35]), Hughes' model starts from a macroscopic approach. It is based on the following three assumptions:

(i) The velocity v of the pedestrians is determined by the density ρ of the surrounding pedestrian flow and the behavioral characteristics of the pedestrians only. Denoting the movement direction by $u \in \mathbb{R}^2$ there holds

$$v = f(\rho) \, u, \quad |u| = 1,$$

where $f(\rho)$ is a non-increasing function that attains the value one at $\rho = 0$, is positive for $0 < \rho < 1$ and zero at $\rho = 1$. It models the effect that the larger the surrounding density, the slower the individual motion becomes.

(ii) Pedestrians have a common sense of the task (called potential ϕ), i. e., they aim to reach their common destination by

$$u = -\frac{\nabla \phi}{|\nabla \phi|}.$$

(iii) Pedestrians seek to minimize their (accurately) estimated travel time, but modify their velocity to avoid high densities. The potential is thus a solution of the Eikonal equation
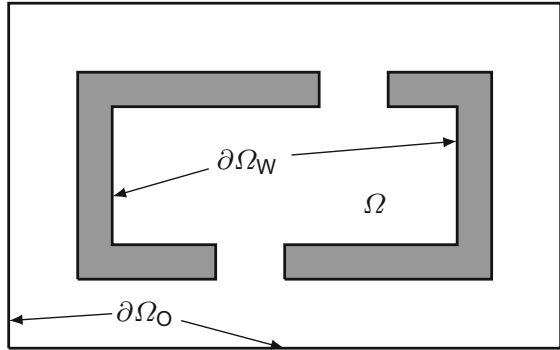
$$|\nabla \phi| = \frac{1}{f(\rho)}.$$

Combining these rules yields the model (1.1). For brevity we write

$$\widetilde{\beta}(\rho, \phi) := f(\rho)^2 \, \nabla \phi. \tag{1.2}$$

Due to the previous explanations it is clear that $-\widetilde{\beta}(\rho, \phi)|_{(t,x)}$ is the direction of the individuals in the point $(t, x) \in Q_T$.

Fig. 1 Illustration of a possible domain Ω and the boundary parts $\partial\Omega_O$ and $\partial\Omega_W$. The gray region is excluded from Ω

The analysis of this system is quite involved since the derivative of ϕ, being the viscosity solution to (1.1b), has jump discontinuities on a set that depends on ρ and is not known a priori. Since we are going to consider optimal control problems, we shall focus on a regularized version of the forward model (1.1), given by

$$\partial_t \rho - \nabla \cdot (\rho\,\widetilde{\beta}(\rho,\phi)) = \varepsilon\,\Delta\rho \qquad\qquad \text{in } Q_T, \qquad (1.3a)$$

$$-\delta_1\,\Delta\phi + |\nabla\phi|^2 = \frac{1}{f(\rho)^2 + \delta_2} \qquad\qquad \text{in } Q_T. \qquad (1.3b)$$

Here ε, δ_1 and δ_2 are positive and fixed regularization parameters. The model is supplemented with initial conditions

$$\rho(x,0) = \rho_0(x) \quad \text{in } \Omega \qquad\qquad (1.4)$$

and mixed boundary conditions

$$\begin{aligned} -(\varepsilon\,\nabla\rho + \rho\,\widetilde{\beta}(\rho,\phi)) \cdot n = \eta\,\rho, &\qquad \phi = 0 \quad \text{on } \Sigma_O, \\ (\varepsilon\,\nabla\rho + \rho\,\widetilde{\beta}(\rho,\phi)) \cdot n = 0, &\qquad \nabla\phi \cdot n = 0 \quad \text{on } \Sigma_W, \end{aligned} \qquad (1.5)$$

where we assume that the boundary consists of parts which act as "walls" $\partial\Omega_W$ not allowing the pedetrians to traverse, and a "truncation" or "outer boundary" $\partial\Omega_O$, at which pedestrians can exit with a given outflow velocity $\eta > 0$. We assume $\partial\Omega_O \cup \partial\Omega_W = \partial\Omega$ and $\partial\Omega_O \cap \partial\Omega_W = \emptyset$ and set $\Sigma_W := (0,T) \times \partial\Omega_W$ and $\Sigma_O := (0,T) \times \partial\Omega_O$. An example domain is illustrated in Fig. 1.

Remark 1.1 (On the regularizations (I)) As multiple regularisations are carried out, let us briefly discuss their meaning in terms of modeling and mathematical necessity.

δ_1: The elliptic regularization in the Eikonal equation (1.3b) is such that in the limit $\delta_1 \to 0$, we would recover the unique viscosity solution. For $\delta_1 = 0$, $\nabla\phi$ becomes discontinuous on a set that depends on ρ. In this situation, existence for the original Hughes' model is only known in one spatial dimension which would be a serious limitation for our application.

δ_2: While δ_2 prevents blow-up of the right hand side in (1.3b), setting $\delta_1 = 0$, its effect on the modulus of the velocity $|\tilde{\beta}(\rho, \phi)|$ is as follows: For $\delta_2 = 0$, one obtains $|\tilde{\beta}(\rho, \phi)| = f(\rho)$ while otherwise one has

$$|\beta(\rho, \phi)| = \frac{f(\rho)^2}{\delta_2 + f(\rho)} = f(\rho)\frac{f(\rho)}{\delta_2 + f(\rho)}.$$

As δ_2, the behavior of $|\beta(\rho, \phi)|$ is essentially unchanged except for values of ρ that are close to one (i.e. when $f(\rho)$ becomes very small). On the other hand, due to the properties of f, at $\rho = 1$, $|\beta|$ vanishes and thus this modification is not effective in the dynamics.

ε: The additional diffusion added to the evolution equation of the density serves two purposes. First, as we will see below, it allows for a continuous solution. This is essential as we will couple the system to an additional ODE later on that requires point evaluation of ρ. From the modeling point of view, the term introduces additional randomness in the motion of the pedestrians that seems a very reasonable addition to the directed motion of the convection term.

1.1 Optimal Control Problem

Our optimal control problem is based on the following scenario. We assume that there is a small, given number $M > 0$ of agents (guides), who are able to locally influence the motion of pedestrians in their vicinity. Think, for instance, of tourist guides or marked security personnel at large sports events. The i-th agent's position is described by a function $x_i(t) \in \mathbb{R}^2$, $i = 1, \ldots, M$, and the interaction is modeled by a radially symmetric and decreasing kernel $K(x) = k(|x|)$, which enters (1.3a) as an additional potential given by

$$\phi_K(\boldsymbol{x}; x) = \sum_{i=1}^{M} K(x - x_i(t)) = \sum_{i=1}^{M} k(|x - x_i(t)|). \tag{1.6}$$

Here and throughout, $\boldsymbol{x} = \boldsymbol{x}(t) = (x_1(t), \ldots, x_M(t))^{\mathrm{T}}$ is the function collecting all agent positions. Furthermore, we need to modify the model to insure that the maximal velocity of the crowd is still normalized to the velocity model $f(\rho)$ despite the agents' presence. Indeed, for the unregularized model,

$$f(\rho)^2 |\nabla \phi| = f(\rho)$$

holds. When ϕ is replaced by $\phi + \phi_K$, this is no longer true. Thus we explicitly normalize the transport direction, i. e., instead of (1.2), we define the transport velocity by

$$\beta(\rho, \phi, \boldsymbol{x}) := f(\rho) h(\nabla(\phi + \phi_K(\boldsymbol{x}; \cdot))). \tag{1.7}$$

Here, h is a smoothed projection onto the unit ball. Throughout this article we will use

$$h(y) = \min_\varepsilon\{1, |y|\}\frac{y}{|y|}$$

with $\min_\varepsilon$ a fixed smooth approximation of the minimum function.

Remark 1.2 (On the regularizations (II)) We emphasise that the regularisation of the function h serves no mathematical purpose (in fact, all of our analysis works without it) but is rather needed to preserve the crucial property of our model that the direction of motion is determined by a vector with unit length while the speed of motion is solely determined by the surrounding density via the function $f(\rho)$.

We assume that the agents move with maximal possible velocity towards a pre-scribed directions $u_i(t) \in \mathbb{R}^2$, $i = 1, \ldots, M$, which act as control in the system. Since the agents are also part of the crowd, their effective velocity will depend on the surrounding density in the same way as it does for all other individuals in the crowd. We therefore assume $|u_i(t)| \leq 1$ and the law of motion for the agents will be

$$\dot{x}_i(t) = f(\overline{\rho}(t, x_i(t)))\, u_i(t), \quad i = 1, \ldots, M, \tag{1.8}$$

where $\overline{\rho}$ is a suitable extension of ρ from Ω to $\mathbb{R}^2$ which will be detailed later. While this extension is necessary to ensure the existence of solutions since the agents may leave the domain Ω on which ρ is defined, the precise choice of the extension is clearly irrelevant in terms of modeling. Notice also that the ODE (1.8) does not prevent agents from walking through walls or to attract people from behind a wall. The first issue can be avoided by imposing additional state constraints in an optimal control problem. The second issue can be avoided when assuming the wall thickness to be larger than the agents' attraction radius.

The complete forward system which we are going to consider finally reads

$$\partial_t \rho - \nabla \cdot (\rho\, \beta(\rho, \phi, x)) = \varepsilon\, \Delta\rho \quad \text{in } Q_T, \tag{1.9a}$$

$$-\delta_1\, \Delta\phi + |\nabla\phi|^2 = \frac{1}{f(\rho)^2 + \delta_2} \quad \text{in } Q_T, \tag{1.9b}$$

$$\dot{x}_i(t) = f(\overline{\rho}(t, x_i(t)))\, u_i(t) \quad \text{for } t \in (0, T), \quad i = 1, \ldots, M, \tag{1.9c}$$

together with initial condition (1.4) on ρ and

$$x_i(0) = x_{i,0} \quad \text{for } i = 1, \ldots, M, \tag{1.10}$$

as well as boundary conditions (1.5).

The aim of the present paper is to investigate several optimal control problems for this coupled system. We seek an optimal control function u such that the solution triple $y = (\rho, \phi, x)$ is optimal in a certain sense. Depending on the application in mind it remains to define a suitable objective functional. We particularize the following two examples:

– *Minimal evacuation time*: In this case one seeks to minimize the time required for the evacuation of a room. The exits where individuals can leave the domain are located at the truncation boundary $\partial\Omega_O$ which is sufficiently far away from the actual domain of interest. As time-optimal control problems with PDEs are rather challenging, see, e.g., [6, 40, 47, 52], we consider a simpler but closely related model. We fix a reasonably large final time $T > 0$ and minimize

$$J(\rho, \phi, \boldsymbol{x}; \boldsymbol{u}) := c_1 \int_{\Omega} \rho(T, x) \, \mathrm{d}x$$

$$+ c_2 \int_0^T \int_{\Omega} t \, \rho(t, x) \, \mathrm{d}x \, \mathrm{d}t + \frac{\alpha}{2T} \sum_{i=1}^M \|u_i\|_{H^1(0,T)}^2 \tag{1.11}$$

with weighting parameters $c_1, c_2 > 0$ and a regularization parameter $\alpha > 0$. The first term in J penalizes individuals remaining in the room at time T. The second term encourages individuals to leave the room as early as possible. The last terms provides the required regularity for the control variables, so that the forward system (1.9) is well-defined. From the modeling point of view, these terms also avoid unrealistic trajectories of the agents.

– *Optimal binding of a crowd*: In some applications it might be desired to keep the group of individuals together, i.e., trying to maintain a single group during an evacuation. This is also motivated by a similar approach which has been used to model the repulsive interaction of dogs in a flock of sheeps, see [13]. To this end, we define the center of mass and variance of ρ as

$$E_\rho(t) := \frac{1}{M(t)} \int_{\Omega} x \rho(t, x) \, \mathrm{d}x \quad \text{and} \quad V_\rho(t) := \frac{1}{M(t)} \int_{\Omega} \rho |x - E_\rho(t)|^2 \, \mathrm{d}x,$$

with total mass $M(t) = \int_{\Omega} \rho(t, x) \, \mathrm{d}x$. A crowd is optimally kept together when the functional

$$J(\rho, \phi, \boldsymbol{x}; \boldsymbol{u}) := \frac{1}{2T} \int_0^T V_\rho(t) \, \mathrm{d}t + \frac{\alpha}{2T} \sum_{i=1}^M \|u_i\|_{H^1(0,T)}^2 \tag{1.12}$$

is minimized.

Clearly, it is also possible to use a combination of the objective functionals (1.11) and (1.12).

1.2 Related Work

We briefly review the literature regarding the analysis of the original Hughes model and related optimal control problems.

A first contribution on existence of solutions for the Hughes model in the one-dimensional case is [26]. There it was shown, starting from the regularized version (1.3), that in the limit $\varepsilon \to 0$ a suitable unique entropy solution ρ exists. The proof

is based on a vanishing viscosity argument and Kruzkov's doubling of variables technique to show uniqueness. The results were complemented by more detailed studies on the unregularized problem, also numerically. For instance generalizations to higher spatial dimensions can be found in [18], even for a slightly more general class of models. Further articles examine Riemann-type solutions to the unregularized problem; see [3, 4, 28] in one spatial dimension. As far as modeling is concerned, slightly different models were derived in [12] based on a mean field games approach. In [16], a modified approach using multiple local potentials ϕ_i instead of one global potential ϕ is introduced, removing the possibly unrealistic assumption that every pedestrian has complete information of the entire crowd. Moreover, in [15], a discrete pedestrian model in a graph network is studied. Further complex scenarios based on a regularized version of Hughes' model are investigated in [19, 25].

From a broader perspective, the optimal control of (1.9) falls into the class of the optimal control of coupled ODE-PDE systems. Such problems, with models from a range of different applications have been analyzed, for instance, in [17, 36, 38, 39, 51]. We want to mention a few contributions dealing with optimal control of coupled PDE-ODE models in the context of pedestrian dynamics. In [30, 31] the feedback control in a multi-agent system is considered aiming at an optimization of the dynamics of a population. The individuals are either followers or leaders and the interaction is modeled by an ODE system. Moreover, the authors study also the limit case where the number of followers tends to infinity, which results in a couples system of ODEs and a PDE of Vlasov type. A similar approach is studied in [1]. Therein, the interaction between individuals is a short range retraction and a long range attraction. While leaders are not visible to the remaining crowd, they still influence it by taking part in these interactions. Then an optimal control problem arises as an external force acts on the leaders. The authors also consider, in the limit of many individuals (grazing interaction limit), macroscopic Boltzmann type equations for this interaction, while the number of leaders remains fixed and finite. Furthermore, in [13] external agents act as control. Again, they start from a microscopic ODE model and subsequently obtain a continuous model for the uncontrolled population by means of a mean field limit. For a general overview on interacting particle systems and control, we refer the reader to [46].

Closer to our approach is the work of [7]. There, a system of hyperbolic conservation laws for the density of different pedestrian groups is coupled to ODEs accounting for agents. Due to the low regularity of solutions to the hyperbolic equations, a regularization in the ODEs, similar to Lemma 3.6 in our case, is used. See also [8, 9] for a similar approach in different settings. Also related to our approach are the studies [20, 21] where evacuation is optimized by shape optimization of obstacles in a room. A more detailed survey on the control of several models for pedestrian dynamics can be found in [2]. Finally, we want to mention our proceeding article [45] where we develop a numerical solution approach for our model based on a structure preserving finite volume method and a projected gradient algorithm.

This paper is organized as follows. In Sect. 2 we collect the required notation, introduce some assumptions and state a precise existence and uniqueness result for the regularized system (1.9). The full forward system involving also the ordinary differential equation (1.8) is investigated in Sect. 3 and the linearized forward system

in Sect. 4. The latter is required to establish the differentiability of the control-to-state map, which in turn is the basis of optimality conditions. Based on this we discuss the optimal control problem in Sect. 5 and derive first-order necessary optimality conditions. The presentation of numerical results will be postponed to a forthcoming publication.

2 Mathematical Preliminaries

Let us first state the assumptions on the domain and data.

(A1) $\Omega \subset \mathbb{R}^2$ is an open, bounded domain with C^4-boundary $\partial\Omega$.
(A2) There exist two measurable sets $\partial\Omega_O, \partial\Omega_W \subset \partial\Omega$ s.t. $\partial\Omega_O \cup \partial\Omega_W = \partial\Omega$ and $\partial\Omega_O \cap \partial\Omega_W = \emptyset$. Moreover, $\partial\Omega_O$ has positive measure with respect to the Lebesgue measure on $\partial\Omega$.
(A3) The initial density satisfies $\rho_0 \in W^{3/2,4}(\Omega)$ and $0 \le \rho_0 \le 1$ a.e. in Ω.
(A4) There holds $f \in W^{3,\infty}(\mathbb{R}) \cap C_c(\mathbb{R})$ with $f(0) = 1$, $f(1) = 0$ and $f(\rho) > 0$ for all $\rho \in [0, 1)$.

Moreover, we require some assumptions on the potential functions (1.6) of the agents, which depend on the kernel K.

(K1) The kernel $K \colon \mathbb{R}^2 \to \mathbb{R}$ is radially symmetric, i.e., $K(x) = k(|x|)$, where $k \colon [0, \infty) \to \mathbb{R}$ is nonnegative and decreasing.
(K2) The kernel satisfies $K \in W^{3,\infty}(\mathbb{R}^2)$.

Finally, we consider an assumption on the velocity controls of the agents.

(C1) There holds $u_i \in L^\infty(0, T; \mathbb{R}^2)$ and $\|u_i\|_{L^\infty(0,T)} \le 1$ for $i = 1, \ldots, M$.

Throughout this article we frequently exploit the boundedness and Lipschitz continuity of the functions K, f, h and $g(x) := x f(x)$ and its derivatives. When doing so, we denote the bounds and Lipschitz constants by C_K, C_f, C_h, C_g and L_K, L_f, L_h, L_g, respectively.

Remark 2.1 (Assumptions)

(i) Our results extend to the case $d = 3$ upon adopting the Sobolev embeddings used in several places.
(ii) Assumption (A2) essentially means that the door and wall parts of the boundary "do not meet", i.e., we do not consider a truly mixed boundary value problem in order to avoid technical conditions ensuring sufficient regularity of solutions. (A2) can be replaced, e.g., in two spatial dimensions, by suitable angle conditions on the points where the two parts of the boundary meet. The interested reader is referred to results in [33].
(iii) The optimal regularity for the initial datum ρ_0 is the Besov space $B_{pp}^{2-2/p}(\Omega)$; see [22].
(iv) While in Assumption (A4) f is defined on all of $\mathbb{R}$, as far as the modeling is concerned, only $f|_{[0,1]}$ is relevant. Indeed, we will later see that the solution to (1.9a) satisfies $0 \le \rho \le 1$.

(v) A reasonable choice in (K1) and (K2) is the kernel function

$$K(x - x_i(t)) = \begin{cases} s \exp\left(-\frac{R^2}{R^2 - |x - x_i(t)|^2}\right), & \text{if } |x - x_i(t)| < R, \\ 0, & \text{otherwise}, \end{cases} \tag{2.1}$$

where $s > 0$ is an intensity factor and $R > 0$ is related to an attraction radius.

Notations. The space-time cylinder and its lateral surface are denoted by $Q_T := (0, T) \times \Omega$ and $\Sigma_T = (0, T) \times \partial\Omega$, respectively. The boundary surface can be divided into

$$\Sigma_{O,T} = (0, T) \times \partial\Omega_O \quad \text{and} \quad \Sigma_{W,T} = (0, T) \times \partial\Omega_W.$$

The (Frobenius) inner product of two matrices $A, B \in \mathbb{R}^{n \times n}$ is denoted by

$$A : B := \text{trace}(A^T B) = \sum_{i,j=1}^{n} a_{ij} b_{ij}.$$

The Jacobian of a function $h \colon \mathbb{R}^2 \to \mathbb{R}^2$ is denoted by Dh and the Hessian of a function $u \colon \mathbb{R}^2 \to \mathbb{R}$ is denoted by $\nabla^2 u$. Furthermore, $\varphi_\gamma \in C_c^\infty(\mathbb{R}^2)$ is a standard mollifier, see, e.g., Ch. 4.4 in [11], i.e., a function satisfying

$$\text{supp}\, \varphi_\gamma \subset \overline{B_\gamma(0)} \quad \text{and} \quad \int_{\mathbb{R}^2} \varphi_\gamma \, dx = 1. \tag{2.2}$$

Note that

$$\|f - \varphi_\gamma * f\|_{C(\overline{\Omega})} \to 0 \text{ as } \gamma \to 0 \tag{2.3}$$

for every continuous function $f \in C(\overline{\Omega})$, cf. Prop. 4.21 in [11].

Finally, we introduce the subspace of $H^1(\Omega)$ incorporating the Dirichlet boundary conditions as

$$H_D^1(\Omega) := \{v \in H^1(\Omega) | v = 0 \text{ a.e. on } \partial\Omega_O\}$$

and, for $p \in [1, \infty]$, we denote the subspace of $W^{2,p}(\Omega)$ fulfilling the boundary conditions (1.5) by

$$W_{ND}^{2,p}(\Omega) := \{v \in W^{2,p}(\Omega) | v = 0 \text{ a.e. on } \partial\Omega_O, \nabla v \cdot n = 0 \text{ a.e. on } \partial\Omega_W\}.$$

For time-dependent functions we introduce, for $p \in (1, \infty)$ and $r, s \in \mathbb{N}_0$, the spaces

$$W_p^{r,s}(Q_T) := L^p(0, T; W^{r,p}(\Omega)) \cap W^{s,p}(0, T; L^p(\Omega)),$$

equipped with the natural norm $\left(\|\cdot\|_{L^p(0,T;W^{r,p}(\Omega))}^p + \|\cdot\|_{W^{s,p}(0,T;L^p(\Omega))}^p\right)^{1/p}$. Spaces with non-integral r and s are defined, as usual, as (real) interpolation spaces. Of particular interest in our application is the space $W_p^{2,1}(Q_T)$ with $p > d = 2$, which fulfills the embedding

$$W_p^{2,1}(Q_T) \hookrightarrow C([0, T]; W^{1,p}(\Omega)) \hookrightarrow C(\overline{Q_T}). \tag{2.4}$$

This is needed in order to allow point evaluations of the density ρ, required in the ordinary differential equation (1.9c). Finally, for functions from the Hölder space $C^{1,\alpha}(\Omega)$ we introduce the norm

$$\|u\|_{C^{1,\alpha}(\Omega)} := \|u\|_{C^1(\Omega)} + \max_{|\beta|=1} |D^\beta u|_{C^{0,\alpha}(\Omega)},$$

$$\text{where } |u|_{C^{0,\alpha}(\Omega)} = \sup_{x \neq y \in \Omega} \frac{|u(x) - u(y)|}{|x - y|^\alpha}.$$

In the following we collect some important properties of the function spaces used in this article.

Lemma 2.2 *For each $\theta \in [0, 1]$, $p \in (1, \infty)$ and $0 \leq s < r$, the continuous embedding*

$$W^{1,p}(0, T; W^{s,p}(\Omega)) \cap L^p(0, T; W^{r,p}(\Omega)) \hookrightarrow W^{\theta,p}(0, T; W^{\theta s + (1-\theta)r, p}(\Omega)).$$

holds.

Proof See Lemma 4.3 in [24]. ∎

Similarly as above we define the Sobolev spaces

$$W_p^{r,s}(\Sigma_T) := L^p(0, T; W^{r,p}(\partial\Omega)) \cap W^{s,p}(0, T; L^p(\partial\Omega))$$

on the lateral boundary $\Sigma_T = (0, T) \times \partial\Omega$ of the space-time cylinder Q_T.

The following trace theorem is proved in [22, Lem. 3.5]; see also Sect. 2 in [23]:

Lemma 2.3 *For $p > 1$, the trace operators*

$$\gamma_0 \colon W_p^{2,1}(Q_T) \to W_p^{2-1/p, 1-1/(2p)}(\Sigma_T),$$

$$\gamma_1 \colon W_p^{2,1}(Q_T) \to W_p^{1-1/p, 1/2-1/(2p)}(\Sigma_T)$$

defined by $\gamma_0\rho = \rho|_{\Sigma_T}$ and $\gamma_1\rho = \nabla\rho \cdot n_{\Sigma_T}$ are bounded and have a continuous right inverse.

Recall the differential equation (1.8), where an extension to $\mathbb{R}^2$ of the density function ρ is used. For theoretical purposes we will use an extension operator fulfilling the following result from Lemma 6.37 in [32]:

Lemma 2.4 *Let $\alpha \in (0, 1)$ be a fixed number. There exists a linear, continuous extension operator*

$$E \colon C^{1,\alpha}(\overline{\Omega}) \to C^{1,\alpha}(\mathbb{R}^2)$$

such that

$$\|Ef\|_{C^{1,\alpha}(\mathbb{R}^2)} \leq C_E \|f\|_{C^{1,\alpha}(\Omega)} \quad \text{and} \quad \|Ef\|_{L^\infty(\mathbb{R}^2)} \leq C_{E,\infty} \|f\|_{L^\infty(\Omega)},$$

holds for all $f \in C^{1,\alpha}(\overline{\Omega})$. For brevity we will write $\overline{f} := Ef$.

3 Analysis of the Forward System

This section is devoted to showing the existence of strong solutions to the forward system (1.9) with boundary and initial conditions (1.4), (1.5), (1.10). We proceed in two steps. First we provide auxiliary results on equation (1.9b) as well as on linear parabolic equations. Then we prove existence of solutions to the complete forward system.

3.1 Preliminary Results

First, we study the regularized Eikonal equation (1.9b).

Lemma 3.1 *For given $\widetilde{\rho}\colon Q_T \to \mathbb{R}$ consider the equation*

$$-\delta_1 \, \Delta\phi + |\nabla\phi|^2 = \frac{1}{f(\widetilde{\rho})^2 + \delta_2} \quad \text{in } Q_T,$$

with boundary conditions (1.5). We have:

(i) *If $\widetilde{\rho} \in C([0, T]; L^2(\Omega)) \cap H^1(0, T; H^1(\Omega)^*)$, then there exists a unique strong solution which satisfies $\phi \in L^\infty(0, T; W_{ND}^{2,p}(\Omega)) \cap H^1(0, T; H^1(\Omega))$ for all $2 \le p < \infty$. Moreover, the a priori estimates*

$$\|\phi\|_{L^\infty(0,T;W^{2,p}(\Omega))} \le \widetilde{C}_\phi \quad \text{and} \quad \|\phi\|_{H^1(0,T;H^1_D(\Omega))} \le \widetilde{C}_\phi \, \|\widetilde{\rho}\|_{H^1(0,T;H^1(\Omega)^*)} \tag{3.1}$$

hold with a positive constant $\widetilde{C}_\phi$ depending on p, δ_1, δ_2, Ω and T only.

(ii) *If $\widetilde{\rho} \in W_p^{2,1}(Q_T)$ holds, the strong solution ϕ additionally belongs to $W_p^{4,1}(Q_T)$ and satisfies the a priori estimate*

$$\|\phi\|_{W^{1,p}(0,T;W^{2,p}(\Omega))} \le \overline{C}_\phi \, \|\widetilde{\rho}\|_{W_p^{2,1}(Q_T)}, \tag{3.2}$$

with a constant $\overline{C}_\phi$ depending on p, δ_1, δ_2, Ω and T only.

(iii) *For any $\widetilde{\rho}_1, \widetilde{\rho}_2 \in C([0, T]; L^2(\Omega))$, the corresponding solutions ϕ_1 and ϕ_2 satisfy the Lipschitz estimate*

$$\|\phi_1(\cdot, t) - \phi_2(\cdot, t)\|_{W^{2,2}(\Omega)} \le \widehat{C}_\phi \, \|\widetilde{\rho}_1(\cdot, t) - \widetilde{\rho}_2(\cdot, t)\|_{L^2(\Omega)} \tag{3.3}$$

for all $t \in [0, T]$, with $\widehat{C}_\phi$ depending on p, δ_1, δ_2, Ω and T only.

From now on we shall use the definition $C_\phi := \max\{\widetilde{C}_\phi, \overline{C}_\phi, \widehat{C}_\phi\}$.

Proof We first show assertion (i). Note that due to the continuity of $\widetilde{\rho}$ in time it makes sense to define, for fixed $t \in [0, T]$, the function

$$q_t(x) := \frac{1}{\delta_1^2} \frac{1}{f(\widetilde{\rho}(t, x))^2 + \delta_2}. \tag{3.4}$$

By assumption (A4) on f we then have $q_t \in L^\infty(\Omega)$.

Step 1: Existence. First note that an application of the transformation

$$\psi(t, \cdot) = e^{-\phi(t, \cdot)/\delta_1} - 1 \tag{3.5}$$

to (1.9b) yields for all $t \in [0, T]$ the linear problem

$$\begin{aligned}
-\Delta\psi(t, \cdot) + q_t\,\psi(t, \cdot) &= -q_t(\cdot) && \text{in } \Omega, \\
\psi(t, \cdot) &= 0 && \text{on } \partial\Omega_O, \\
\partial_n\psi(t, \cdot) &= 0 && \text{on } \partial\Omega_W.
\end{aligned} \tag{3.6}$$

As $q_t \in L^\infty(\Omega)$ and $q_t > 0$ a.e. in Ω, the Lax-Milgram lemma yields for every $t \in [0, T]$ the existence of a unique weak solution $\psi(t, \cdot) \in H_D^1(\Omega)$ and the estimate

$$\|\psi(t, \cdot)\|_{H^1(\Omega)} \le C\,\|q_t\|_{L^2(\Omega)}.$$

Elliptic regularity theory (see for instance Theorem 3.17 in [49]) then implies $\psi(t, \cdot) \in W^{2,q}(\Omega)$ for any $2 \le q < \infty$, and there holds

$$\|\psi(t, \cdot)\|_{W^{2,q}(\Omega)} \le C\,\|q_t\|_{L^q(\Omega)} =: C_\phi \tag{3.7}$$

for $t \in [0, T]$. For given $x \in \Omega$, the function $t \mapsto q_t(x)$ belongs to $L^\infty(0, T)$ again by (A4). Taking the supremum over $0 \le t \le T$ yields the regularity $\psi \in L^\infty(0, T; W^{2,q}(\Omega))$. Formally differentiating equation (3.6) with respect to time we see that $\partial_t\psi$ satisfies

$$-\Delta\partial_t\psi(t, \cdot) + q_t\partial_t\psi(t, \cdot) = -(1 + \psi(t, \cdot))(\partial_t q_t) \quad \text{in } \Omega.$$

Using its definition and the fact that $\widetilde{\rho} \in H^1(0, T; H^1(\Omega)^*)$, the derivative of $q_t(x)$ with respect to time is an element of $L^2(0, T; H^1(\Omega)^*)$. Therefore, Sect. 2.2.2, Corollary, p. 99 in [49] (after replacing $\partial_t q_t(x)$ by a continuous in time approximation and passing to the limit) yields the existence of $\partial_t\psi \in L^2(0, T; H^1(\Omega))$ with

$$\|\partial_t\psi\|_{L^2(0,T;H^1(\Omega))} \le C\,\|\partial_t\widetilde{\rho}\|_{L^2(0,T;H^1(\Omega)^*)}. \tag{3.8}$$

Step 2: Strict lower bound. In order to invert (3.5) and obtain a solution to (1.9b), we need to ensure that $\psi > -1$ holds. This follows from Theorem 4 in [43], provided that we can show $\psi \ge -1$, $\|\psi\|_{L^\infty(Q_T)} < \infty$ and Hölder continuity of $\psi(t, \cdot)$ for a.a. $t \in (0, T)$. The last two assertions are a direct consequence of the embeddings $W^{2,p}(\Omega) \hookrightarrow L^\infty(\Omega)$ and $W^{2,p}(\Omega) \hookrightarrow C^{0,1/2}(\Omega)$ as $p \ge 2$ and we are in dimension 2, combined with (3.7). To show the first part we choose $\phi = (\psi + 1)_-$, i.e., the negative part of $\psi + 1$, as test function in the weak formulation of (3.6) and obtain

$$\int_\Omega |\nabla(\psi(t, \cdot) + 1)_-|^2\,dx + \int_\Omega q_t(x)(\psi + 1)_-^2\,dx = 0.$$

Since q_t is strictly positive, this implies that $(\psi + 1)_- = 0$ holds a.e. in Ω. Thus there exists a positive constant C_ψ s.t.

$$\psi \geq C_\psi > -1 \text{ a.e. in } \Omega.$$

We can therefore invert the transformation (3.5) and conclude that the solution of (1.9b), (1.5) fulfills the desired regularity, as the regularity is unaffected by the transformation (using $\psi \in L^\infty(0, T; L^\infty(\Omega))$ and $\psi > -1$). This ends the proof of assertion (i). To conclude (ii) we only have to show the additional regularity $\partial_t \psi \in L^p(0, T; W^{2,p}(\Omega))$ which follows under the assumption $\widetilde{\rho} \in W_p^{2,1}(Q_T)$ and standard elliptic regularity theory, see [49, Theorem 3.17].

Similarly, in case $\widetilde{\rho} \in W_p^{2,1}(Q_T)$ we obtain $\partial_t \psi \in L^p(0, T; W^{2,p}(\Omega))$, which gives the desired estimate.

Step 3: Lipschitz estimate. To show (iii), denote by $q_{t,1}$, $q_{t,2}$ the respective coefficients for $\widetilde{\rho}_1$ and $\widetilde{\rho}_2$ as in (3.4) and, analogously, let ψ_1 and ψ_2 be the respective solutions to (3.6). Then $\overline{\psi} = \psi_1 - \psi_2$ satisfies

$$
\begin{aligned}
-\Delta\overline{\psi}(t, \cdot) + q_{t,1}(x)\,\overline{\psi}(t, \cdot) &= (1 - \psi_2(t, \cdot))\,\overline{q}_t(x) & &\text{in } \Omega, \\
\overline{\psi}(t, \cdot) &= 0 & &\text{on } \partial\Omega_O, \\
\partial_n\overline{\psi}(t, \cdot) &= 0 & &\text{on } \partial\Omega_W,
\end{aligned}
\tag{3.9}
$$

for $t \in [0, T]$ with $\overline{q}_t = q_{t,1} - q_{t,2}$. Noting that q_t is Lipschitz continuous as a function of $\widetilde{\rho}$ (due to (A4)) and applying the a priori estimate (3.7) to (3.9) yields

$$\|\overline{\psi}(t, \cdot)\|_{W^{2,2}(\Omega)} \leq C\,\|\overline{q}_t\|_{L^2(\Omega)} \leq C\,\|\widetilde{\rho}_1(t, \cdot) - \widetilde{\rho}_2(t, \cdot)\|_{L^2(\Omega)},$$

where we used the boundedness of $(1 - \psi_2(t, \cdot))$ in $L^\infty(\Omega)$ and, again, $W^{2,p}(\Omega) \hookrightarrow L^\infty(\Omega)$ and (3.7). This implies (iii) and completes the proof. $\qquad\square$

To obtain the desired $W_p^{2,1}(Q_T)$-regularity of the density function ρ we will need the following lemma taken from [22] but adopted to our notation.

Lemma 3.2 *Let assumptions (A1)–(A4) hold, $p \in [2, \infty)$ and $\varepsilon > 0$. Suppose that $c \in L^p(Q_T)$, $b \in L^p(0, T; L^\infty(\Omega))$, $r \in W_p^{1-1/p, 1/2-1/(2p)}(\Sigma_T)$ and $\rho_0 \in W^{2-2/p,p}(\Omega)$ are given. Then the problem*

$$
\begin{aligned}
\partial_t \rho - \varepsilon\Delta\rho + b \cdot \nabla\rho &= c & &\text{in } Q_T, \\
\varepsilon\nabla\rho \cdot n &= r & &\text{on } \Sigma_{W,T}, \\
\rho &= 0 & &\text{on } \Sigma_{O,T}, \\
\rho(0) &= \rho_0 & &\text{in } \Omega
\end{aligned}
$$

admits a unique strong solution $\rho \in W_p^{2,1}(Q_T)$ depending continuously on the input data c, r and ρ_0.

__Proof__ See Theorem 2.1 in [22]. $\qquad\square$

Finally, we need the following regularity result for $p = 2$ with flux boundary conditions.

Lemma 3.3 *Given $h \in H^1(0, T; H^1(\Omega)) \cap L^\infty(Q_T)$, $g \in L^\infty(\mathbb{R})$ and $\rho_0 \in H^1(\Omega)$, the variational problem*

$$\int_{Q_T} \partial_t \rho \, \xi \, \mathrm{d}x \, \mathrm{d}t + \varepsilon \int_{Q_T} \nabla \rho \cdot \nabla \xi \, \mathrm{d}x \, \mathrm{d}t - \int_{Q_T} g(\rho) \, h \cdot \nabla \xi \, \mathrm{d}x \, \mathrm{d}t$$

$$= -\eta \int_{\Sigma_T} \chi_{\partial \Omega_O} \, \rho \, \xi \, \mathrm{d}s_x \, \mathrm{d}t \quad \text{for all } \xi \in L^2(0, T; H^1(\Omega)) \tag{3.10}$$

has a unique solution $\rho \in L^\infty(0, T; H^1(\Omega)) \cap H^1(0, T; L^2(\Omega))$ with $\rho(0) = \rho_0$.

The proof mainly uses standard methods but since, to the best of the authors' knowledge, a proof matching our boundary conditions is not available in the literature, we included it into Sect. A.

Next we define our notion of solution for the ODE (1.9c).

Definition 3.4 Fix $2 < p < \infty$ and $\rho \in W_p^{2,1}(Q_T)$. Then for given f and u satisfying assumptions (A4) and (C1) and an initial value $x_0 \in \mathbb{R}^2$ we say that $x \colon [0, T] \to \mathbb{R}^2$ is a solution to

$$\dot{x}(t) = f(\overline{\rho}(x(t), t)) \, u(t), \tag{3.11}$$

if it is absolutely continuous, satisfies (3.11) for a.a. $t \in [0, T]$ and $x(0) = x_0$.

We have the following result about the existence of a solution of (3.11).

Lemma 3.5 *For given $2 < p < 4$, $u \in L^\infty(0, T; \mathbb{R}^2)$ satisfying assumption (C1) and $\rho \in W_p^{2,1}(Q_T)$, there exists a unique, absolutely continuous solution $x \colon [0, T] \to \mathbb{R}^2$ to (3.11) satisfying $x(0) = x_0$. Furthermore, $x \in W^{1,\infty}(0, T)$ holds. For $\rho_1, \rho_2 \in W_p^{2,1}(Q_T)$, the corresponding solutions x_1 and x_2 satisfy*

$$\|x_1 - x_2\|_{L^\infty(0,T;\mathbb{R}^2)} \leq C_s \, \|\rho_1 - \rho_2\|_{L^\infty(Q_T)}, \tag{3.12}$$

where the constant C_s depends on u, T, C_E, $C_{E,\infty}$ and the Lipschitz constants of f and ρ.

Proof First note that $\rho \in L^p(0, T; W^{2,p}(\Omega)) \hookrightarrow L^p(0, T; C^{1,\alpha}(\Omega))$, for some $\alpha > 0$, so that the application of the extension operator from Lemma 2.4 is well-defined. Since also f is Lipschitz continuous with Lipschitz constant L_f by assumption (A4), the function $f(\overline{\rho}(t, x))$ satisfies the Carathéodory conditions (see Definition A.1 in the Appendix) and thus there exists a solution in the sense of Definition 3.4; see Ch. I, Theorem 5.1 in [34]. Furthermore, as we also have $\overline{\rho} \in L^p(0, T; C^{0,1}(\mathbb{R}^2))$, we obtain with $C_u := \|u\|_{L^\infty(0,T)}$, $C_\rho(t) = \|\rho(t)\|_{W^{2,p}(\Omega)}$, and the property $C_\rho \in L^p(0, T)$ that

$$|f(\overline{\rho}(t, x)) - f(\overline{\rho}(t, y))| \leq L_f \, |\overline{\rho}(t, x) - \overline{\rho}(t, y)| \leq L_f \, \|\overline{\rho}(t)\|_{C^{1,\alpha}(\mathbb{R}^2)} \, |x - y|$$

$$\leq L_f \, C_E \, \|\rho(t)\|_{C^{1,\alpha}(\Omega)} \, |x - y|$$

$$\leq L_{f,\mathrm{E},\rho}(t)\,|x - y| \quad \text{for} \quad \text{a.a.} \quad t \in [0, T] \text{ and all } x, y \in \mathbb{R}^2, \tag{3.13}$$

where $L_{f,\mathrm{E},\rho}(t) := L_f\,C_E\,C_\infty\,C_\rho(t)$ and with C_∞ the embedding constant for $W^{2,p}(\Omega) \hookrightarrow C^{1,\alpha}(\overline{\Omega})$. The estimate (3.13) implies uniqueness by Ch. I, Thm. 5.3 in [34]. The additional regularity $x \in W^{1,\infty}(0, T; \mathbb{R}^2)$ is a consequence of the boundedness in $L^\infty(0, T; \mathbb{R}^2)$ of the right-hand side of (3.11). To establish the stability estimate we show

$$
\begin{aligned}
|x_1(t) - x_2(t)| &\leq \int_0^t |f(\overline{\rho}_1(s, x_1(s))) - f(\overline{\rho}_2(s, x_2(s)))|\,|u(s)|\,\mathrm{d}s \\
&\leq \int_0^t |f(\overline{\rho}_1(s, x_1(s))) - f(\overline{\rho}_1(s, x_2(s)))|\,|u(s)|\,\mathrm{d}s \\
&\quad + \int_0^t |f(\overline{\rho}_1(s, x_2(s))) - f(\overline{\rho}_2(s, x_2(s)))|\,|u(s)|\,\mathrm{d}s \\
&\leq \int_0^t C_u\,L_{f,\rho,\mathrm{E}}(s)\,|x_1(s) - x_2(s)|\,\mathrm{d}s + L_f\,C_u\,C_E\,C_{\mathrm{E},\infty}\,t\,\|\rho_1 - \rho_2\|_{L^\infty(Q_t)},
\end{aligned}
$$

where we used that the extension E is also continuous with respect to the L^∞-norm. An application of Gronwall's inequality in integral form then yields, for $t \in (0, T)$,

$$
\begin{aligned}
|x_1(t) - x_2(t)| &\leq L_f\,C_u\,C_E\,C_{\mathrm{E},\infty}\,t\,\|\rho_1 - \rho_2\|_{L^\infty(Q_T)}\,\exp\left(\int_0^t C_u\,L_{f,\rho,\mathrm{E}}(r)\,\mathrm{d}r\right) \\
&\leq C_s(u, L_f, L_{f,\rho,\mathrm{E}}, C_E, C_{\mathrm{E},\infty}, T)\,\|\rho_1 - \rho_2\|_{L^\infty(Q_T)}.
\end{aligned}
$$

Next, we state an existence and stability result for a regularized version of (3.11). Note that the following result requires less regularity for the density function ρ.

Lemma 3.6 *Fix* $2 < p < \infty$ *and* $\rho \in C([0, T]; L^2(\Omega))$. *Then for given* $u \in L^\infty(0, T)$, f *satisfying assumption (A4) and (C1), and every* $\gamma > 0$, *there exists, a unique, absolutely continuous solution* $x\colon [0, T] \to \mathbb{R}^2$ *to*

$$\dot{x}(t) = f((\overline{\varphi_\gamma * \rho})(t, x(t)))\,u(t) \tag{3.14}$$

satisfying $x(0) = x_0 \in \mathbb{R}^2$ *and* $x \in W^{1,\infty}(0, T)$. *Here,* φ_γ *is a standard mollifier as in* (2.2) *and* $*$ *denotes the convolution w.r.t. to the x-variable. Furthermore, for* $\rho_1, \rho_2 \in C([0, T]; L^2(\Omega))$, *the corresponding solutions* x_1 *and* x_2 *satisfy*

$$|x_1(t) - x_2(t)| \leq C_1\,T\,e^{C_2\,t}\,\|\rho_1 - \rho_2\|_{L^\infty(0, T; L^2(\Omega))}, \tag{3.15}$$

where the constants C_1, C_2 *depend on* $u, \gamma, C_E, C_{\mathrm{E},\infty}$ *and the Lipschitz constants of* f *and* ρ_1, ρ_2.

Proof As f and $\overline{\varphi_\gamma * \rho}$ are Lipschitz continuous (see assumption (A4) and Lemma A.2), the function $f((\overline{\varphi_\gamma * \rho})(t, x))$ satisfies the Carathéodory conditions of Definition 3.4

and thus there exists a solution in the sense of Definition 3.4; see Ch. I, Thm. 5.1 in [34]. In particular there exists a positive function $L_{f,\gamma,E} > 0$ such that

$$|f((\overline{\varphi_\gamma * \rho})(t, x)) - f((\overline{\varphi_\gamma * \rho})(t, y))| \leq L_{f,\gamma,E} \|\rho(t, \cdot)\|_{L^1(\Omega)} |x - y| \quad (3.16)$$

for a.a. $t \in [0, T]$ and all $x, y \in \Omega$. This implies uniqueness by Ch. I, Thm. 5.3 in [34].

To show the stability estimate (3.15) we proceed as in the proof of Lemma Lemma 3.5, employ the estimate (3.16), Young's inequality for convolutions $\|\varphi_\gamma * \rho\|_{L^\infty(\Omega)} \leq \|\varphi_\gamma\|_{L^2(\Omega)} \|\rho\|_{L^2(\Omega)}$, and obtain

$$
\begin{aligned}
|x_1(t) - x_2(t)| &\leq \int_0^t |f((\overline{\varphi_\gamma * \rho_1})(s, x_1(s))) - f((\overline{\varphi_\gamma * \rho_2})(s, x_1(s)))||u(s)| \, ds \\
&\quad + \int_0^t |f((\overline{\varphi_\gamma * \rho_2})(s, x_1(s))) - f((\overline{\varphi_\gamma * \rho_2})(s, x_2(s)))||u(s)| \, ds \\
&\leq L_f \, C_{u,E} \|\varphi_\gamma\|_{L^2(\Omega)} \, t \, \|\rho_1 - \rho_2\|_{L^\infty(0,T;L^2(\Omega))} \\
&\quad + L_{f,\gamma,E} \, C_u \|\rho_2\|_{L^\infty(0,T;L^2(\Omega))} \int_0^t |x_1(s) - x_2(s)| \, ds.
\end{aligned}
$$

Then, (3.15) follows from the Gronwall Lemma. $\qquad\square$

We continue with the following Lipschitz estimate for the transport term in (1.9a), which is needed in the theorem right after the next.

Lemma 3.7 *Given $\rho_1, \rho_2 \in C([0, T]; H^1(\Omega))$, $\phi_1, \phi_2 \in L^\infty(0, T; W^{2,p}(\Omega))$ and $x_1, x_2 \in L^\infty(0, T; \mathbb{R}^2)$ the function $\beta(\rho, \phi, x)$ satisfies the Lipschitz inequality*

$$
\begin{aligned}
&\|\rho_1 \, \beta(\rho_1, \phi_1, x_1) - \rho_2 \, \beta(\rho_2, \phi_2, x_2)\|_{L^2(Q_T)} \\
&\leq L_\beta \left(\|\rho_1 - \rho_2\|_{L^2(Q_T)} + \|\nabla\phi_1 - \nabla\phi_2\|_{L^2(Q_T)} + \|x_1 - x_2\|_{L^\infty(0,T;\mathbb{R}^2)^M} \right),
\end{aligned}
$$

where L_β depends on the Lipschitz and boundedness constants of f, h and K.

Proof We define $g(\rho) = \rho \, f(\rho)$ and write $\rho_i \, \beta(\rho_i, \phi_i, x_i) = g(\rho_i) \, h(\nabla(\phi_i + \phi_K(x_i; \cdot)))$, $i = 1, 2$. Using the Lipschitz continuity of g, h and ϕ_K we obtain

$$
\begin{aligned}
&\|\rho_1 \, \beta(\rho_1, \phi_1, x_1) - \rho_2 \, \beta(\rho_2, \phi_2, x_2)\|_{L^2(0,T;L^2(\Omega))} \\
&\leq L_g \|\rho_1 - \rho_2\|_{L^2(Q_T)} \|h(\nabla(\phi_1 + \phi_K(x_1; \cdot)))\|_{L^\infty(Q_T)} \\
&\quad + C_g \, L_h (\|\nabla(\phi_1 - \phi_2)\|_{L^2(Q_T)} + L_{\phi_K} \|x_1 - x_2\|_{L^\infty(0,T;\mathbb{R}^2)^M}).
\end{aligned}
$$

3.2 Existence for the Full Forward System

We are now in a position to show the following existence and uniqueness result.

Theorem 3.8 *Let assumptions (A1), (A2), (A3) (A4), (K1), (K2) (C1) hold and fix $2 < p < \infty$. Then for any given control $u = (u_1, \ldots, u_M)^T \in L^\infty(0, T; \mathbb{R}^2)^M$ and any*

$T > 0$, *there exists a unique solution* $(\rho, \phi, \boldsymbol{x})$ *to* (1.9) *with initial and boundary conditions* (1.4), (1.5), (1.10) *which satisfies* $\rho \in W_p^{2,1}(Q_T)$, $\phi \in L^\infty(0, T; W^{2,p}(\Omega))$ *and* $\boldsymbol{x}$ *is a solution to* (1.9c) *in the sense of Definition 3.4. Moreover, the a priori estimate*

$$\|\rho\|_{W_p^{2,1}(Q_T)} + \|\phi\|_{L^\infty(0,T;W^{2,p}(\Omega))} \le C \|\rho_0\|_{W^{1,p}(\Omega)},$$

holds with C depending only on the domain, the bounds for the coefficients and the respective kernel.

The structure of (1.9a)–(1.9b) is very similar to chemotaxis models with volume filling, see for instance [44], except for the additional nonlinearity of the squared gradient term in (1.9b), which can, however, be handled using Lemma 3.1. Therefore, the existence and uniqueness of solutions can be proved using Banach's fixed point theorem, similar to, e. g., Thm. 3.1 in [27]. The main issue in our situation is the additional coupling to the system of ODEs (1.9c), which requires ρ to be regular enough to allow point evaluations. Our strategy is to introduce an additional regularization in (1.9c) in order to be able to perform the fixed point argument in the relatively "large" space $C([0, T]; L^2(\Omega))$. We then show additional regularity and pass to the limit to recover the original system.

Proof The proof consists of two parts. First we show existence with (1.9c) replaced by the regularized version (3.14). Then, we perform the limit $\gamma \to 0$ to recover the original problem.

Step 1: Fixed point argument: We consider the fixed-point operator

$$S \colon C([0, T]; L^2(\Omega)) \to C([0, T]; L^2(\Omega)), \quad \widetilde{\rho} \mapsto \rho_\gamma, \qquad (3.17)$$

where, for fixed $\gamma > 0$, ρ_γ is the unique weak solution to the system

$$\partial_t \rho_\gamma - \nabla \cdot (\widetilde{\rho} \, \beta(\widetilde{\rho}, \phi_\gamma, \boldsymbol{x}_\gamma)) = \varepsilon \, \Delta \rho_\gamma \quad \text{in } Q_T, \qquad (3.18)$$

$$-\delta_1 \, \Delta \phi_\gamma + |\nabla \phi_\gamma|^2 = \frac{1}{f(\widetilde{\rho})^2 + \delta_2} \quad \text{in } Q_T, \qquad (3.19)$$

$$\dot{x}_{i,\gamma}(t) = f((\overline{\varphi_\gamma * \widetilde{\rho}})(t, x_{i,\gamma}(t))) \, u_i(t) \quad \text{for } t \in [0, T], \quad i = 1, \ldots, M \qquad (3.20)$$

subject to the boundary conditions

$$\varepsilon \, \nabla \rho_\gamma \cdot n = -\widetilde{\rho} \, \beta(\widetilde{\rho}, \phi_\gamma, \boldsymbol{x}_\gamma) \cdot n - \eta \, \rho_\gamma \, \chi_{\partial \Omega_O} \quad \text{on } \Sigma_T.$$

The boundary conditions for ϕ and the initial conditions for all variables are as in (1.5) and (1.4), (1.10), respectively. The equation for ρ_γ is understood in the weak sense, i.e.,

$$\int_\Omega \partial_t \rho_\gamma \, \xi \, \mathrm{d}x + \varepsilon \int_\Omega \nabla \rho_\gamma \cdot \nabla \xi \, \mathrm{d}x = -\int_\Omega \widetilde{\rho} \, \beta(\widetilde{\rho}, \phi_\gamma, \boldsymbol{x}_\gamma) \cdot \nabla \xi \, \mathrm{d}x - \eta \int_{\partial \Omega_O} \rho_\gamma \, \xi \, \mathrm{d}s_x$$
$$(3.21)$$

for all $\xi \in H^1(\Omega)$ and a.a. $t \in [0, T]$. Applying Lemma 3.3 with $g(\rho) = \rho f(\rho)$ and $h = h(\nabla(\phi_\gamma + \phi_K(\boldsymbol{x}_\gamma ; \cdot)))$ shows that there exists a unique weak solution

$$\rho_\gamma \in L^\infty(0, T; H^1(\Omega)) \cap H^1(0, T; L^2(\Omega)) \hookrightarrow C([0, T]; L^2(\Omega)).$$

Choosing ρ itself as a test function in the weak formulation (3.21) yields, after an application of the weighted Young's inequality with parameter $\varepsilon/2$, the trace inequality for H^1-functions, and an integration in time, the a priori estimate

$$\|\rho_\gamma\|_{C([0,T];L^2(\Omega))} \leq C(\|\widetilde{\rho}\,\beta(\widetilde{\rho}, \phi_\gamma, \boldsymbol{x}_\gamma)\|_{L^2(0,T;L^2(\Omega))} + \|\rho_0\|_{L^2(\Omega)})$$
$$\leq C(\sqrt{T\,|\Omega|} + \|\rho_0\|_{L^2(\Omega)}),$$

where we used $|\tilde{\rho}\,\beta(\tilde{\rho}, \phi_\gamma, \boldsymbol{x}_\gamma)| \leq 1$.

Next, we show that S is a contraction. Let $\widetilde{\rho}_1, \widetilde{\rho}_2 \in C([0, T]; L^2(\Omega))$ be arbitrary. The corresponding densities, potentials and agent positions are denoted by $\rho_{\gamma,1}, \rho_{\gamma,2}$, $\phi_{\gamma,1}, \phi_{\gamma,2}$ and $\boldsymbol{x}_{\gamma,1}, \boldsymbol{x}_{\gamma,2}$, respectively. Using Lemma 3.7 we have

$$\|\rho_{\gamma,1} - \rho_{\gamma,2}\|_{C([0,T];L^2(\Omega))}$$
$$\leq \|\widetilde{\rho}_1\,\beta(\widetilde{\rho}_1, \phi_{\gamma,1}, \boldsymbol{x}_{\gamma,1}) - \widetilde{\rho}_2\,\beta(\widetilde{\rho}_2, \phi_{\gamma,2}, \boldsymbol{x}_{\gamma,2})\|_{L^2(0,T;L^2(\Omega))}$$
$$\leq L_\beta(\|\widetilde{\rho}_1 - \widetilde{\rho}_2\|_{L^2(0,T;L^2(\Omega))} + \|\nabla(\phi_{\gamma,1} - \phi_{\gamma,2})\|_{L^2(0,T;L^2(\Omega))}$$
$$+ \|\boldsymbol{x}_{\gamma,1} - \boldsymbol{x}_{\gamma,2}\|_{L^\infty(0,T)^M}). \tag{3.22}$$

Applying Lemmas 3.1 and 3.6 to the second and third terms on the right-hand side, respectively, yields

$$\|\rho_{\gamma,1} - \rho_{\gamma,2}\|_{C([0,T];L^2(\Omega))} \leq C\,\max\{\sqrt{T}, T\,e^{C_2\,T}\}\,\|\widetilde{\rho}_1 - \widetilde{\rho}_2\|_{C([0,T];L^2(\Omega))}.$$

Thus, we can again find T small enough so that S is a contraction and Banach's fixed point theorem asserts the existence of a unique solution.

The box constraints $0 \leq \rho_\gamma \leq 1$ a.e. in Q_T follow by applying Lem. 3 in [27], i.e., by testing with smoothed versions of the positive part of $\rho - 1$ and $-\rho$, respectively. In view of these uniform estimates a standard continuation argument yields existence for arbitrary $T > 0$.

So far we have shown

$$\rho_\gamma \in L^2(0, T; H^1(\Omega)) \cap H^1(0, T; H^1(\Omega)^*) \cap L^\infty(Q_T),$$

since the fixed-point satisfies the weak formulation (3.21). Note in particular that ρ_γ is bounded in $L^2(0, T; H^1(\Omega)) \cap H^1(0, T; H^1(\Omega)^*) \cap L^\infty(Q_T)$ by a constant independent of γ. Furthermore, we denote the corresponding potential and agent trajectories by

$$\phi_\gamma \in L^\infty(0, T; W^{2,p}(\Omega)) \cap H^1(0, T; H^1(\Omega)), \quad \boldsymbol{x}_\gamma \in W^{1,\infty}(0, T)^M.$$

Step 2: Additional regularity: To shorten the notation we write the nonlinear term in the form

$$\rho_\gamma \, \beta(\rho_\gamma, \phi_\gamma, \boldsymbol{x}_\gamma) = g(\rho_\gamma) \, h(\Phi_\gamma)$$

with $g(\rho_\gamma) := \rho_\gamma \, f(\rho_\gamma)$ and $\Phi_\gamma := \nabla(\phi_\gamma + \phi_K(\boldsymbol{x}_\gamma; \cdot))$. From the product rule we obtain the following representation for the divergence,

$$\nabla \cdot (\rho_\gamma \, \beta(\rho_\gamma, \phi_\gamma, \boldsymbol{x}_\gamma)) = g'(\rho_\gamma) \, \nabla \rho_\gamma \cdot h(\Phi_\gamma) + \rho_\gamma \, f(\rho_\gamma) \, \nabla \cdot h(\Phi_\gamma).$$

Freezing the nonlinear terms allows us to understand (3.18) as a linear equation of the form

$$
\begin{aligned}
\partial_t \rho_\gamma - \varepsilon \, \Delta \rho_\gamma + b(t, x) \cdot \nabla \rho_\gamma + c(t, x) \, \rho_\gamma &= 0 \quad && \text{in } Q_T, \\
\varepsilon \nabla \rho_\gamma \cdot n &= r \quad && \text{on } \Sigma_T, \\
\rho(0) &= \rho_0 \quad && \text{in } \Omega
\end{aligned}
\tag{3.23}
$$

with

$$
\begin{aligned}
b(t, x) &:= g'(\rho_\gamma) \, h(\Phi_\gamma) \in L^p(0, T; L^\infty(\Omega)), \\
c(t, x) &:= f(\rho_\gamma) \, \nabla \cdot h(\Phi_\gamma) \in L^p(Q_T),
\end{aligned}
\tag{3.24}
$$

and

$$r(t, x) := -g(\rho_\gamma) \, h(\Phi_\gamma) \cdot n - \eta \, \rho_\gamma \, \chi_{\partial\Omega_O} \in W_p^{\kappa, \kappa/2}(\Sigma_T) \tag{3.25}$$

with $\kappa = 1 - 1/p$. It remains to show the regularity claimed for b, c and r. Again, this follows from the Hölder inequality and the regularity already shown for ρ_γ, ϕ_γ, $\boldsymbol{x}_\gamma$. Together with the product and the chain rule this leads to

$$\|b\|_{L^\infty(Q_T)} \le \|g'\|_{L^\infty(\mathbb{R})} \|h\|_{L^\infty(\mathbb{R}^2)} \le C_b, \tag{3.26}$$

$$\|c\|_{L^p(Q_T)} \le \|f\|_{W^{1,\infty}(\mathbb{R})} \|Dh\|_{W^{1,\infty}(\mathbb{R}^2)} (\|\phi_\gamma\|_{L^\infty(0,T;W^{2,p}(\Omega))} + C_{\phi_K}) \le C_c. \tag{3.27}$$

Note that C_b and C_c are independent of γ as, in particular, ϕ_γ can be bounded independently of ρ_γ and thus of γ. To show the required regularity for r we proceed as follows. First, we show the estimates

$$
\begin{aligned}
&\|\partial_{x_i} (g(\rho_\gamma) \, h(\Phi_\gamma))\|_{L^2(Q_T)} \\
&\quad \le C \, (\|g\|_{L^\infty(\mathbb{R})} \|Dh\|_{L^\infty(\mathbb{R}^{2\times 2})} \|\nabla^2(\phi_\gamma + \phi_K(\boldsymbol{x}_\gamma; \cdot))\|_{L^2(\Omega_T)} \\
&\quad\quad + \|g'\|_{L^\infty(\mathbb{R})} \|\nabla \rho_\gamma\|_{L^2(Q_T)} \|h\|_{L^\infty(\mathbb{R}^2)}) \le C
\end{aligned}
\tag{3.28}
$$

for $i = 1, 2$, as well as

$$
\begin{aligned}
&\|\partial_t (g(\rho_\gamma) \, h(\Phi_\gamma))\|_{L^2(Q_T)} \\
&\quad \le C \, (\|g\|_{L^\infty(\mathbb{R})} \|Dh\|_{L^\infty(\mathbb{R}^{2\times 2})} \|\partial_t \nabla(\phi_\gamma + \phi_K(\boldsymbol{x}_\gamma; \cdot))\|_{L^2(Q_T)} \\
&\quad\quad + \|g'\|_{L^\infty(\mathbb{R})} \|\partial_t \rho_\gamma\|_{L^2(Q_T)} \|h\|_{L^\infty(\mathbb{R}^2)}) \le C.
\end{aligned}
\tag{3.29}
$$

This follows from the regularity already shown for ρ_γ, ϕ_γ (see Lemma 3.1) and $\boldsymbol{x}_\gamma$ (see Lemma 3.6). With these considerations we conclude

$$-g(\rho_\gamma)\,h(\Phi_\gamma) - \eta\,\rho_\gamma\,\chi_{\partial\Omega_O} \in W_2^{1,1}(Q_T) \hookrightarrow W_2^{1,1/2}(Q_T).$$

This allows us to apply the trace Lemma 2.3, which provides

$$\|r\|_{W_2^{1/2,1/4}(\Sigma_T)} \leq C_r. \tag{3.30}$$

Collecting the properties (3.26)–(3.30) an application of Thm. 2.1 in [22] and Lemma 3.1 implies

$$\rho_\gamma \in W_2^{2,1}(Q_T) \quad \text{and} \quad \phi_\gamma \in L^\infty(0, T; H^3(\Omega)) \cap H^1(0, T; H^1(\Omega)).$$

We can even further improve the regularity of ρ_γ. Analogous to (3.28) and (3.29) we show

$$\|\partial_{x_i}(g(\rho_\gamma)\,h(\Phi_\gamma))\|_{L^p(Q_T)} + \|\partial_t(\partial_{x_i}g(\rho_\gamma)\,h(\Phi_\gamma))\|_{L^2(0,T;L^p(\Omega))} \leq C$$

and together with the trace Lemma 2.3 we deduce

$$r \in L^p(0, T; W^{1,p}(\Omega)) \cap W^{1,2}(0, T; L^p(\Omega))$$
$$\hookrightarrow W_p^{1,1/2}(Q_T) \hookrightarrow W_p^{1-1/p,1/2-1/(2p)}(\Sigma_T).$$

This, (3.26) and (3.27) allow a further application of Thm. 2.1 in [22] with $p \leq 4$, from which we infer the desired regularity

$$\rho_\gamma \in W_p^{2,1}(Q_T).$$

The functions ρ_γ, ϕ_γ and $\boldsymbol{x}_\gamma$ are thus a strong solution of the system

$$\partial_t \rho_\gamma - \nabla \cdot (\rho_\gamma\,\beta(\rho_\gamma, \phi_\gamma, \boldsymbol{x}_\gamma)) = \varepsilon\,\Delta\rho_\gamma \quad \text{in } Q_T, \tag{3.31}$$

$$-\delta_1\,\Delta\phi_\gamma + |\nabla\phi_\gamma|^2 = \frac{1}{f(\rho_\gamma)^2 + \delta_2} \quad \text{in } Q_T, \tag{3.32}$$

$$\dot{x}_{i,\gamma} = f((\overline{\varphi_\gamma * \rho_\gamma})(\cdot, x_{i,\gamma}(\cdot)))\,u_i \quad \text{in } [0, T], \quad i = 1, \ldots, M \tag{3.33}$$

that satisfy the boundary and initial conditions pointwise almost everywhere.

Step 3: Limit $\gamma \to 0$: In order to recover a solution to (1.9), it remains to pass to the limit $\gamma \to 0$ in (3.31)–(3.33). As a first step in this direction, note that $b(t, x)$, $c(t, x)$ and $r(t, x)$ defined in (3.24) and (3.25) are bounded independently of γ. Thus, understanding (3.31) as the linear equation (3.23) we obtain the following estimate, uniformly in γ,

$$\|\rho_\gamma\|_{W_p^{2,1}(Q_T)} \leq C.$$

This implies the existence of a sequence γ_k with $\gamma_k \to 0$ that

$$\rho_{\gamma_k} \rightharpoonup \rho \text{ in } W_p^{2,1}(Q_T) \quad \text{and} \quad \rho_{\gamma_k} \to \rho \text{ in } L^p(0, T; W^{1,p}(\Omega)), \qquad (3.34)$$

where the second convergence is a consequence of the Aubin-Lions lemma, Thm. II.5.16 in [10]. As, moreover, ϕ_γ is also uniformly bounded in $L^p(0, T; W^{2,p}(\Omega))$ by C_ϕ, we also have

$$
\begin{aligned}
D^2\phi_{\gamma_k} \rightharpoonup D^2\phi &\quad \text{in } L^p(0, T; L^p(\Omega; \mathbb{R}^{d \times d})), \\
\phi_{\gamma_k} \to \phi &\quad \text{in } L^p(0, T; W^{1,p}(\Omega)),
\end{aligned}
\qquad (3.35)
$$

where the second convergence follows from Lemma A.3. We omit the index k in the following to shorten the notation. Passing to the limit $\gamma \to 0$, in the sense of distributions, in equation (1.9b) yields the validity of this equation also for the limit values ϕ and ρ. As for (1.9c), denote by $x_i(t)$ the solution of $\dot{x}_i(t) = f(\overline{\rho}(x_i(t), t)) u_i(t)$, $t \in [0, T]$, and $x_i(0) = x_{i,0}$, with ρ denoting the limit from (3.34). Arguing similarly as in Lemma 3.6 and using $|u_i(t)| \le 1$ for a.a. $t \in (0, T)$ we obtain

$$
\begin{aligned}
&|x_i(t) - x_{i,\gamma}(t)| \\
&\le \int_0^t |f(\overline{\rho}(s, x_i(s))) - f((\overline{\varphi_\gamma * \rho_\gamma})(s, x_{i,\gamma}(s)))| \, ds \\
&\le \int_0^t |f(\overline{\rho}(s, x_i(s))) - f(\overline{\rho}(s, x_{i,\gamma}(s)))| \, ds \\
&\quad + \int_0^t |f(\overline{\rho}(s, x_{i,\gamma}(s))) - f((\overline{\varphi_\gamma * \rho})(s, x_{i,\gamma}(s)))| \, ds \\
&\quad + \int_0^t |f((\overline{\varphi_\gamma * \rho})(s, x_{i,\gamma}(s))) - f((\overline{\varphi_\gamma * \rho_\gamma})(s, x_{i,\gamma}(s)))| \, ds \\
&\le \int_0^t L_{f,\rho,\mathrm{E}}(s) |x_i(s) - x_{i,\gamma}(s)| \, ds \\
&\quad + L_f(t \|\rho - \varphi_\gamma * \rho\|_{C([0,T]\times\overline{\Omega})} + \sqrt{t} \, \|\varphi_\gamma\|_{L^1(\Omega)} \|\rho - \rho_\gamma\|_{L^2(0,T;L^\infty(\Omega))}).
\end{aligned}
$$

As $\|\varphi_\gamma\|_{L^1(\Omega)} = 1$, another application of Gronwall's inequality yields

$$
\begin{aligned}
&|x_i(t) - x_{i,\gamma}(t)| \\
&\quad \le C \, L_f(t \|\rho - \varphi_\gamma * \rho\|_{C([0,T]\times\overline{\Omega})} + \sqrt{t} \, \|\rho - \rho_\gamma\|_{L^2(0,T;L^\infty(\Omega))}) \\
&\quad \cdot \exp\left(\int_0^t L_{f,\rho,\mathrm{E}}(s) \, ds \right),
\end{aligned}
$$

and due to (2.3), (3.34) we see that as $\gamma \to 0$, the right-hand side converges to 0 and thus

$$x_{i,\gamma}(t) \to x_i(t) \text{ for every } t \in (0, T), \quad i = 1, \ldots, M. \qquad (3.36)$$

Next, we pass to the limit in (3.31). The convergence of the linear terms is a direct consequence of (3.34). For the convection term we note that it can be written as

$$\nabla \cdot (\rho_\gamma \, \beta(\rho_\gamma, \phi_\gamma, \boldsymbol{x}_\gamma)) = \nabla \cdot (g(\rho_\gamma) \, h(\Phi_\gamma))$$
$$= g'(\rho_\gamma) \, \nabla \rho_\gamma \cdot h(\Phi_\gamma) + g(\rho_\gamma) \, Dh(\Phi_\gamma) : \nabla^2 (\phi_\gamma + \phi_K(\boldsymbol{x}_\gamma; \cdot)).$$

As both ρ_γ and h are uniformly bounded, the convergences (3.34), (3.35) and (3.36) imply

$$\nabla \cdot (\rho_\gamma \, \beta(\rho_\gamma, \phi_\gamma, \boldsymbol{x}_\gamma)) \rightharpoonup \nabla \cdot (\rho \, \beta(\rho, \phi, \boldsymbol{x})) \quad \text{in } L^p(Q_T).$$

It remains to pass to the limit in the boundary conditions. The trace Lemma 2.3 implies that $\varepsilon \, \nabla \rho_\gamma \cdot n$ is bounded in $L^p(\Sigma_T)$ and thus converges weakly. The compact embedding $W_p^{2-1/p, 1-1/(2p)}(\Sigma_T) \hookrightarrow L^p(\Sigma_T)$ implies the convergences

$$\rho|_{\Sigma_T} \to \rho|_{\Sigma_T} \quad \text{and} \quad \nabla \phi_\gamma|_{\Sigma_T} \to \nabla \phi|_{\Sigma_T} \quad \text{in } L^p(\Sigma_T),$$

so that using the uniform boundedness of both ρ_γ and h as well as (3.36) we conclude

$$\varepsilon \, \nabla \rho_\gamma \cdot n + \rho_\gamma \, \beta(\rho_\gamma, \phi_\gamma, \boldsymbol{x}_\gamma) \cdot n \rightharpoonup \varepsilon \, \nabla \rho \cdot n + \rho \, \beta(\rho, \phi, \boldsymbol{x}) \cdot n \quad \text{in } L^p(\Sigma_T).$$

Thus, the weak limit of ρ_γ, namely ρ, is the strong solution of (1.9) with boundary conditions (1.5). This completes the proof. $\qquad\square$

The previous theorem allows us to introduce the solution operator

$$S: \mathcal{U} \to \mathcal{Y}, \quad \boldsymbol{u} \mapsto S(\boldsymbol{u}) := (\rho, \phi, \boldsymbol{x})$$

of (1.9) with boundary conditions (1.5) and initial conditions $\boldsymbol{x}(0) = \boldsymbol{x}^0$ and $\rho(\cdot, 0) = \rho_0$. The control and state spaces are

$$\mathcal{U} := L^\infty(0, T; \mathbb{R}^2)^M,$$
$$\mathcal{Y} := W_p^{2,1}(Q) \times (L^\infty(0, T; W_{\mathrm{ND}}^{2,p}(\Omega)) \cap W^{1,p}(0, T; W^{1,p}(\Omega))) \times W^{1,s}(0, T; \mathbb{R}^2)^M$$

with $1/s = 1/2 + 1/p$. Later, the operator S is referred to as *control-to-state operator*.

We conclude this section with an auxiliary result required later to show the existence of global solutions to an optimal control problem.

Lemma 3.9 *The operator $S: \mathcal{U} \to \mathcal{Y}$ is weakly sequentially continuous.*

Proof Given a weakly convergent sequence $(u_n)_{n \in \mathbb{N}} \subset \mathcal{U}$ with $u_n \rightharpoonup u$ one has to show that the corresponding states $(\rho_n, \phi_n, \boldsymbol{x}_n) = S(u_n)$ converge weakly in $\mathcal{Y}$ to $(\rho, \phi, \boldsymbol{x}) = S(u)$. This follows from the same arguments as those in step 3 of the proof of Theorem 3.8, together with the uniqueness of solutions.

4 The Linearized System

In order to prove necessary optimality conditions for the optimal control problems introduced later, we investigate the differentiability of the control-to-state operator S. The desired results follow from the implicit function theorem applied to the equation $e(y, u) = 0$ with $y = S(u) = (\rho, \phi, x)$. Here, e corresponds to the strong formulation of our forward system (1.9a)–(1.9c), more precisely, there holds

$$e : \mathcal{Y} \times \mathcal{U} \to \mathcal{Z}$$

with

$$\mathcal{Z} = L^p(Q_T) \times W_p^{1-1/p, 1/2 - 1/(2p)}(\Sigma_W) \times W^{2(1-1/p), p}(\Omega)$$
$$\times (L^\infty(0, T; L^p(\Omega)) \cap W^{1,p}(0, T; W^{1,p}(\Omega)^*)) \times W^{1,s}(0, T; \mathbb{R}^2)^M$$

defined by

$$
\begin{aligned}
e_1(y, u) &:= \partial_t \rho - \varepsilon \, \Delta \rho - \nabla \cdot (\rho \, \beta(y)) \\
e_2(y, u) &:= (\varepsilon \nabla \rho + \rho \, \beta(y)) \cdot n + \chi_{\partial \Omega_O} \, \eta \, \rho \\
e_3(y, u) &:= \rho(0) - \rho_0 \\
e_4(y, u) &:= -\delta_1 \, \Delta \phi + |\nabla \phi|^2 - \frac{1}{f(\rho)^2 + \delta_2} \\
e_{5,i}(y, u)(t) &:= x_i(t) - x_{i,0} - \int_0^t f(\overline{\rho}(s, x_i(s))) \, u_i(s) \, \mathrm{d}s, \quad i = 1, \ldots, M.
\end{aligned}
\tag{4.1}
$$

Notice that from here on we write $\beta(y)$ in place of $\beta(\rho, \phi, x)$. Recall that $p \in (2, 4)$ is a fixed number and the integrability index s of the space of agent trajectories is chosen such that $\frac{1}{s} = \frac{1}{2} + \frac{1}{p}$.

Lemma 4.1 *Let $y = (\rho, \phi, x) \in \mathcal{Y}$ and define $\Phi := \nabla(\phi + \phi_K(x; \cdot))$. Then a constant $C_h > 0$ exists such that*

$$\|h(\Phi)\|_{L^\infty(Q_T)} + \|h(\Phi)\|_{L^\infty(0, T; W^{1,p}(\Omega))} + \|\partial_t h(\Phi)\|_{L^p(Q_T)} \le C_h$$
$$\|D^k h(\Phi)\|_{L^\infty(Q_T)} + \|D^k h(\Phi)\|_{L^\infty(0, T; W^{1,p}(\Omega))} \le C_h, \quad k = 1, 2.$$

Note that
$$Dh(y) = \chi_{\{|y| \le 1\}, \varepsilon} \, \mathrm{id}$$

is a smoothed characteristic function. Moreover, the functions h and $D^k h$, $k = 1, 2$, are Lipschitz continuous. In particular, for any two vector fields $\Phi_1, \Phi_2 \in L^\infty(0, T; W^{1,p}(\Omega))$ there holds

$$\|\nabla \cdot (Dh(\Phi_1) - Dh(\Phi_2))\|_{L^\infty(0, T; L^p(Q_T))}$$
$$\le C_h (1 + \|D\Phi_1\|_{L^\infty(Q_T)}) \|D\Phi_1 - D\Phi_2\|_{L^\infty(0, T; L^p(\Omega))}. \tag{4.2}$$

Proof The result follows directly from the regularity properties of h.

To shorten the notation we also introduce the following constants,

$$C_\rho := \|\rho\|_{W_p^{2,1}(Q_T)}, \quad C_\phi := \|\phi\|_{L^\infty(0,T;W^{2,p}(\Omega))} + \|\phi\|_{W^{1,p}(0,T;W^{1,p}(\Omega))},$$
$$C_{\boldsymbol{x}} := \|\boldsymbol{x}\|_{W^{1,s}(0,T)}, \tag{4.3}$$

whose boundedness is guaranteed due to assumptions (A4), (K1), (K2) and (C1).

We confirm the assumptions of the implicit function theorem in the following lemmas.

Lemma 4.2 *The operator $e\colon \mathcal{Y} \times \mathcal{U} \to \mathcal{Z}$ is continuously Fréchet differentiable.*

Proof We start with equation e_1. Again, we write the nonlinear term in the form

$$\rho\,\beta(\boldsymbol{y}) = g(\rho)\,h(\Phi) \text{ with } \Phi = \nabla(\phi + \phi_K(\boldsymbol{x};\cdot)).$$

For given $(\rho, \phi, \boldsymbol{x}) \in \mathcal{Y}$ and $(\widetilde{\rho}, \widetilde{\phi}, \widetilde{\boldsymbol{x}}) \in \mathcal{Y}$, Taylor's formula with integral remainder yields

$$e_1(\rho + \widetilde{\rho}, \phi, \boldsymbol{x}; \boldsymbol{u}) - e_1(\rho, \phi, \boldsymbol{x}; \boldsymbol{u})$$
$$= \partial_t \widetilde{\rho} - \varepsilon\Delta\widetilde{\rho} - \nabla\cdot(g'(\rho)\,\widetilde{\rho}\,h(\Phi)) - \int_0^1 \nabla\cdot[(g'(\rho + s\,\widetilde{\rho}) - g'(\rho))\,\widetilde{\rho}\,h(\Phi)]\,\mathrm{d}s,$$

$$e_1(\rho, \phi + \widetilde{\phi}, \boldsymbol{x}; \boldsymbol{u}) - e_1(\rho, \phi, \boldsymbol{x}; \boldsymbol{u})$$
$$= -\nabla\cdot(g(\rho)\,Dh(\Phi)\,\nabla\widetilde{\phi}) - \int_0^1 \nabla\cdot[g(\rho)(Dh(\widetilde{\Phi}_\phi) - Dh(\Phi))\nabla\widetilde{\phi}]\,\mathrm{d}s,$$

$$e_1(\rho, \phi, \boldsymbol{x} + \widetilde{\boldsymbol{x}}; \boldsymbol{u}) - e_1(\rho, \phi, \boldsymbol{x}; \boldsymbol{u})$$
$$= \nabla\cdot(g(\rho)\,Dh(\Phi)\nabla^2 K(\cdot - x_i)\,\widetilde{x}_i)$$
$$+ \int_0^1 \nabla\cdot(g(\rho)[Dh(\widetilde{\Phi}_{\boldsymbol{x}})\,\nabla^2 K(\cdot - x_i - s\,\widetilde{x}_i) - Dh(\Phi)\,\nabla^2 K(\cdot - x_i)]\widetilde{x}_i)\,\mathrm{d}s,$$

for $\widetilde{\boldsymbol{x}} = (0, \ldots 0, \widetilde{x}_i, 0, \ldots 0)^{\mathrm{T}}$, $i = 1, \ldots, M$.

In the above equations we used the notation $\widetilde{\Phi}_\phi = \nabla(\phi + s\,\widetilde{\phi} + \phi_K(\boldsymbol{x};\cdot))$ and $\widetilde{\Phi}_{\boldsymbol{x}} = \nabla(\phi + \phi_K(\boldsymbol{x} + s\,\widetilde{\boldsymbol{x}};\cdot))$. In the following we derive bounds for the remainder terms (the terms depending nonlinearly on $\widetilde{\rho}$, $\widetilde{\phi}$ and $\widetilde{\boldsymbol{x}}$ in the above equations), which we denote by $r_{1,\rho}(\widetilde{\rho})$, $r_{1,\phi}(\widetilde{\phi})$ and $r_{1,x_i}(\widetilde{\boldsymbol{x}})$, $i = 1, \ldots, M$.

First, we apply the product rule, the Hölder inequality and the embedding $W_p^{2,1}(Q_T) \hookrightarrow L^\infty(Q_T)$, where we denote by C_∞ the maximum of the embedding constant and 1. These arguments yield

$$\|r_{1,\rho}(\widetilde{\rho})\|_{L^p(Q_T)}$$
$$\leq \int_0^1 \Big([\|(g''(\rho + s\,\widetilde{\rho}) - g''(\rho))\|_{L^\infty(Q_T)}\|\nabla\rho\|_{L^p(Q_T)}$$

$$+ s \, \|g''(\rho + s \, \widetilde{\rho})\|_{L^\infty(Q_T)} \|\nabla \widetilde{\rho}\|_{L^p(Q_T)}] \cdot \|\widetilde{\rho}\|_{L^\infty(Q_T)} \|h(\Phi)\|_{L^\infty(Q_T)}$$
$$+ \|g'(\rho + s \, \widetilde{\rho}) - g'(\rho)\|_{L^\infty(Q_T)}$$
$$\cdot \, (\|\nabla \widetilde{\rho}\|_{L^p(Q_T)} \|h(\Phi)\|_{L^\infty(Q_T)} + \|\widetilde{\rho}\|_{L^\infty(Q_T)} \|\nabla \cdot h(\Phi)\|_{L^p(Q_T)}) \Big) \, \mathrm{d}s$$
$$\leq C_g \, C_h (C_\rho + 3 \, C_\infty) \, \|\widetilde{\rho}\|^2_{W^{2,1}_p(Q_T)} = \mathrm{o}(\|\widetilde{\rho}\|_{W^{2,1}_p(Q_T)}). \tag{4.4}$$

Second, to show differentiability with respect to ϕ we confirm

$$r_{1,\phi}(\widetilde{\phi}) = \int_0^1 \Big[g'(\rho)\nabla\rho(Dh(\widetilde{\Phi}_\phi) - Dh(\Phi))\nabla\widetilde{\phi}$$
$$+ g(\rho)(s \, \nabla \cdot (Dh(\widetilde{\Phi}_\phi) - Dh(\Phi)) \cdot \nabla\widetilde{\phi} + (Dh(\widetilde{\Phi}_\phi) - Dh(\Phi)) : \nabla^2\widetilde{\phi} \Big] \, \mathrm{d}s.$$

With the Hölder inequality, the Lipschitz properties of Dh and D^2h, in particular (4.2), and the usual embeddings we can show

$$\|r_{1,\phi}(\widetilde{\phi})\|_{L^p(Q_T)} \leq C_g \, C_h \, C_\infty \, (C_\rho + 2) \, \|\widetilde{\phi}\|^2_{L^\infty(0,T;W^{2,p}(\Omega))}$$
$$= \mathcal{O}(\|\widetilde{\phi}\|_{L^\infty(0,T;W^{2,p}(\Omega))}).$$

Third, we derive an estimate for $r_{1,x_i}(\widetilde{\boldsymbol{x}})$ with $\widetilde{\boldsymbol{x}} = (0, \ldots, 0, \widetilde{x}_i, 0, \ldots, 0)$. We use the notation $\phi_K = \phi_K(\boldsymbol{x}; \cdot)$ and $\widetilde{\phi}_K = \phi_K(\boldsymbol{x} + s\widetilde{\boldsymbol{x}}; \cdot)$ as well as $K_i = K(\cdot - x_i)$ and $\widetilde{K}_i = K(\cdot - x_i - s \, \widetilde{x}_i)$, reformulate the remainder term by applying the product and chain rule and obtain

$$r_{1,x_i}(\widetilde{x}_i)$$
$$\leq \int_0^1 \Big(|g'(\rho) \, \nabla\rho| \, (|Dh(\widetilde{\Phi}_{\boldsymbol{x}})| \, |\nabla^2 K_i - \nabla^2 \widetilde{K}_i| + |Dh(\widetilde{\Phi}_{\boldsymbol{x}}) - Dh(\Phi)| \, |\nabla^2 K_i|)$$
$$+ g(\rho)\Big[|\nabla \cdot Dh(\widetilde{\Phi}_{\boldsymbol{x}})| \, |\nabla^2 \widetilde{K}_i - \nabla^2 K_i| + |Dh(\widetilde{\Phi})| \, |\nabla^3 \widetilde{K}_i - \nabla^3 K_i|$$
$$+ |\nabla^3 K_i| \, |Dh(\widetilde{\Phi}_{\boldsymbol{x}}) - Dh(\Phi)| + |\nabla^2 K_i| \, |\nabla \cdot (Dh(\widetilde{\Phi}_{\boldsymbol{x}}) - Dh(\Phi))| \Big] \Big) \, |\widetilde{x}_i| \, \mathrm{d}s.$$

Here, $|\cdot|$ is an arbitrary vector, matrix or tensor norm, depending on the argument. With the Hölder inequality, the regularity of $(\rho, \phi, \boldsymbol{x})$, in particular $\rho \in W^{2,1}_p(Q_T) \hookrightarrow L^\infty(0, T; W^{1,p}(\Omega)) \cap L^p(0, T; W^{1,\infty}(\Omega))$, and the Lipschitz properties of h and K we deduce

$$\|r_{1,x_i}(\widetilde{\boldsymbol{x}})\|_{L^p(Q_T)} \leq C_g \, (C_\rho + 4) \, C_h \, C_{\phi_K} \, 2 \, (1 + C_{\phi_K}) \|\widetilde{x}_i\|^2_{L^\infty(0,T;\mathbb{R}^2)}$$
$$= \mathrm{o}(\|\widetilde{x}_i\|_{L^\infty(0,T;\mathbb{R}^2)}).$$

The differentiability of e_2 can be shown with similar arguments. The Taylor formula yields

$$e_2(\rho + \widetilde{\rho}, \phi, \boldsymbol{x}; \boldsymbol{u}) - e_2(\rho, \phi, \boldsymbol{x}; \boldsymbol{u})$$

$$= (\varepsilon \nabla \widetilde{\rho} + \widetilde{\rho} \, \beta(\boldsymbol{y})) \cdot n + \chi_{\partial \Omega_O} \, \eta \, \widetilde{\rho}$$

$$+ \, g'(\rho) \, \widetilde{\rho} \, h(\varPhi) \cdot n + \int_0^1 (g'(\rho + s \, \widetilde{\rho}) - g'(\rho)) \, \widetilde{\rho} \, h(\varPhi) \cdot n \, \mathrm{d}s,$$

$$e_2(\rho, \phi + \widetilde{\phi}, \boldsymbol{x}; \boldsymbol{u}) - e_2(\rho, \phi, \boldsymbol{x}; \boldsymbol{u})$$

$$= \, g(\rho) \, Dh(\varPhi) \, \nabla \widetilde{\phi} \cdot n + \int_0^1 g(\rho)(Dh(\widetilde{\varPhi}_\phi) - Dh(\varPhi)) \nabla \widetilde{\phi} \cdot n \, \mathrm{d}s,$$

$$e_2(\rho, \phi, \boldsymbol{x} + \widetilde{\boldsymbol{x}}; \boldsymbol{u}) - e_2(\rho, \phi, \boldsymbol{x}; \boldsymbol{u})$$

$$= -g(\rho) \, Dh(\varPhi) \, \nabla^2 K(\cdot - x_i) \, \widetilde{x}_i \cdot n$$

$$- \int_0^1 g(\rho)(Dh(\widetilde{\varPhi}_{\boldsymbol{x}}) \, \nabla^2 K(\cdot - x_i - s \, \widetilde{x}_i) - Dh(\varPhi) \, \nabla^2 K(\cdot - x_i)) \, \widetilde{x}_i \cdot n \, \mathrm{d}s,$$

and it remains to estimate the remainder terms, i. e., the last terms on the right-hand sides of the previous equations. We abbreviate these terms by $r_{2,\rho}(\widetilde{\rho}) \cdot n$, $r_{2,\phi}(\widetilde{\phi}) \cdot n$ and $r_{2,\boldsymbol{x}}(\widetilde{\boldsymbol{x}}) \cdot n$, respectively.

First, we show $\|r_{2,\rho}(\widetilde{\rho}) \cdot n\|_{W_p^{1-1/p,1/2-1/(2p)}(\Sigma_T)} = \mathrm{o}(\|\widetilde{\rho}\|_{W_p^{2,1}(Q_T)})$. We wish to apply the trace Lemma 2.3, for which we have to show the required time and space regularity of the extension onto Q_T. First, note that there holds

$$\|r_{2,\rho}(\widetilde{\rho})\|_{L^p(0,T;W^{1,p}(\Omega))} = \mathrm{o}(\|\widetilde{\rho}\|_{W_p^{2,1}(Q_T)}), \tag{4.5}$$

which can be concluded from the same arguments as in (4.4). Moreover, for the time derivative we show

$$\partial_t r_{2,\rho}(\widetilde{\rho}) = \int_0^1 \Big([(g''(\rho + s \widetilde{\rho}) - g''(\rho)) \, \partial_t \rho + s \, g''(\rho + s \widetilde{\rho}) \, \partial_t \widetilde{\rho}] \widetilde{\rho} \, h(\varPhi)$$

$$+ \, (g'(\rho + s \, \widetilde{\rho}) - g'(\rho)) \, (\partial_t \widetilde{\rho} \, h(\varPhi) + \widetilde{\rho} \, \partial_t h(\varPhi)) \Big) \, \mathrm{d}s$$

and with the usual arguments we obtain for $s^{-1} = 2^{-1} + p^{-1}$

$$\|r_{2,\rho}(\widetilde{\rho})\|_{W^{1,s}(0,T;L^p(\Omega))} \le C_\infty \, C_g \, C_h \, (3 + C_\rho) \, \|\widetilde{\rho}\|_{W_p^{2,1}(Q_T)}^2 = \mathrm{o}(\|\widetilde{\rho}\|_{W_p^{2,1}(Q_T)}). \tag{4.6}$$

With the estimates (4.5) and (4.6), the embedding

$$L^p(0, T; W^{1,p}(\Omega)) \cap W^{1,s}(0, T; L^p(\Omega)) \hookrightarrow W_p^{1,1/2}(Q_T) \tag{4.7}$$

and the trace Lemma 2.3 we deduce

$$\|r_{2,\rho}(\widetilde{\rho}) \cdot n\|_{W_p^{1-1/p,1/2-1/(2p)}(\Sigma_T)} = \mathrm{o}(\|\widetilde{\rho}\|_{W_p^{2,1}(Q_T)}),$$

which implies the differentiability of e_2 with respect to ρ.

To show an estimate for $r_{2,\phi}(\widetilde{\phi}) \cdot n$ we proceed in a similar fashion. With analogous arguments we deduce the estimates

$$\|\partial_{x_i} r_{2,\phi}(\widetilde{\phi})\|_{L^p(Q_T)} \leq C \, \|\nabla \widetilde{\phi}\|^2_{L^\infty(0,T;W^{1,p}(\Omega))},$$

$$\|\partial_t r_{2,\phi}(\widetilde{\phi})\|_{L^s(0,T;L^p(Q_T))} \leq C \left(C_\rho \, \|\nabla \widetilde{\phi}\|^2_{L^\infty(Q_T)} + \|\partial_t \nabla \widetilde{\phi}\|_{L^p(Q_T)} \|\nabla \widetilde{\phi}\|_{L^\infty(Q_T)} \right)$$

and exploit again (4.7) to arrive at

$$\|r_{2,\phi}(\widetilde{\phi})\|_{W_p^{1-1/p,1/2-1/(2p)}(\Sigma_T)} = \mathcal{O}(\|\widetilde{\phi}\|_{L^\infty(0,T;W^{2,p}(\Omega))} + \|\widetilde{\phi}\|_{W^{1,p}(0,T;W^{1,p}(\Omega))}),$$

which confirms the differentiability of e_2 w.r.t. ϕ.

An analogous procedure is used to deduce an estimate for the remainder term $r_{2,x}(\widetilde{x}) \cdot n$. For each direction $\widetilde{x} = (0, \ldots 0, \widetilde{x}_i, 0, \ldots 0)^{\mathrm{T}}$, $i = 1, \ldots, M$, a direct calculation taking into account Lemma 4.1 and the Lipschitz continuity of the derivatives of K and h yields

$$\|\partial_{x_i} r_{2,x}(\widetilde{x})\|_{L^p(Q_T)} \leq C \, \|\widetilde{x}_i\|^2_{L^\infty(0,T;\mathbb{R}^2)},$$

$$\|\partial_t r_{2,x}(\widetilde{x})\|_{L^s(0,T;L^p(\Omega))} \leq C \, \|\widetilde{x}_i\|^2_{W^{1,s}(0,T;\mathbb{R}^2)}.$$

Using again (4.7) and the trace Lemma 2.3 leads to

$$\|r_{2,x_i}(\widetilde{x}) \cdot n\|_{W_p^{1-1/p,1/2-1/(2p)}(\Sigma_T)} = \mathcal{O}(\|\widetilde{x}_i\|_{W^{1,s}(0,T;\mathbb{R}^2)}),$$

which proves the differentiability of e_2 w.r.t. x.

The component e_3 is trivially differentiable. For e_4 we show

$$e_4(\rho + \widetilde{\rho}, \phi, x) - e_4(\rho, \phi, x)$$
$$= \frac{2 f(\rho) \, f'(\rho)}{(f^2(\rho) + \delta_2)^2} \widetilde{\rho} + \int_0^1 \left(\frac{2 \, f(\rho + s\widetilde{\rho}) \, f'(\rho + s\widetilde{\rho})}{(f(\rho + s\widetilde{\rho})^2 + \delta_2)^2} - \frac{2 \, f(\rho) \, f'(\rho)}{(f(\rho)^2 + \delta_2)^2} \right) \widetilde{\rho} \, ds,$$
$$e_4(\rho, \phi + \widetilde{\phi}, x) - e_4(\rho, \phi, x)$$
$$= \delta_1 \Delta \widetilde{\phi} + 2\nabla \phi^{\mathrm{T}} \nabla \widetilde{\phi} + |\nabla \widetilde{\phi}|^2.$$

Again, we denote the remainder terms (the terms which are nonlinear in $\widetilde{\rho}$ and $\widetilde{\phi}$) by $r_{4,\rho}(\widetilde{\rho})$ and $r_{4,\phi}(\widetilde{\phi})$. We show

$$\|r_{4,\phi}(\widetilde{\phi})\|_{L^\infty(0,T;L^p(\Omega))} \leq C\|\nabla \widetilde{\phi}\|^2_{L^\infty(Q_T)} \leq C\|\widetilde{\phi}\|^2_{L^\infty(0,T;W^{2,p}(\Omega))},$$

$$\|\partial_t r_{4,\phi}(\widetilde{\phi})\|_{L^p(0,T;W^{1,p}(\Omega)^*)} \leq C\|\nabla \widetilde{\phi}\|_{L^\infty(0,T;W^{1,p}(\Omega))} \|\partial_t \nabla \widetilde{\phi}\|_{L^p(Q_T)}.$$

and both estimates together imply

$$\|r_{4,\phi}(\widetilde{\phi})\|_{\mathcal{Z}_4} = \mathcal{O}(\|\widetilde{\phi}\|_{\mathcal{Y}_2}).$$

With similar arguments we can show the differentiability of e_4 with respect to ρ.

Finally we consider e_5, which is nonlinear in ρ and x_i. We confirm by a simple computation that

$$
e_5(\rho + \widetilde{\rho}, \phi, \mathbf{x}; \mathbf{u}; t) - e_5(\rho, \phi, \mathbf{x}; \mathbf{u}; t)
$$
$$
= \int_0^t f'(\overline{\rho}(\tau, x_i(\tau)))\, \overline{\widetilde{\rho}}(\tau, x_i(\tau))\, u_i(\tau)\, \mathrm{d}\tau
$$
$$
+ \int_0^1 \int_0^t (f'(\overline{(\rho + s\,\widetilde{\rho})}(\tau, x_i(\tau))) - f'(\overline{\rho}(\tau, x_i(\tau))))\overline{\widetilde{\rho}}(\tau, x_i(\tau))\, u_i(\tau)\, \mathrm{d}\tau\, \mathrm{d}s,
$$
$$
e_5(\rho, \phi, \mathbf{x} + \widetilde{\mathbf{x}}; \mathbf{u}; t) - e_5(\rho, \phi, \mathbf{x}; \mathbf{u}; t)
$$
$$
= \widetilde{x}_i(t) - \int_0^t f'(\rho(\tau, x_i(\tau)))\, \nabla\rho(\tau, x_i(\tau)) \cdot \widetilde{x}_i(\tau)\, u_i(\tau)\, \mathrm{d}\tau
$$
$$
- \int_0^1 \int_0^t \Big(f'(\overline{\rho}(\tau, (x_i + s\,\widetilde{x}_i)(\tau)))\, \nabla\overline{\rho}(\tau, (x_i + s\,\widetilde{x}_i)(\tau))
$$
$$
- f'(\overline{\rho}(\tau, x_i(\tau)))\, \nabla\overline{\rho}(\tau, x_i(\tau)) \Big) \cdot \widetilde{x}_i(\tau)\, u_i(\tau)\, \mathrm{d}\tau\, \mathrm{d}s
$$

for $\widetilde{\mathbf{x}} = (0 \ldots \widetilde{x}_i \ldots 0)^{\mathrm{T}}, i = 1, \ldots, M$. Again, the remainder terms (the terms depending nonlinearly on $\widetilde{\rho}$ and $\widetilde{\mathbf{x}}$ in the above equation) are denoted by $r_{5,\rho}(\widetilde{\rho})$ and $r_{5,\mathbf{x}}(\widetilde{\mathbf{x}})$.

Using the Lipschitz continuity of f', assumption (C1) and $\|\overline{\rho}\|_{L^\infty(\mathbb{R}^2)} \le \|\rho\|_{L^\infty(\Omega)}$ (see Lemma 2.4), we obtain

$$
|r_{5,\rho}(\widetilde{\rho}; t)| \le C_f\, C_{\mathrm{E},\infty}^2\, \|\widetilde{\rho}\|_{L^2(0,T;L^\infty(\Omega))}^2 = \mathrm{o}(\|\widetilde{\rho}\|_{W_p^{2,1}(Q_T)}). \tag{4.8}
$$

With similar arguments and the Hölder inequality with $1/s > 2/p$ we deduce for the temporal derivative of $r_{5,\rho}$

$$
\|\partial_t r_{5,\rho}(\widetilde{\rho})\|_{L^s(0,T;\mathbb{R}^2)}
$$
$$
\le \int_0^1 \|(f'(\overline{(\rho + s\widetilde{\rho})}(\cdot, x_i(\cdot))) - f'(\overline{\rho}(\cdot, x_i(\cdot))))\, \overline{\widetilde{\rho}}(\cdot, x_i(\cdot))\, u_i(\cdot)\|_{L^s(0,T;\mathbb{R}^2)}\, \mathrm{d}s
$$
$$
\le C_f\, \|\overline{\widetilde{\rho}}\|_{L^p(0,T;L^\infty(\mathbb{R}^2))}^2 \le C_f\, C_{\mathrm{E},\infty}^2\, \|\widetilde{\rho}\|_{L^p(0,T;L^\infty(\Omega))}^2 = \mathrm{o}(\|\widetilde{\rho}\|_{W_p^{2,1}(Q_T)}).
$$

For the remainder term $r_{5,\mathbf{x}}(\widetilde{\mathbf{x}})$ with $\widetilde{\mathbf{x}} = (0, \ldots 0, \widetilde{x}_i, 0, \ldots 0)^{\mathrm{T}}$ we show

$$
|r_{5,\mathbf{x}}(\widetilde{\mathbf{x}}; t)|
$$
$$
\le \|\widetilde{x}_i\|_{L^\infty(0,T;\mathbb{R}^2)} \cdot
$$
$$
\cdot \int_0^1 \int_0^t \Big(|f'(\overline{\rho}(\tau, (x_i + s\,\widetilde{x}_i)(t))) - f'(\overline{\rho}(\tau, x_i(t)))||\nabla\overline{\rho}(\tau, (x_i + s\,\widetilde{x}_i)(\tau))|
$$
$$
+ |f'(\overline{\rho}(\tau, x_i(\tau)))||\nabla\overline{\rho}(\tau, (x_i + s\,\widetilde{x}_i)(\tau)) - \nabla\overline{\rho}(\tau, x_i(\tau))| \Big)\, \mathrm{d}\tau\, \mathrm{d}s. \tag{4.9}
$$

Using Lipschitz estimates for f' and the mean value theorem we conclude

$$
\begin{aligned}
|f'(\overline{\rho}(\tau, (x_i + s\,\widetilde{x}_i)(\tau))) - f'(\overline{\rho}(\tau, x_i(\tau)))| &\leq C_f \operatorname{Lip}(\overline{\rho}(\tau, \cdot))\,|\widetilde{x}_i(\tau)| \\
&\leq C_f\,C_E\,\|\rho(\tau, \cdot)\|_{W^{2,p}(\Omega)}\,|\widetilde{x}_i(\tau)|,
\end{aligned}
\tag{4.10}
$$

$$
\begin{aligned}
|\nabla\overline{\rho}(\tau, (x_i + s\,\widetilde{x}_i)(\tau)) - \nabla\overline{\rho}(\tau, x_i(\tau))| &\leq \|\nabla\overline{\rho}(\tau, \cdot)\|_{C^{0,\alpha}(\mathbb{R}^2)}\,|\widetilde{x}_i(\tau)|^\alpha \\
&\leq C_E\,\|\rho(\tau, \cdot)\|_{W^{2,p}(\Omega)}\,|\widetilde{x}_i(\tau)|^\alpha.
\end{aligned}
\tag{4.11}
$$

In the above estimates we used $\|\overline{\rho}\|_{C^{0,1}(\mathbb{R}^2)} \leq \|\overline{\rho}\|_{C^{1,\alpha}(\Omega)}$, the continuity of $E\colon C^{1,\alpha}(\overline{\Omega}) \to C^{1,\alpha}(\mathbb{R}^2)$, see Lemma 2.4, and the embedding $W^{2,p}(\Omega) \hookrightarrow C^{1,\alpha}(\overline{\Omega})$, which is valid for $\alpha \in (0, 1/2]$ due to $p > 2$.

The insertion of (4.10) and (4.11) into (4.9) yields

$$
\begin{aligned}
|r_{5,\boldsymbol{x}}(\widetilde{\boldsymbol{x}}; t)| &\leq C_E\,C_f\,C_\rho\,\max\{t^{1-2/p}, t^{1-1/p}\}\left(C_\rho\,\|\widetilde{x}_i\|_{L^\infty(0,T;\mathbb{R}^2)}^2 + \|\widetilde{x}_i\|_{L^\infty(0,T;\mathbb{R}^2)}^{1+\alpha}\right) \\
&= o(\|\widetilde{x}_i\|_{L^\infty(0,T;\mathbb{R}^2)}).
\end{aligned}
$$

The above also uses an application of the Hölder inequality in time and $2/p < 1$.

Finally, the Hölder inequality in time with $1/s = 1/2 + 1/p$ yields the estimate

$$
\begin{aligned}
&\|\partial_t r_{5,\boldsymbol{x}}(\widetilde{\boldsymbol{x}})\|_{L^s(0,T;\mathbb{R}^2)} \\
&\quad \leq C_f\Big(\|\overline{\rho}(\cdot, (x_i + s\,\widetilde{x}_i)(\cdot)) - \overline{\rho}(\cdot, x_i(\cdot))\|_{L^2(0,T)}\,\|\nabla\overline{\rho}(\cdot, (x_i + s\,\widetilde{x}_i)(\cdot))\|_{L^p(0,T)} \\
&\qquad + \|\overline{\rho}(\tau, x_i(\tau))\|_{L^2(0,T)}\|\nabla\overline{\rho}(\cdot, (x_i + s\,\widetilde{x}_i)(\cdot)) \\
&\qquad - \nabla\overline{\rho}(\cdot, x_i(\cdot))\|_{L^p(0,T)}\Big)\|\widetilde{x}_i\|_{L^\infty(0,T;\mathbb{R}^2)} \\
&\quad = o(\|\widetilde{x}_i\|_{L^\infty(0,T;\mathbb{R}^2)}),
\end{aligned}
$$

where the last step follows again from (4.10) and (4.11). This confirms the partial differentiability of e_5.

Collecting the previous estimates proves the partial Fréchet differentiability of the operator e. Based on the properties of f, h and ϕ_K using again Lipschitz continuity and boundedness, the operators $\partial_\rho e$, $\partial_\phi e$ and $\partial_{x_i} e$ depend continuously on $(\rho, \phi, \boldsymbol{x})$, which implies the continuous Fréchet differentiability. $\qquad\square$

Next, we show that the operator $\partial_{\boldsymbol{y}} e(\boldsymbol{y}, \boldsymbol{u})$ is invertible.

Lemma 4.3 *The operator $\partial_y e(y; u)$ is bijective. That is, given $y = (\rho, \phi, x) \in \mathcal{Y}$ and $F = (F_1, \ldots, F_5)^{\mathrm{T}} \in \mathcal{Z}$, the system $\partial_y e(y, u)\, \widetilde{y} = F$ given by*

$$\partial_t \widetilde{\rho} - \varepsilon\, \Delta\widetilde{\rho} - \nabla \cdot (\widetilde{\rho}\, \beta(y) + \rho\, \frac{\partial\beta(y)}{\partial y}\widetilde{y}) = F_1 \quad in\ Q_T,$$

$$(\varepsilon\, \nabla\widetilde{\rho} + \widetilde{\rho}\, \beta(y) + \rho\, \frac{\partial\beta(y)}{\partial y}\widetilde{y}) \cdot n + \chi_{\partial\Omega_O}\, \eta\, \widetilde{\rho} = F_2 \quad on\ \Sigma_T,$$

$$\widetilde{\rho}(0) = F_3 \quad in\ \Omega,$$

$$-\delta_1\, \Delta\widetilde{\phi} + 2\nabla\phi^{\mathrm{T}}\nabla\widetilde{\phi} + \frac{2f(\rho)\, f'(\rho)}{(f^2(\rho) + \delta_2)^2}\widetilde{\rho} = F_4 \quad in\ Q_T,$$

$$\widetilde{x}_i(t) - \int_0^t f'(\overline{\rho}(s, x_i(s)))(\nabla\overline{\rho}(s, x_i(s))^{\mathrm{T}}\widetilde{x}_i(s) + \overline{\widetilde{\rho}}(s, x_i(s)))u_i(s)\, \mathrm{d}s = F_{5,i}(t),$$

$$(4.12)$$

for $t \in (0, T)$ and $i = 1, \ldots, M$, possesses a unique solution $\widetilde{y} = (\widetilde{\rho}, \widetilde{\phi}, \widetilde{x}) \in \mathcal{Y}$.

Proof The strategy of the proof is to apply Banach's fixed point theorem to the linear system (4.12) to avoid technicalities that arise from the fact that $\widetilde{x}(t)$ depends on $\widetilde{\rho}$ non-locally in time. To this end, we introduce three solution operators.

First, there is $F_\rho \colon \mathcal{Y} \to W_p^{2,1}(Q_T)$ which maps $\widetilde{y} = (\widetilde{\rho}, \widetilde{\phi}, \widetilde{x})$ to $\widehat{\rho} \in W_p^{2,1}(Q_T)$, which is defined as the solution to

$$\partial_t \widehat{\rho} - \varepsilon\Delta\widehat{\rho} - \nabla \cdot (\widetilde{\rho}\, \beta(y) + \rho\, \frac{\partial\beta(y)}{\partial y}\widetilde{y}) = F_1 \quad in\ Q_T,$$

$$(\varepsilon\nabla\widehat{\rho} + \widetilde{\rho}\, \beta(y) + \rho\, \frac{\partial\beta(\widetilde{y})}{\partial y}) \cdot n + \chi_{\partial\Omega_O}\, \eta\, \widetilde{\rho} = F_2 \quad on\ \Sigma_{\mathrm{W}}, \qquad (4.13)$$

$$\widehat{\rho}(0) = F_3 \quad in\ \Omega.$$

Second, we define the operator

$$F_\phi \colon W_p^{2,1}(Q_T) \to L^\infty(0, T; W_{\mathrm{ND}}^{2,p}(\Omega)) \cap H^1(0, T; H^1(\Omega)),$$

which maps $\widetilde{\rho}$ to the solution $\widetilde{\phi}$ of

$$-\delta_1\Delta\widetilde{\phi} + 2\nabla\phi^{\mathrm{T}}\nabla\widetilde{\phi} + \frac{2f(\rho)\, f'(\rho)}{(f^2(\rho) + \delta_2)^2}\widetilde{\rho} = F_4 \quad in\ Q_T, \qquad (4.14)$$

together with the boundary conditions (1.5) (these are incorporated in the function space).

Third, we introduce $F_x \colon W_p^{2,1}(Q_T) \to W^{1,\infty}([0, T]; \mathbb{R}^n)^M$ mapping $\widetilde{\rho}$ to the solution $\widetilde{x} = (\widetilde{x}_1, \ldots, \widetilde{x}_M)^{\mathrm{T}}$ of

$$\widetilde{x}_i(t) = \int_0^t f'(\overline{\rho}(s, x_i(s)))(\nabla\overline{\rho}(s, x_i(s))^{\mathrm{T}}\widetilde{x}_i(s) + \overline{\widetilde{\rho}}(s, x_i(s)))u_i(s)\, \mathrm{d}s + F_{5,i}(t),$$

for $i = 1, \ldots, M$.

The idea of the proof is to apply Banach's fixed point theorem in the space $W_p^{2,1}(Q_T)$ to the operator

$$\widehat{\rho} = F(\widetilde{\rho}) := F_\rho(\widetilde{\rho}, F_\phi(\widetilde{\rho}), F_x(\widetilde{\rho})). \tag{4.15}$$

As the operator F is affine, it suffices to show the boundedness of the linear part of F by a constant smaller than 1, which will imply that $F(\widetilde{\rho})$ is a contraction.

Step 1: Estimate for F_ϕ. Due to $\phi(t) \in W^{2,p}(\Omega) \hookrightarrow W^{1,\infty}(\Omega)$ for $p > 2$ and a.a. $t \in (0, T)$, we confirm that $\nabla\phi \in L^\infty(Q_T)$ holds. Moreover, taking into account $f \in W^{1,\infty}(\mathbb{R})$ and $\rho \in W^{2,p}(Q_T) \hookrightarrow L^\infty(Q_T)$, we get

$$\|\frac{2f(\rho)\,f'(\rho)}{(f^2(\rho) + \delta_2)^2}\,\widetilde{\rho}\|_{L^\infty(0,T;L^p(\Omega))} \le 2\,C_f\,C_\rho\,\delta_2^{-2}\,\|\widetilde{\rho}\|_{L^\infty(0,T;L^p(\Omega))}.$$

From Thm. 3.17 in [49], see also Prop. 2 in [41] for an application to the linearized Eikonal equation, we then deduce that the problem (4.14) possesses for a.a. $t \in (0, T)$ a unique solution $\widetilde{\phi}(t) \in W_{\mathrm{ND}}^{2,p}(\Omega)$, which fulfills the inequality

$$\|\widetilde{\phi}\|_{L^\infty(0,T;W^{2,p}(\Omega))} \le C(\|\widetilde{\rho}\|_{L^\infty(0,T;L^p(\Omega))} + \|F_4\|_{L^\infty(0,T;L^p(\Omega))}). \tag{4.16}$$

Furthermore, formal differentiation of (4.14) with respect to time and exploiting $\partial_t F_4 \in L^p(0, T; W^{1,p}(\Omega)^*)$ yields

$$\|\partial_t\widetilde{\phi}\|_{L^p(0,T;W^{1,p}(\Omega))}$$
$$\le C\,(\|\widetilde{\rho}\|_{W_p^{2,1}(Q_T)} + \|F_4\|_{L^\infty(0,T;L^p(\Omega))} + \|\partial_t F_4\|_{L^p(0,T;W^{1,p}(\Omega)^*)}). \tag{4.17}$$

In the second estimate C may depend on T, via norms of $\partial_t\phi$, $\partial_t\rho$, etc.. Since this implies that C is decreasing when T is decreasing, this does not affect the contraction argument.

Step 2: Estimate for F_x. Next, we show the boundedness of the solution operator F_x. To this end, we define $C_\rho(t) := \|\rho(t, \cdot)\|_{W^{2,p}(\Omega)}$ and deduce from the last equation in (4.12), using the Hölder inequality with $1/p + 1/q = 1$ and the continuity of $\mathrm{E}\colon W^{2,p}(\Omega) \hookrightarrow C^{1,\alpha}(\overline{\Omega}) \to C^{1,\alpha}(\mathbb{R}^2)$, see Lemma 2.4:

$$|\widetilde{x}_i(t)| \le \int_0^t |f'(\overline{\rho}(s, x_i(s)))\left(\nabla\overline{\rho}(s, x_i(s))^{\mathrm{T}}\widetilde{x}_i(s) + \overline{\widetilde{\rho}}(s, x_i(s))\right) u_i(s)|\,\mathrm{d}s + |F_{5,i}(t)|$$

$$\le C_f\,C_{\mathrm{E}}\,C_\infty\big(\int_0^t C_\rho(s)\,|\widetilde{x}_i(s)|\,\mathrm{d}s + t^{1/q}\,\|\widetilde{\rho}\|_{W_p^{2,1}(Q_t)}\big) + \|F_{5,i}\|_{L^\infty(0,t)}.$$

Due to the Gronwall inequality we obtain

$$|\widetilde{x}_i(t)| \le \big(C_f\,C_{\mathrm{E}}\,C_\infty t^{1/q}\,\|\widetilde{\rho}\|_{W_p^{2,1}(Q_T)} + \|F_{5,i}\|_{L^\infty(0,t)}\big)\exp(C_f\,C_{\mathrm{E}}\,C_\infty\int_0^t C_\rho(s)\,\mathrm{d}s).$$

We can further estimate $\int_0^t C_\rho(s)\,\mathrm{d}s \le t^{1/q}\,\|\rho\|_{W_p^{2,1}(Q_t)}$. The constants depending only on input data and the linearization point ρ are shifted into the generic constants c, C, which depend on T in monotonically decreasing way only. This implies

$$\|\widetilde{x}_i\|_{L^\infty(0,T)} \le C\,e^{c\,T^{1/q}}(T^{1/q}\,\|\widetilde{\rho}\|_{W_p^{2,1}(Q_T)} + \|F_{5,i}\|_{L^\infty(0,T)}). \qquad (4.18)$$

From now on we will choose $T > 0$ sufficiently small such that $C\,T^{1/q}\,e^{cT^{1/q}} < 1$.

Furthermore, we can show L^s-regularity for the time derivative of $\widetilde{x}_i$. From the classical formulation of the ODE and the usual Hölder arguments we deduce

$$
\begin{aligned}
|\dot{\widetilde{x}}_i(t)| &= |f'(\overline{\rho}(t, x_i(t)))(\nabla\overline{\rho}(t, x_i(t))^{\mathrm{T}}\widetilde{x}_i(t) + \overline{\widetilde{\rho}}(t, x_i(t)))u_i(t) + \dot{F}_5(t)| \\
&\le C_f\,C_{\mathrm{E}}(C_\rho(t)\,|\widetilde{x}_i(t)| + C_{\widetilde{\rho}}(t)) + |\dot{F}_{5,i}(t)|
\end{aligned}
$$

for a.a. $t \in [0, T]$. Taking the $L^s(0, T; \mathbb{R}^2)$-norm yields, after insertion of the estimate (4.18) and an application of the Hölder inequality with $1/s = 1/2 + 1/p$,

$$
\begin{aligned}
\|\widetilde{x}_i\|_{W^{1,s}(0,T;\mathbb{R}^2)} &\le C_f\,C_{\mathrm{E}}\,T^{1/2}(\|\rho\|_{W_p^{2,1}(Q_T)}\,\|\widetilde{x}_i\|_{L^\infty(0,T;\mathbb{R}^2)} + \|\widetilde{\rho}\|_{W_p^{2,1}(Q_T)}) \\
&\quad + \|F_{5,i}\|_{W^{1,s}(0,T;\mathbb{R}^2)} \\
&\le C\,T^{1/2}\,\|\widetilde{\rho}\|_{W_p^{2,1}(Q_T)} + (1 + C\,e^{c\,T^{1/q}})\|F_{5,i}\|_{W^{1,s}(0,T;\mathbb{R}^2)}. \quad (4.19)
\end{aligned}
$$

Step 3: Estimate for F_ρ: Next, we consider the solution operator of (4.13). By Lemma 3.2 we have the a priori estimate

$$
\begin{aligned}
\|\widehat{\rho}\|_{W_p^{2,1}(Q_T)} \le C\Bigg[&\|F_1\|_{L^p(Q_T)} + \|F_2\|_{W_p^{1-1/p,1/2-1/(2p)}(\Sigma_{\mathrm{W}})} + \|F_3\|_{W^{2(1-1/p),p}(\Omega)} \\
&+ \|\nabla \cdot (\widetilde{\rho}\,\beta(\mathbf{y}) + \rho\,\frac{\partial\beta(\mathbf{y})}{\partial y}\widetilde{y})\|_{L^p(Q_T)} \\
&+ \|(\widetilde{\rho}\,\beta(\mathbf{y}) + \rho\,\frac{\partial\beta(\mathbf{y})}{\partial y}\widetilde{y}) \cdot n + \chi_{\partial\Omega_{\mathrm{O}}}\,\eta\,\widetilde{\rho}\|_{W_p^{1-1/p,1/2-1/(2p)}(\Sigma_T)} \Bigg]. \quad (4.20)
\end{aligned}
$$

It remains to discuss the last two terms on the right-hand side. First, we confirm with the product rule, the Hölder inequality and Lemma 4.1 that the estimate

$$
\begin{aligned}
\|\nabla \cdot (\widetilde{\rho}\,\beta(\mathbf{y}))\|_{L^p(Q_T)} &\le C_h\,T^{1/p}(\|\nabla\widetilde{\rho}\,f(\rho) + \widetilde{\rho}\,f'(\rho)\,\nabla\rho\|_{L^\infty(0,T;L^p(\Omega))} + \|\widetilde{\rho}\,f(\rho)\|_{L^\infty(Q_T)}) \\
&\le C\,T^{1/p}\,\|\widetilde{\rho}\|_{W_p^{2,1}(Q_T)}
\end{aligned}
\qquad (4.21)
$$

holds. Note that we also exploited regularity results for the linearization point, more precisely, $f(\rho), \rho \in W_p^{2,1}(Q_T) \hookrightarrow L^\infty(0, T; W^{1,p}(\Omega)) \hookrightarrow L^\infty(Q_T)$ (see (2.4)), $\phi \in L^\infty(0, T; W^{2,p}(\Omega))$, and the assumptions on f, h and ϕ_K.

Next, we show an estimate for the term involving $\rho \frac{\partial \beta(\mathbf{y})}{\partial \mathbf{y}} \widetilde{\mathbf{y}}$ and we study all partial derivatives of β separately. For the derivative with respect to ρ we get

$$
\begin{aligned}
&\|\nabla \cdot (\rho \frac{\partial \beta(\mathbf{y})}{\partial \rho} \widetilde{\rho})\|_{L^p(Q_T)} \\
&= \|\nabla \cdot (\rho \, f'(\rho) \, \widetilde{\rho} \, h(\Phi))\|_{L^p(Q_T)} \\
&\leq C_h(\|\widetilde{\rho} \, \nabla \rho \, (f'(\rho) + \rho \, f''(\rho)) + \rho \, \nabla \widetilde{\rho} \, f'(\rho)\|_{L^p(Q_T)} \\
&\quad +)\|\rho \, f'(\rho) \, \widetilde{\rho}\|_{L^p(0,T;L^\infty(\Omega))} \\
&\leq C \, T^{1/p} \|\widetilde{\rho}\|_{W_p^{2,1}(Q_T)},
\end{aligned}
\tag{4.22}
$$

where the last step follows from the Hölder inequality, $f \in W^{2,\infty}(\mathbb{R})$, Lemma 4.1 and the embedding $W_p^{2,1}(Q_T) \hookrightarrow L^\infty(0,T;W^{1,p}(\Omega)) \hookrightarrow L^\infty(Q_T)$, taking into account the regularity of the linearization point $(\rho, \phi, \mathbf{x})$.

For the terms involving $\frac{\partial \beta}{\partial \phi}$ we first confirm by a simple computation

$$
\frac{\partial h(\Phi)}{\partial \phi} \widetilde{\phi} = Dh(\Phi)\nabla\widetilde{\phi}.
$$

This implies with the product rule

$$
\begin{aligned}
&\|\nabla \cdot (\rho \frac{\partial \beta(\mathbf{y})}{\partial \phi} \widetilde{\phi})\|_{L^p(Q_T)} \\
&\leq \|g'(\rho) \nabla\rho \cdot (Dh(\Phi) \nabla\widetilde{\phi}) + g(\rho) \nabla \cdot (Dh(\Phi) \nabla\widetilde{\phi})\|_{L^p(Q_T)}
\end{aligned}
\tag{4.23}
$$

with $g(\rho) := \rho \, f(\rho)$ and $g'(\rho) = \rho \, f'(\rho) + f(\rho)$. Due to $\widetilde{\phi} \in L^\infty(0,T;W^{2,p}(\Omega))$ and thus, $\nabla\widetilde{\phi} \in L^\infty(0,T;W^{1,p}(\Omega)) \hookrightarrow L^\infty(Q_T)$, as well as $\nabla\rho \in L^\infty(0,T;L^p(\Omega))$ and Lemma 4.1, we get for the first term on the right-hand side of (4.23)

$$
\begin{aligned}
&\|g'(\rho)\nabla\rho \cdot (Dh(\Phi) \nabla\widetilde{\phi})\|_{L^p(Q_T)} \\
&\leq T^{1/p} \, C_g \, C_h \, \|\nabla\rho\|_{L^\infty(0,T;L^p(\Omega))}\|\nabla\widetilde{\phi}\|_{L^\infty(Q_T)} \leq C \, T^{1/p} \, \|\widetilde{\phi}\|_{L^\infty(0,T;W^{2,p}(\Omega))}.
\end{aligned}
$$

To bound the second term in (4.23), we insert the identity

$$
\nabla \cdot (Dh(\Phi)\nabla\widetilde{\phi}) = (\nabla \cdot Dh(\Phi)) \cdot \nabla\widetilde{\phi} + Dh(\Phi) : \nabla^2\widetilde{\phi}
$$

and obtain, together with the bounds for Dh, from Lemma 4.1

$$
\begin{aligned}
&\|g(\rho) \, \nabla \cdot (Dh(\Phi) \, \nabla\widetilde{\phi})\|_{L^p(Q_T)} \\
&\leq C_g \, C_h \, T^{1/p} \, (\|\nabla\widetilde{\phi}\|_{L^\infty(Q_T)} + \|\nabla^2\widetilde{\phi}\|_{L^\infty(0,T;L^p(\Omega))}) \\
&\leq C \, T^{1/p} \, \|\widetilde{\phi}\|_{L^\infty(0,T;W^{2,p}(\Omega))}.
\end{aligned}
$$

Insertion of the two previous estimates into (4.23) then yields

$$\|\nabla \cdot (\rho \frac{\partial \beta(\boldsymbol{y})}{\partial \phi} \widetilde{\phi})\|_{L^p(Q_T)} \le C \, T^{1/p} \, \|\widetilde{\phi}\|_{L^\infty(0,T;W^{2,p}(\Omega))}. \tag{4.24}$$

The derivative with respect to x_i, $i = 1, \ldots, M$, is handled using the chain rule, Lemma 4.1 and assumption (K2), leading to

$$\|\nabla \cdot (\rho \frac{\partial \beta(\boldsymbol{y})}{\partial x_i} \widetilde{x}_i)\|_{L^p(Q_T)} = \|\nabla \cdot (g(\rho) \, Dh(\Phi) \, \nabla^2 K(\cdot - x_i) \, \widetilde{x}_i)\|_{L^p(Q_T)}$$

$$\le C \, T^{1/p} \, \|\widetilde{x}_i\|_{L^\infty(0,T;\mathbb{R}^2)}. \tag{4.25}$$

The estimates (4.21), (4.22), (4.24) and (4.25) give an estimate for the fourth term in (4.20), namely

$$\|\nabla \cdot (\widetilde{\rho} \, \beta(\boldsymbol{y}) + \rho \frac{\partial \beta(\boldsymbol{y})}{\partial \boldsymbol{y}} \widetilde{\boldsymbol{y}})\|_{L^p(Q_T)} \le C \, T^{1/p} \, \|\widetilde{\boldsymbol{y}}\|_{\mathcal{Y}}. \tag{4.26}$$

It remains to show the boundedness of the fifth term in (4.20) and we apply a trace theorem as in the proof of Lemma 4.2, see (4.7). This requires us to show appropriate bounds in the $L^p(0, T; W^{1,p}(\Omega))$- and $W^{1,s}(0, T; L^p(\Omega))$-norms.

First, when replacing $\nabla \cdot$ by an arbitrary partial derivative ∂_{x_j} in the previous considerations, we get analogously to (4.26)

$$\|\widetilde{\rho} \, \beta(\boldsymbol{y}) + \rho \frac{\partial \beta(\boldsymbol{y})}{\partial \boldsymbol{y}} \widetilde{\boldsymbol{y}}\|_{L^p(0,T;W^{1,p}(\Omega))} \le C \, T^{1/p} \, \|\widetilde{\boldsymbol{y}}\|_{\mathcal{Y}}. \tag{4.27}$$

Furthermore, we study the temporal derivative of the expressions occurring in the fifth term of (4.20). The Hölder inequality and Lemma 4.1 yield

$$\|\partial_t (\widetilde{\rho} \, \beta(\boldsymbol{y}))\|_{L^s(0,T;L^p(\Omega))}$$
$$\le T^{1/2} \, \|\partial_t (\widetilde{\rho} \, f(\rho) \, h(\Phi))\|_{L^p(Q_T)}$$
$$\le C_f \, C_h \, T^{1/2} (\|\partial_t \widetilde{\rho}\|_{L^p(Q_T)} + \|\widetilde{\rho}\|_{L^\infty(Q_T)} \, (\|\partial_t \rho\|_{L^p(Q_T)} + 1))$$
$$\le C_f \, C_h \, C_\infty \, (C_\rho + 1) \, T^{1/2} \, \|\widetilde{\rho}\|_{W_p^{2,1}(Q_T)}.$$

For the term involving $\frac{\partial \beta(\boldsymbol{y})}{\partial \rho} \widetilde{\rho}$ we get

$$\left\|\partial_t (\rho \frac{\partial \beta(\boldsymbol{y})}{\partial \rho} \widetilde{\rho})\right\|_{L^s(0,T;L^p(\Omega))}$$
$$\le T^{1/2} \, \|\partial_t (\rho \, f'(\rho) \, \widetilde{\rho} \, h(\Phi))\|_{L^p(Q_T)}$$
$$\le C_f \, C_h \, T^{1/2} \, (2 \, \|\partial_t \rho\|_{L^p(Q_T)} \, \|\widetilde{\rho}\|_{L^\infty(Q_T)}$$
$$\quad +) \|\rho\|_{L^\infty(Q_T)} \, (\|\partial_t \widetilde{\rho}\|_{L^p(Q_T)} + \|\widetilde{\rho}\|_{L^\infty(Q_T)})$$
$$\le C_f \, C_h \, C_\rho \, C_\infty \, T^{1/2} \, \|\widetilde{\rho}\|_{W_p^{2,1}(Q_T)}.$$

For the term depending on $\frac{\partial \beta(\mathbf{y})}{\partial \phi} \widetilde{\phi}$ we obtain

$$
\begin{aligned}
\left\| \partial_t \left(\rho \, \frac{\partial \beta(\mathbf{y})}{\partial \phi} \widetilde{\phi} \right) \right\|_{L^s(0,T;L^p(\Omega))} &\leq T^{1/2} \| \partial_t (g(\rho) \, Dh(\Phi) \, \nabla \widetilde{\phi}) \|_{L^p(Q_T)} \\
&\leq C_g \, C_h \, T^{1/2} \left((\| \partial_t \rho \|_{L^p(Q_T)} + 1) \| \nabla \widetilde{\phi} \|_{L^\infty(Q_T)} + \| \partial_t \nabla \widetilde{\phi} \|_{L^p(Q_T)} \right) \\
&\leq C_g \, C_h (2 + C_\rho) \, T^{1/2} \left(\| \widetilde{\phi} \|_{L^\infty(0,T;W^{2,p}(\Omega))} + \| \partial_t \widetilde{\phi} \|_{L^p(0,T;W^{1,p}(\Omega))} \right).
\end{aligned}
$$

Finally, for the derivative with respect to x_i we derive

$$
\begin{aligned}
\left\| \partial_t \left(\rho \, \frac{\partial \beta(\mathbf{y})}{\partial x_i} \widetilde{x}_i \right) \right\|_{L^s(0,T;L^p(\Omega))} & \\
= \| \partial_t (g(\rho) \, Dh(\Phi) \, \nabla^2 K(\cdot - x_i) \, \widetilde{x}_i) \|_{L^s(0,T;L^p(\Omega))} & \\
\leq C_g \, C_h \, C_K \, C_x \big((C_\rho + 1) \, T^{1/2} \, \| \widetilde{x}_i \|_{L^\infty(0,T;\mathbb{R}^2)} &+ 2 \, \| \widetilde{x}_i \|_{W^{1,s}(0,T;\mathbb{R}^2)} \big).
\end{aligned}
$$

The previous four estimates and an embedding yield the desired property

$$
\begin{aligned}
\left\| \widetilde{\rho} \, \beta(\mathbf{y}) + \rho \, \frac{\partial \beta(\mathbf{y})}{\partial \mathbf{y}} \widetilde{\mathbf{y}} \right\|_{W^{1/2,p}(0,T;L^p(\Omega))} &\leq C \left\| \widetilde{\rho} \, \beta(\mathbf{y}) + \rho \, \frac{\partial \beta(\mathbf{y})}{\partial \mathbf{y}} \widetilde{\mathbf{y}} \right\|_{W^{1,s}(0,T;L^p(\Omega))} \\
&\leq C \, T^{1/2} (\| \widetilde{\rho} \|_{\mathcal{Y}_1} + \| \widetilde{\phi} \|_{\mathcal{Y}_2}) + C \, \| \widetilde{\mathbf{x}} \|_{\mathcal{Y}_3}.
\end{aligned}
\tag{4.28}
$$

Next, we may combine the estimates (4.28) and (4.27) and apply the trace Lemma 2.3 to arrive at

$$
\begin{aligned}
\left\| \left(\widetilde{\rho} \, \beta(\mathbf{y}) + \rho \, \frac{\partial \beta(\mathbf{y})}{\partial \mathbf{y}} \widetilde{\mathbf{y}} \right) \cdot n + \chi_{\partial \Omega_O} \, \eta \, \widetilde{\rho} \right\|_{W_p^{1-1/p,1/2-1/(2p)}(\Sigma_T)} & \\
\leq C \, T^{1/p} (\| \widetilde{\rho} \|_{\mathcal{Y}_1} + \| \widetilde{\phi} \|_{\mathcal{Y}_2}) + C \, \| \widetilde{\mathbf{x}} \|_{\mathcal{Y}_3}. &
\end{aligned}
\tag{4.29}
$$

From the estimates (4.26) and (4.29) we deduce that there exists a unique solution $\widehat{\rho} \in W_p^{2,1}(Q_T)$ of (4.13). Moreover, together with (4.20) using $p > 2$, we obtain for sufficiently small $T > 0$ the a priori estimate

$$
\begin{aligned}
\| \widehat{\rho} \|_{W_p^{2,1}(Q_T)} &\\
\leq C \Big(T^{1/p} (\| \widetilde{\rho} \|_{\mathcal{Y}_1} + \| \widetilde{\phi} \|_{\mathcal{Y}_2}) + \| \widetilde{\mathbf{x}} \|_{\mathcal{Y}_3} &\\
+ \| F_1 \|_{L^p(Q_T)} + \| F_2 \|_{W_p^{1-1/p,1/2-1/(2p)}(\Sigma_W)} &+ \| F_3 \|_{W^{2(1-1/p),p}(\Omega)} \Big).
\end{aligned}
\tag{4.30}
$$

Step 4: The fixed-point argument

Now, we can combine the estimates (4.16), (4.17), (4.19) and (4.30) to deduce the continuity of the operator $F(\widetilde{\rho}) = F_\rho(\widetilde{\rho}, F_\phi(\widetilde{\rho}), F_x(\widetilde{\rho}))$. This gives the estimate

$$
\| F(\widetilde{\rho}) \|_{W_p^{2,1}(Q_T)} \leq C \, \| F \|_{\mathcal{Z}} + C \, T^{1/p} \, \| \widetilde{\rho} \|_{W_p^{2,1}(Q_T)}
$$

for $T \le 1$.

As F is an affine linear operator we directly conclude the contraction property provided that T is sufficiently small. Thus, Banach's fixed-point theorem provides the existence of a unique solution $\widetilde{\rho} \in W_p^{2,1}(Q_T)$ which can be extended to an arbitrary time-scale with a concatenation argument.

The regularities of the corresponding potential $\widetilde{\phi}$ and the agent trajectories $\widetilde{x}_i$, $i = 1, \ldots, M$, follow from (4.16), (4.17) and (4.19). $\qquad \blacksquare$

The previous two lemmas, together with the implicit function theorem, imply the differentiability of the control-to-state operator. The derivative in a direction $\widetilde{u} \in L^\infty(0, T; \mathbb{R}^2)^M$ can be computed by means of

$$e_y(y, u)\,\widetilde{y} = -e_u(y, u)\,\widetilde{u}.$$

We summarize the final result in the following theorem.

Theorem 4.4 *The control-to-state operator* $S: \mathcal{U} \to \mathcal{Y}$ *is Fréchet-differentiable and* $\widetilde{y} = S'(u)\,\widetilde{u}$ *for given* $\widetilde{u} \in \mathcal{U}$ *is characterized by the unique solution of*

$$
\begin{aligned}
&\partial_t \widetilde{\rho} - \varepsilon \Delta \widetilde{\rho} \\
&\quad - \nabla \cdot \left(\widetilde{\rho}\,\beta(\rho, \phi, x) + \rho \left(\frac{\partial \beta(\rho, \phi, x)}{\partial \rho}\widetilde{\rho} + \frac{\partial \beta(\rho, \phi, x)}{\partial \phi}\widetilde{\phi} + \frac{\partial \beta(\rho, \phi, x)}{\partial x}\widetilde{x} \right) \right) = 0, \\
&\quad - \delta_1 \Delta \widetilde{\phi} + 2\nabla \phi \cdot \nabla \widetilde{\phi} + \frac{2 f(\rho)\,f'(\rho)}{(f^2(\rho) + \delta_2)^2}\widetilde{\rho} = 0, \\
&\quad \dot{\widetilde{x}}_i - f'(\overline{\rho}(\cdot, x_i))\left(\nabla \overline{\rho}(\cdot, x_i)^{\mathsf{T}}\widetilde{x}_i + \overline{\widetilde{\rho}}(\cdot, x_i) \right) u_i = f(\overline{\rho}(\cdot, x_i))\,\widetilde{u}_i,
\end{aligned}
$$
$$(4.31)$$

for $i = 1, \ldots, M$, *together with the boundary conditions* (1.5) *and homogeneous initial conditions*

$$\widetilde{\rho}(\cdot, 0) = 0 \quad and \quad \widetilde{x}_i(0) = 0, \quad i = 1, \ldots, M.$$

5 Optimal Control Problem

In this section we come back to the optimal control problem outlined in Sect. 1.1. We consider objective functionals of the form

$$J(\rho, \phi, x; u) := \Psi(\rho, \phi, x) + \frac{\alpha}{2T}\|u\|_{H^1(0,T;\mathbb{R}^2)^M}^2, \qquad (5.1)$$

where $\alpha > 0$ is a regularization parameter and the functional $\Psi: \mathcal{Y} \to \mathbb{R}$ fulfills the following assumption:

(J1) The functional $\Psi: \mathcal{Y} \to \mathbb{R}$ is weakly lower semi-continuous and bounded from below on $\{(\rho, \phi, x) \in \mathcal{Y}: \rho \ge 0 \text{ a.e. in } Q_T\}$.

The assumed weak lower semi-continuity is fulfilled, e.g., when Ψ is convex and continuous. As the density part ρ of each solution of the forward model (1.9) is nonnegative it suffices to assume boundedness of Ψ on a subset only.

We consider the following optimal control problem:

$$
\begin{cases}
\text{Minimize} \quad J(\rho, \phi, \boldsymbol{x}; \boldsymbol{u}) \quad \text{where } (\rho, \phi, \boldsymbol{x}; \boldsymbol{u}) \in \mathcal{Y} \times \mathcal{U} \\
\text{subject to} \quad (\rho, \phi, \boldsymbol{x}) = S(\boldsymbol{u}) \\
\qquad \text{and} \quad \boldsymbol{u} \in \mathcal{U}_{\text{ad}}.
\end{cases}
\tag{5.2}
$$

The set of admissible controls is defined by

$$
\mathcal{U}_{\text{ad}} := \{ \boldsymbol{u} \in \mathcal{U} \,|\, \|u_i\|_{L^\infty(0,T)} \leq 1 \text{ for } i = 1, \ldots, M \}.
$$

This general framework covers in particular the two applications mentioned in Sect. 1

With standard arguments, see, e. g., Ch. 4.4 in [50], (J1) and Lemma 3.9 imply the following existence result for (5.2).

Theorem 5.1 *Let assumptions (A1)–(A4), (K1)–(K2), (C1) and (J1) hold. Then problem (5.2) possesses at least one global solution $\boldsymbol{u}^* \in \mathcal{U}_{ad}$.*

Moreover, with the properties shown for S in Sects. 3 and 4, in particular the Fréchet differentiability, we may deduce the following first-order necessary optimality condition, which is a direct consequence of the chain rule applied to the reduced optimization problem with objective $j(\boldsymbol{u}) := J(S(\boldsymbol{u}); \boldsymbol{u})$ and subject to $\boldsymbol{u} \in \mathcal{U}_{\text{ad}}$:

Theorem 5.2 *Let assumptions (A1)–(A4), (K1)–(K2), (C1) and (J1) hold. Then each local minimizer $(\boldsymbol{y}^*, \boldsymbol{u}^*) \in \mathcal{Y} \times \mathcal{U}_{ad}$ of (5.2) fulfills*

$$
\langle \Psi'(\boldsymbol{y}^*), \delta \boldsymbol{y} \rangle_{\mathcal{Y}^* \times \mathcal{Y}} + \frac{\alpha}{T} (\boldsymbol{u}^*, \delta \boldsymbol{u})_{H^1(0,T)} \geq 0 \quad \text{for all } \delta \boldsymbol{u} \in \mathcal{T}_{\mathcal{U}_{ad}}(\boldsymbol{u}^*),
$$

where $\boldsymbol{y}^* = S(\boldsymbol{u}^*)$ and $\delta \boldsymbol{y} = S'(\boldsymbol{u}^*)\delta \boldsymbol{u}$ and S' is characterized by the system from Theorem 4.4. Moreover, $\mathcal{T}_{\mathcal{U}_{ad}}(\boldsymbol{u}^*)$ denotes the tangent cone to $\mathcal{U}_{ad}$ at $\boldsymbol{u}^*$.

The optimality conditions from Theorem 5.2 can be rewritten by expressing $\langle \Psi'(\boldsymbol{y}^*), \delta \boldsymbol{y} \rangle_{\mathcal{Y}^* \times \mathcal{Y}}$ in terms of a suitably defined adjoint state. This then serves as the starting point for gradient-based optimization methods. Such methods, as well as appropriate discretization of the forward and adjoint systems and numerical results will be discussed in a forthcoming publication.

Declarations

A Auxiliary Results

First we present the proof of Lemma 3.3. It is based on choosing derivatives of the solution as test functions and a Galerkin approximation as well as the uniqueness of weak solutions, see also Ch. III, Sect. 6 in [42].

Proof of Lemma 3.3 Existence in $L^2(0, T; H^1(\Omega)) \cap H^1(0, T; H^1(\Omega)^*)$ follows by a standard Galerkin approximation. To this end we introduce $\rho^N(t, x) = \sum_{i=0}^{N} d_i(t) \varphi_i(x)$, where $(\varphi_i)_{i \in \mathbb{N}}$ is an orthonormal basis in $H^1(\Omega)$ which is also orthogonal w.r.t. $L^2(\Omega)$. The coefficients satisfy $d_i(0) = (\rho_0, \varphi_i)_{H^1(\Omega)}$. Choosing ρ^N as a test function results in

$$\int_0^T \frac{\mathrm{d}}{\mathrm{d}t} \int_\Omega (\rho^N)^2 \, \mathrm{d}t \, \mathrm{d}x + \varepsilon \int_{Q_T} |\nabla \rho^N|^2 \, \mathrm{d}t \, \mathrm{d}x$$
$$\leq \int_{Q_T} |\nabla \rho^N| |g(\rho^N) \, h| \, \mathrm{d}t \, \mathrm{d}x - \eta \int_{\Sigma_O} (\rho^N)^2 \, \mathrm{d}t \, \mathrm{d}s_x$$
$$\leq \frac{\varepsilon}{2} \int_{Q_T} |\nabla \rho^N|^2 \, \mathrm{d}t \, \mathrm{d}x + \frac{1}{2\varepsilon} \int_{Q_T} |g(\rho^N) \, h|^2 \, \mathrm{d}t \, \mathrm{d}x, \qquad \text{(A.1)}$$

i.e.

$$\|\rho^N\|_{L^\infty(0,T;L^2(\Omega))} + \|\nabla \rho^N\|_{L^2(Q_T)} \leq C(\|h\|_{L^2(0,T;L^2(\Omega))} + \|\rho_0\|_{L^2(\Omega)}), \quad \text{(A.2)}$$

where we used $\|\rho^N(\cdot, 0)\|_{L^2(\Omega)} \leq C \|\rho_0\|_{L^2(\Omega)}$. This readily implies that ρ^N is bounded, uniformly in N, in

$$L^2(0, T; H^1(\Omega)) \cap H^1(0, T; H^1(\Omega)^*) \cap C([0, T]; L^2(\Omega)),$$

see, e.g., Ch. 7.1.2, Thm. 2 in [29]. A subsequent application of the Aubin-Lions Lemma (Thm. II.5.16 in [10]), ensures the convergence of ρ^N strongly in $L^2(Q_T)$ to ρ and thus the convergence of the nonlinear term

$$\int_{Q_T} g(\rho^N) \, h \cdot \nabla \xi \, \mathrm{d}t \, \mathrm{d}x \to \int_{Q_T} g(\rho) \, h \cdot \nabla \xi \, \mathrm{d}t \, \mathrm{d}x$$

as $N \to \infty$ which follows by dominated convergence, noting that $g(\rho^N) h \cdot \nabla \xi$ is uniformly bounded in $L^1(0, T; L^1(\Omega))$. Hence the limit satisfies the weak formulation (3.10).

To show uniqueness, assume that there are two solutions ρ_1 and ρ_2. Testing each equation with $\rho_1 - \rho_2$, integration w.r.t. to t, and using the Lipschitz continuity of g together with an application of the weighted Cauchy inequality yields

$$
\frac{1}{2} \int_\Omega (\rho_1(t, \cdot) - \rho_2(t, \cdot))^2 \, dx + \frac{\varepsilon}{2} \int_{Q_t} |\nabla(\rho_1 - \rho_2)|^2 \, dt \, dx
$$

$$
\leq \frac{1}{2\varepsilon} \|h\|_{L^\infty(0,t;L^\infty(\Omega))} \int_{Q_t} (\rho_1 - \rho_2)^2 \, dt \, dx - \eta \int_{\Sigma_{D,t}} (\rho_1 - \rho_2)^2 \, dt \, ds_x
$$

for a.a. $t \in (0, T)$. As the last term on the right-hand side is non-positive, the uniqueness follows from Gronwall's inequality as both ρ_1 and ρ_2 have the same initial datum.

To show the additional regularity, we choose, for finite N, the test function $\partial_t \rho^N$ and integrate w.r.t. time. We obtain

$$
\int_{Q_t} (\partial_t \rho^N)^2 \, dt \, dx + \varepsilon \int_{Q_t} \nabla \rho^N \cdot \nabla(\partial_t \rho^N) \, dt \, dx - \int_{Q_t} g(\rho^N) h \cdot \nabla(\partial_t \rho^N) \, dt \, dx
$$

(A.3)

$$
= -\eta \int_{\Sigma_{D,t}} \rho_N \, \partial_t \rho_N \, dt \, ds_x.
$$

(A.4)

For the second term on the left-hand side we have

$$
\varepsilon \int_{Q_t} \nabla \rho^N \cdot \nabla(\partial_t \rho^N) \, dt \, dx = \frac{\varepsilon}{2} \int_0^t \frac{d}{dt} \int_\Omega |\nabla \rho^N|^2 \, dt \, dx
$$

$$
= \frac{\varepsilon}{2} \int_\Omega |\nabla \rho^N(t, \cdot)|^2 \, dx - \frac{\varepsilon}{2} \int_\Omega |\nabla \rho^N(0, x)|^2 \, dx,
$$

while for the third one, integration by parts in time together with the (weighted) Cauchy's inequality gives

$$
- \int_{Q_t} g(\rho^N) h \cdot \nabla(\partial_t \rho^N) \, dt \, dx
$$

$$
= \int_{Q_t} \partial_t \rho^N g'(\rho^N) h \cdot \nabla \rho^N \, dt \, dx
$$

$$
+ \int_{Q_t} g(\rho^N) \partial_t h \cdot \nabla \rho^N \, dt \, dx - \int_\Omega |g(\rho^N) h \cdot \nabla \rho^N \, dx_0^t
$$

$$
\leq \frac{1}{2} \|\partial_t \rho^N\|^2_{L^2(0,T;L^2(\Omega))} + \frac{1}{2}(\|g'(\rho^N)h\|^2_{L^\infty(0,T;L^\infty(\Omega))} + 1)\|\nabla \rho^N\|^2_{L^2(0,T;L^2(\Omega))}
$$

$$
+ \|g(\rho^N) \partial_t h\|^2_{L^2(0,T;L^2(\Omega))} + \frac{\varepsilon}{4} \|\nabla \rho^N(t, \cdot)\|^2_{L^2(\Omega)} + \frac{1}{4\varepsilon} \|g(\rho^N) h\|^2_{L^\infty(0,T;L^2(\Omega))}
$$

$$
+ \|g(\rho(0, \cdot)^N) h(0, \cdot) \nabla \rho^N(0, \cdot)\|^2_{L^2(\Omega)}.
$$

The term on the right-hand side of (A.3) becomes

$$-\eta \int_{\Sigma_{D,t}} \rho_N \partial_t \rho_N \, dt \, ds_x = -\frac{\eta}{2} \int_{\Omega} (\rho^N(t,x))^2 \, dx + \frac{\eta}{2} \int_{\Omega} (\rho^N(x,0))^2 \, dx.$$

Combining all these estimates, an application of the trace theorem applied to $\|\rho_0\|_{L^2(\partial\Omega_O)}$ and using

$$\|\rho^N(\cdot, 0)\|_{H^1(\Omega)} \le \|\rho_0\|_{H^1(\Omega)}$$

yields

$$\frac{1}{2} \|\partial_t \rho^N\|^2_{L^2(0,T;L^2(\Omega))} + \frac{\varepsilon}{4} \|\nabla \rho^N\|^2_{L^\infty(0,T;L^2(\Omega))}$$
$$\le C\|\rho_0^N\|^2_{H^1(\Omega)} + c \|\nabla \rho^N\|^2_{L^2(0,T;L^2(\Omega))} + d, \tag{A.5}$$

where

$$c = \frac{1}{2}(\|g'(\rho^N) \, h\|^2_{L^\infty(0,T;L^\infty(\Omega))} + 1),$$
$$C = \|g(\rho(0,\cdot)^N) \, h(0,\cdot)\|^2_{L^\infty(\Omega)},$$
$$d = \frac{1}{2}\|g(\rho^N) \, \partial_t h\|^2_{L^2(0,T;L^2(\Omega))} + \frac{1}{4\varepsilon}\|g(\rho^N) \, h\|^2_{L^\infty(0,T;L^2(\Omega))}.$$

Using (A.2), we can estimate the second term on the right-hand side which yields the desired regularity for ρ^N. We also infer the existence of a weakly converging subsequence in $L^2(0, T; L^2(\Omega)) \cap L^\infty(0, T; H^1(\Omega))$ which, by the weak lower semi-continuity of the norms, yields the bound also for the limit ρ. Reinserting this into (A.5) yields the assertion.

We also state the definition of Carathéodory conditions.

Definition A.1 We say that a function $r : Q_T \to \mathbb{R}^2$ satisfies the Carathéodory conditions whenever

- for each fixed $x \in \mathbb{R}^d$, the function $t \mapsto r(t, x)$ is measurable,
- for a.a. $t \in [0, T]$, the function $x \mapsto r(t, x)$ is continuous,
- there exists an integrable function m s.t.

$$|r(t, x)| \le m(t) \text{ for all } x \in \Omega \text{ and a.a.} t \in [0, T].$$

Lemma A.2 *Recall that φ_γ is a standard mollifier. For given $\rho \in L^\infty(0, T; L^1(\Omega))$, the function $\rho_\gamma := \varphi_\gamma * \rho$ is Lipschitz continuous in x with the Lipschitz constant depending on γ and $\|\rho\|_{L^1(\Omega)} \in L^\infty(0, T)$, only.*

Proof By definition, φ_γ is Lipschitz continuous and we denote its Lipschitz constant by L_γ. Then, for a.a. $t \in (0, T)$, we have

$$|\rho_\gamma(t, x) - \rho_\gamma(t, y)| \leq \int_\Omega |\varphi_\gamma(x - z) - \varphi_\gamma(y - z)| |\rho(t, z)| \, dz$$

$$\leq L_\gamma \int_\Omega |x - y| |\rho(t, z)| \, dz \leq L_\gamma \|\rho(t, \cdot)\|_{L^1(\Omega)} |x - y|.$$

Lemma A.3 *[Compactness of ϕ_γ] Fix $2 \leq p < \infty$ and denote by $(\rho_{\gamma_k})_k \in W_p^{2,1}(Q_T)$ a bounded family of functions, i.e. there exists a constant C_ρ, independent of γ, s.t.*

$$\|\rho_{\gamma_k}\|_{W_p^{2,1}(Q_T)} \leq C_\rho.$$

Denoting by ϕ_{γ_k} the sequence of solutions to (1.9b) with ρ_{γ_k} on the right-hand side, there exists an element $\phi \in L^p(0, T; W^{1,p}(\Omega))$ and a subsequence, again denoted by ϕ_{γ_k}, s.t.

$$\phi_{\gamma_k} \to \phi \text{ in } L^p(0, T; W^{1,p}(\Omega)). \tag{A.6}$$

Proof Instead of (1.9b), we work with the transformed version (3.6) and denote its solution by ψ_γ. We seek to apply a variant of the Aubin-Lions Lemma, introduced in [48], where instead of time derivatives we have to control finite differences in time. Indeed, as $W^{2,p}(\Omega) \hookrightarrow\hookrightarrow W^{1,p}(\Omega)$, it suffices to show

$$\int_0^{T-h} \|\psi_\gamma(t + h, \cdot) - \psi_\gamma(t, \cdot)\|_{W^{1,p}(\Omega)}^p \, dt \leq \mathcal{O}(h).$$

To this end, note that the difference $\psi_\gamma(t + h, \cdot) - \psi_\gamma(t, \cdot)$ satisfies the equation

$$\Delta(\psi_\gamma(t + h, \cdot) - \psi_\gamma(t, \cdot)) + q_t(\psi_\gamma(t + h, \cdot) - \psi_\gamma(t, \cdot)) = (1 - \psi(t, \cdot))(q_{t+h} - q_t),$$

with boundary conditions (3.6) and the corresponding a priori estimate

$$\|\psi_\gamma(t + h, \cdot) - \psi_\gamma(t, \cdot)\|_{W^{1,p}(\Omega)} \leq C(C_\phi)\|q_{t+h} - q_t\|_{L^p(\Omega)}.$$

This yields

$$\int_0^{T-h} \|\psi_\gamma(t + h, \cdot) - \psi_\gamma(t, \cdot)\|_{W^{1,p}(\Omega)}^p \, dt \leq C \int_0^{T-h} \|q_{t+h}(\cdot) - q_t(\cdot)\|_{W^{1,p}(\Omega)}^p \, dt. \tag{A.7}$$

To estimate the right-hand side, note that $\rho_\gamma \in W^{1,p}(0, T; L^p(\Omega)) \hookrightarrow C^{0,1/2}([0, T]; L^p(\Omega))$, and thus there exists a constant L_ρ, depending only on C_ϕ but not on γ, s.t. for all $t \geq 0$ and $h > 0$ satisfying $t + h \leq T$, the inequality

$$\|\rho_\gamma(t + h, \cdot) - \rho_\gamma(t, \cdot)\|_{L^p(\Omega)} \leq L_\rho \sqrt{h}$$

holds. The definition of q_t in (3.4) then directly implies

$$\|q_{t+h}(\cdot) - q_t(\cdot)\|_{L^p(\Omega)} \le CL_\rho \sqrt{h},$$

which, when inserted into the right-hand side of (A.7), yields

$$\int_0^{T-h} \|\psi_\gamma(t+h,\cdot) - \psi_\gamma(t,\cdot)\|_{W^{1,p}(\Omega)}^p \, dt \le (CL_\rho)^p \, (T-h) \, h^{p/2}.$$

Since $2 \le p$ holds, this is the desired estimate. Thus, Lemma on p. 1011 in [48] ensures the existence of a subsequence that strongly converges in $L^p(0, T; W^{1,p}(\Omega))$. By inverting the transformation (3.5), the result for ϕ_γ is obtained. □

References

1. Albi, G., Bongini, M., Cristiani, E., Kalise, D.: Invisible control of self-organizing agents leaving unknown environments. SIAM J. Appl. Math. **76**(4), 1683–1710 (2016). https://doi.org/10.1137/15m1017016
2. Albi, G., Cristiani, E., Pareschi, L., Peri, D.: Mathematical models and methods for crowd dynamics control. In: Crowd Dynamics, vol. 2, pp. 159–197. Springer, Berlin (2020). https://doi.org/10.1007/978-3-030-50450-2_8
3. Amadori, D., Di Francesco, M.: The one-dimensional Hughes model for pedestrian flow: Riemann-type solutions. Acta Math. Sci. **32**(1), 259–280 (2012). https://doi.org/10.1016/s0252-9602(12)60016-2
4. Amadori, D., Goatin, P., Rosini, M.D.: Existence results for Hughes' model for pedestrian flows. J. Math. Anal. Appl. **420**(1), 387–406 (2014). https://doi.org/10.1016/j.jmaa.2014.05.072
5. Bellomo, N., Dogbe, C.: On the modeling of traffic and crowds: a survey of models, speculations, and perspectives. SIAM Rev. **53**(3), 409–463 (2011). https://doi.org/10.1137/090746677
6. Bonifacius, L., Pieper, K., Vexler, B.: A priori error estimates for space-time finite element discretization of parabolic time-optimal control problems. SIAM J. Control Optim. **57**(1), 129–162 (2019). https://doi.org/10.1137/18m1166948
7. Borsche, R., Colombo, R.M., Garavello, M., Meurer, A.: Differential equations modeling crowd interactions. J. Nonlinear Sci. **25**(4), 827–859 (2015). https://doi.org/10.1007/s00332-015-9242-0
8. Borsche, R., Klar, A., Kühn, S., Meurer, A.: Coupling traffic flow networks to pedestrian motion. Math. Models Methods Appl. Sci. **24**(2), 359–380 (2014). https://doi.org/10.1142/S0218202513400113
9. Borsche, R., Meurer, A.: Microscopic and macroscopic models for coupled car traffic and pedestrian flow. J. Comput. Appl. Math. **348**, 356–382 (2019). https://doi.org/10.1016/j.cam.2018.08.037
10. Boyer, F., Fabrie, P.: Mathematical tools for the study of the incompressible Navier-Stokes equations and related models. Applied Mathematical Sciences, vol. 183. Springer, New York (2013). https://doi.org/10.1007/978-1-4614-5975-0
11. Brezis, H.: Functional Analysis, Sobolev Spaces and Partial Differential Equations. Universitext, Springer, New York (2011). https://doi.org/10.1007/978-0-387-70914-7
12. Burger, M., Di Francesco, M., Markowich, P.A., Wolfram, M.T.: Mean field games with nonlinear mobilities in pedestrian dynamics. Discret. Contin. Dyn. Syst. B **19**(5), 1311–1333 (2014). https://doi.org/10.3934/dcdsb.2014.19.1311
13. Burger, M., Pinnau, R., Roth, A., Totzeck, C., Tse, O.: Controlling a self-organizing system of individuals guided by a few external agents—particle description and mean-field limit. arXiv:1610.01325 (2016)
14. Burstedde, C., Klauck, K., Schadschneider, A., Zittartz, J.: Simulation of pedestrian dynamics using a two-dimensional cellular automaton. Phys. A: Stat. Mech. Appl. **295**(3–4), 507–525 (2001). https://doi.org/10.1016/s0378-4371(01)00141-8
15. Carlini, E., Festa, A., Silva, F.J., Wolfram, M.T.: A semi-Lagrangian scheme for a modified version of the Hughes' model for pedestrian flow. Dyn. Games Appl. **7**(4), 683–705 (2016). https://doi.org/10.1007/s13235-016-0202-6

16. Carrillo, J.A., Martin, S., Wolfram, M.T.: An improved version of the Hughes model for pedestrian flow. Math. Models Methods Appl. Sci. **26**(04), 671–697 (2016). https://doi.org/10.1142/s0218202516500147

17. Chudej, K., Pesch, H.J., Wächter, M., Sachs, G., Bras, F.L.: Instationary heat-constrained trajectory optimization of a hypersonic space vehicle by ODE–PDE-constrained optimal control. In: Variational Analysis and Aerospace Engineering, pp. 127–144. Springer, New York (2009). https://doi.org/10.1007/978-0-387-95857-6_8

18. Colombo, R.M., Garavello, M., Lécureux-Mercier, M.: A class of nonlocal models for pedestrian traffic. Math. Models Methods Appl. Sci. **22**(04), 1150023 (2012). https://doi.org/10.1142/s0218202511500230

19. Colombo, R.M., Gokieli, M., Rosini, M.D.: Modeling crowd dynamics through hyperbolic-elliptic equations. In: Non-Linear Partial Differential Equations, Mathematical Physics, and Stochastic Analysis, EMS Series of Congress Reports, pp. 111–128. European Mathematical Society, Zürich (2018). https://doi.org/10.4171/186-1/6

20. Cristiani, E., Peri, D.: Handling obstacles in pedestrian simulations: models and optimization. Appl. Math. Model. **45**, 285–302 (2017). https://doi.org/10.1016/j.apm.2016.12.020

21. Cristiani, E., Priuli, F.S., Tosin, A.: Modeling rationality to control self-organization of crowds: an environmental approach. SIAM J. Appl. Math. **75**(2), 605–629 (2015). https://doi.org/10.1137/140962413

22. Denk, R., Hieber, M., Prüss, J.: Optimal L^p-L^q-estimates for parabolic boundary value problems with inhomogeneous data. Mathematische Zeitschrift **257**(1), 193–224 (2007). https://doi.org/10.1007/s00209-007-0120-9

23. Denk, R., Prüss, J., Zacher, R.: Maximal L_p-regularity of parabolic problems with boundary dynamics of relaxation type. J. Funct. Anal. **255**(11), 3149–3187 (2008). https://doi.org/10.1016/j.jfa.2008.07.012

24. Denk, R., Saal, J., Seiler, J.: Inhomogeneous symbols, the Newton polygon, and maximal L^p-regularity. Russ. J. Math. Phys. **15**(2), 171–191 (2008). https://doi.org/10.1134/s1061920808020040

25. Di Francesco, M., Fagioli, S., Rosini, M.D., Russo, G.: Deterministic particle approximation of the Hughes model in one space dimension. Kinetic Relat. Models **10**(1), 215–237 (2017). https://doi.org/10.3934/krm.2017009

26. Di Francesco, M., Markowich, P.A., Pietschmann, J.F., Wolfram, M.T.: On the Hughes' model for pedestrian flow: the one-dimensional case. J. Differ. Equ. **250**(3), 1334–1362 (2011). https://doi.org/10.1016/j.jde.2010.10.015

27. Egger, H., Pietschmann, J.F., Schlottbom, M.: Identification of chemotaxis models with volume-filling. SIAM J. Appl. Math. **75**(2), 275–288 (2015). https://doi.org/10.1137/140967222

28. El-Khatib, N., Goatin, P., Rosini, M.D.: On entropy weak solutions of Hughes' model for pedestrian motion. Zeitschrift für Angewandte Mathematik und Physik. ZAMP. J. Appl. Math. Phys. Journal de Mathématiques et de Physique Appliquées **64**(2), 223–251 (2013). https://doi.org/10.1007/s00033-012-0232-x

29. Evans, L.C.: Partial Differential Equations. Graduate Studies in Mathematics, vol. 19. American Mathematical Society, Providence (1998). https://doi.org/10.1090/gsm/019

30. Fornasier, M., Piccoli, B., Rossi, F.: Mean-field sparse optimal control. Philos. Trans. R. Soc. A: Math. Phys. Eng. Sci. **372**(2028), 20130400 (2014). https://doi.org/10.1098/rsta.2013.0400

31. Fornasier, M., Solombrino, F.: Mean-field optimal control. ESAIM: Control Optim. Calc. Var. **20**(4), 1123–1152 (2014). https://doi.org/10.1051/cocv/2014009

32. Gilbarg, D., Trudinger, N.S.: Elliptic Partial Differential Equations of Second Order, Classics in Mathematics, vol. 224. Springer, Berlin (2001). https://doi.org/10.1007/978-3-642-61798-0. Reprint of the 1998 edition

33. Grisvard, P.: Elliptic problems in nonsmooth domains. Soc. Ind. Appl. Math. (2011). https://doi.org/10.1137/1.9781611972030

34. Hale, J.K.: Ordinary Differential Equations. Dover Publications, Mineola (2009)

35. Helbing, D., Molnár, P.: Social force model for pedestrian dynamics. Phys. Rev. E **51**(5), 4282–4286 (1995). https://doi.org/10.1103/physreve.51.4282

36. Holtmannspötter, M., Rösch, A., Vexler, B.: A priori error estimates for the space-time finite element discretization of an optimal control problem governed by a coupled linear PDE-ODE system. arXiv:2004.04448 (2020)

37. Hughes, R.L.: A continuum theory for the flow of pedestrians. Transp. Res. Part B: Methodol. **36**(6), 507–535 (2002). https://doi.org/10.1016/s0191-2615(01)00015-7

38. Kimmerle, S.J., Gerdts, M.: Necessary optimality conditions and a semi-smooth Newton approach for an optimal control problem of a coupled system of Saint-Venant equations and ordinary differential equations. Pure Appl. Funct. Anal. **1**(2), 231–256 (2016)

39. Kimmerle, S.J., Gerdts, M., Herzog, R.: Optimal control of an elastic crane-trolley-load system—a case study for optimal control of coupled ODE-PDE systems. Math. Comput. Model. Dyn. Syst. **24**(2), 182–206 (2018). https://doi.org/10.1080/13873954.2017.1405046

40. Knowles, G.: Finite element approximation of parabolic time optimal control problems. SIAM J. Control Optim. **20**(3), 414–427 (1982). https://doi.org/10.1137/0320032

41. Kunisch, K., Neic, A., Plank, G., Trautmann, P.: Inverse localization of earliest cardiac activation sites from activation maps based on the viscous Eikonal equation. J. Math. Biol. **79**(6–7), 2033–2068 (2019). https://doi.org/10.1007/s00285-019-01419-3

42. Ladyzhenskaya, O.A., Ural'tseva, N.N.: Linear and Quasilinear Elliptic Equations, Mathematics in Science and Engineering, vol. 46. Academic Press, New York (1968)

43. Le, D., Smith, H.: Strong positivity of solutions to parabolic and elliptic equations on non-smooth domains. J. Math. Anal. Appl. **275**(1), 208–221 (2002). https://doi.org/10.1016/s0022-247x(02)00314-1

44. Painter, K.J., Hillen, T.: Volume-filling and quorum-sensing in models for chemosensitive movement. Can. Appl. Math. Q. **10**(4), 501–543 (2002)

45. Pietschmann, J.F., Stötzner, A., Winkler, M.: Numerical investigation of agent controlled pedestrian dynamics using a structure preserving finite volume scheme. arXiv:2301.02516 (2023)

46. Pinnau, R., Totzeck, C.: Interacting particles and optimization. PAMM (2018). https://doi.org/10.1002/pamm.201800182

47. Schittkowski, K.: Numerical solution of a time-optimal parabolic boundary value control problem. J. Optim. Theory Appl. **27**(2), 271–290 (1979). https://doi.org/10.1007/BF00933231

48. Simon, J.: Écoulement d'un fluide non homogène avec une densité initiale s'annulant. Comptes Rendus de l'Académie des Sciences Paris. Séries A et B **287**(15), A1009–A1012 (1978)

49. Troianiello, G.M.: Elliptic Differential Equations and Obstacle Problems. Plenum Press, New York (1987). https://doi.org/10.1007/978-1-4899-3614-1

50. Tröltzsch, F.: Optimal Control of Partial Differential Equations. Graduate Studies in Mathematics, vol. 112. American Mathematical Society, Providence (2010). https://doi.org/10.1090/gsm/112

51. Wendl, S., Pesch, H.J., Rund, A.: On a state-constrained PDE optimal control problem arising from ODE-PDE optimal control. In: Recent Advances in Optimization and its Applications in Engineering, pp. 429–438. Springer, Berlin (2010). https://doi.org/10.1007/978-3-642-12598-0_37

52. Zheng, G., Yin, J.: Numerical approximation for a time optimal control problems governed by semi-linear heat equations. Adv. Differ. Equ. (2014). https://doi.org/10.1186/1687-1847-2014-94

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.