ESAIM: M2AN 59 (2025) 3301–3340
https://doi.org/10.1051/m2an/2025092
ESAIM: Mathematical Modelling and Numerical Analysis
www.esaim-m2an.org

CONVERGENCE OF A FINITE VOLUME SCHEME FOR A MODEL FOR ANTS

Maria Bruna[1], Markus Schmidtchen[2] and Oscar de Wit[3,*]

Abstract. We develop and analyse a finite volume scheme for a nonlocal active matter system known to exhibit a rich array of complex behaviours. The model under investigation was derived from a stochastic system of interacting particles describing a foraging ant colony coupled to pheromone dynamics. In this work, we prove that the unique numerical solution converges to the unique weak solution as the mesh size and the time step go to zero. We also show discrete long-time estimates, which prove that certain norms are preserved for all times, uniformly in the mesh size and time step. In particular, we prove higher regularity estimates which provide an analogue of continuum parabolic higher regularity estimates. Finally, we numerically study the rate of convergence of the scheme.

Mathematics Subject Classification. 35K55, 35B36, 65M08, 65M12, 35Q92, 92D50.

Received March 31, 2025. Accepted October 31, 2025.

1. Model introduction and motivation

Active matter exhibits richly complex collective motion, ranging from flocking birds and schooling fish [25] to pattern-forming skin cells [28], turbulence in suspensions of microswimmers [2], and lane-forming pedestrians [3]. There are two main approaches to modelling collective phenomena in active matter: microscopic or agent-based models or macroscopic or PDE-based models [5–7]. The former approach is very high-dimensional due to the large number of individuals typically involved in active matter systems, and quickly becomes analytically and computationally intractable. This is why macroscopic PDE models are customarily favoured to investigate the system's behaviour. Moreover, analytical challenges–commonly rooted in metastability and multiple bifurcation branches–highlight the importance of developing reliable numerical methods for macroscopic active matter models. In this paper, we develop and analyse a finite volume scheme for a nonlinear Fokker–Planck equation derived in [13] as the continuum limit of a microscopic active matter model for ants:

$$\partial_t f = \nabla_{\mathbf{x}} \cdot (D_T \nabla_{\mathbf{x}} f - \mathrm{Pe}\,\mathbf{e}(\theta) f) + \partial_\theta(\partial_\theta f - \gamma B[c] f), \quad f(t=0,\mathbf{x},\theta) = f^0(\mathbf{x},\theta), \tag{1.1a}$$

where $f = f(t,\mathbf{x},\theta) : \mathbb{R}_+ \times \mathbb{T}_1^2 \times \mathbb{T}_{2\pi} \to \mathbb{R}_+$ is a one-particle probability density function that quantifies the probability to find a particle at time t with position $\mathbf{x} \in \mathbb{T}_1^2$ and orientation $\theta \in \mathbb{T}_{2\pi}$, and $B[c]$ is some functional that represents a nonlocal interaction between particles through a chemical field $c = c(t,\mathbf{x})$. In equation (1.1a),

Keywords and phrases. Active matter, emergent behaviour, finite volume methods, numerical analysis.

[1] Mathematical Institute, University of Oxford, Oxford, UK.
[2] Institute of Scientific Computing, Faculty of Mathematics, TU Dresden, Germany.
[3] DAMTP, Centre for Mathematical Sciences, University of Cambridge, Cambridge, UK.
*Corresponding author: odewit8@gmail.com

$D_T > 0$ is the translational diffusion coefficient, Pe > 0 is the Péclet number (or dimensionless self-propulsion speed), $\mathbf{e}(\theta) = (\cos\theta, \sin\theta)^\mathsf{T}$ is the direction of self-propulsion, and $\gamma > 0$ is the interaction strength.

Equation (1.1a) was derived in [13] as the formal mean-field limit of a colony of ants interacting through the laying and sensing of pheromones. In particular, the pheromone field is coupled to the particle distribution through the elliptic equation

$$\Delta c - \alpha c + \rho = 0, \tag{1.1b}$$

where α is a chemical decay rate and particles' spatial probability density is given by

$$\rho(t, \mathbf{x}) = \int_{\mathbb{T}_{2\pi}} f(t, \mathbf{x}, \theta)\, \mathrm{d}\theta. \tag{1.1c}$$

Ants interact nonlocally *via* the pheromone field according to

$$B_\lambda[c] = \mathbf{n}(\theta) \cdot \nabla_\mathbf{x} c(\mathbf{x} + \lambda\mathbf{e}(\theta)), \tag{1.1d}$$

where $\mathbf{n}(\theta) = (-\sin\theta, \cos\theta)^\mathsf{T}$. Equation (1.1d) models how ants adjust their orientation towards higher chemical concentrations by sensing at position $\mathbf{x} + \lambda\mathbf{e}(\theta)$, where $\lambda \geq 0$ represents the sensing distance of ants. This can be thought of as the location of antennas relative to their body centre.

The solutions of model (1.1) were studied *via* a linear stability analysis and time-dependent numerical simulations in [13]. Specifically, for $\lambda > 0$, equation (1.1) admits three types of steady states depending on the model parameters: the trivial (constant) steady state, cluster stationary states, and lane-like steady states. In particular, in the region in the (γ, Pe)-plane where the constant state is linearly unstable, numerical solutions of equation (1.1) were found to either converge to stationary clusters or bidirectional lanes, which a thin region where bistability between the two types of nontrivial steady states is possible.

Similar lane-like steady states were also found in [8], where they used a related interaction term $B[c]$ (which we discuss further below).

Model (1.1) is a nonlinear, nonlocal extension of the Fokker–Planck equation associated with a single active Brownian particle corresponding to setting $\gamma = 0$ in Eq. (1.1a). The self-propulsion term, premultiplied by the Péclet number Pe, disrupts the gradient flow structure $\partial_t f = \nabla \cdot (f \nabla \delta\mathcal{F}/\delta f)$, often encountered in macroscopic PDE models for interacting particles. The non-gradient character of equation (1.1a) complicates the analysis as tools from gradient flows, such as a general LaSalle invariance principle [17], are not available. Despite these challenges, the well-posedness of solutions $f \in L^2(0, T; H^1)$ for any time $T > 0$ was proven in [13]. The estimates rely on $D_T > 0$ and classical Sobolev–Poincaré and Gagliardo–Nirenberg inequalities. Particularly, regularity estimates for the spatial density ρ, which satisfies

$$\partial_t \rho = \nabla \cdot (D_T \nabla\rho - \mathrm{Pe}\,\mathbf{p}), \tag{1.2}$$

where $\mathbf{p}(t, \mathbf{x}) = \int_0^{2\pi} \mathbf{e}(\theta) f(t, \mathbf{x}, \theta)\, \mathrm{d}\theta$ is the polarization (or average orientation), can be used to show regularity estimates for the full density f. The analysis in [13] also gives long-time estimates showing that there cannot be blow up at infinite time, such that $f \in L^\infty(0, +\infty; L^\infty)$. A remaining challenge is the characterisation of the pattern-forming steady states beyond a stability analysis of the homogeneous state.

Model (1.1) is a compelling example of pattern formation in a nonlinear active matter PDE model. Mathematical tools to characterise the existence of pattern formation were established for collectives of pedestrians in [12, 14, 15], flocking birds [11, 21], collectives of migrating cells [16] and suspensions of micro-swimmers [1, 27]. Despite the absence of a gradient flow structure, these papers provide the mathematical foundations for these pattern formations through bifurcation analysis, numerical methods, well-posedness analysis, and (non-)linear stability analysis.

Our primary objective is to develop a numerical scheme that accurately reproduces the behaviour of (1.1). Finite-volume schemes are particularly well-suited for this purpose, as their structure inherently ensures mass conservation and preserves the nonnegativity of the initial data. Extensive research exists on the numerical analysis of finite-volume methods [4, 18, 23, 24, 32], providing rigorous foundations for their convergence and error analysis across various PDE models. In particular, the works [4, 24] are central to our analysis. Reference

[24] demonstrates the convergence of a finite-volume scheme for the parabolic-elliptic Keller–Segel equation, which incorporates (1.1b), using a compactness argument in $L^2(0,T;H^1)$. A similar convergence result is given in [4] for a more general one-dimensional nonlinear aggregation-diffusion equation. Their analysis of general drift terms provides inspiration for treating the drift terms in our model, following *a priori* estimates for the chemical field in the discrete setting. These estimates can be obtained by translating the local well-posedness estimates in $L^2(0,T;H^1)$ of [13] to the discrete setting, using discrete analogues of Sobolev–Poincaré inequalities [9]. In addition to a convergence result, we show the analysis for long-time estimates for the numerical solutions. As part of this, we show that a standard H^1-regularity parabolic estimate has a finite-volume equivalent, allowing us to show L^2- and L^∞-long-time estimates. A crucial component of this analysis is a recent result on the Morrey inequality within the finite volume framework ([30], Thm. 4.1).

The paper is organised as follows. In Section 2, we define the weak solutions of (1.1), introduce the numerical scheme and state the main convergence result. In Section 3, we provide the required *a priori* estimates for the convergence result, which is proven in Section 4. Section 5 covers the long-time estimates for the scheme. Finally, we present a series of numerical results in Section 6 showcasing the performance of our scheme. We show the scheme's order of accuracy and numerically verify the long-time behaviour observed in [8, 13], confirming the convergence to nontrivial steady states. We conclude by showing that quasi-steady states comprising multiple spots or lanes exist and that they eventually destabilise and evolve to a single spot or lane.

2. Definition of the scheme and function spaces

We begin by introducing some notation. Throughout, let $\xi = (\mathbf{x},\theta)$ denote the three-dimensional coordinate vector, $\Omega := \mathbb{T}_1^2$ be a two-dimensional domain with periodic boundaries extending from $[-\frac{1}{2},\frac{1}{2}]$, $\Sigma = \Omega\times\mathbb{T}_{2\pi}$ and $\Sigma_T = \Sigma\times(0,T)$. We will consider the nonlocal interaction (1.1d), as well as its zeroth and first-order Taylor expansions in λ, which we denote by B_0 and B_τ respectively:

$$B_0[c] = \mathbf{n}(\theta)\cdot\nabla_{\mathbf{x}}c(\mathbf{x}), \tag{2.1a}$$
$$B_\tau[c] = \mathbf{n}(\theta)\cdot\nabla_{\mathbf{x}}c(\mathbf{x}) + \tau\mathbf{n}(\theta)\cdot\nabla_{\mathbf{x}}^2 c(\mathbf{x})\mathbf{e}(\theta). \tag{2.1b}$$

The sensing strategy B_0 corresponds to ants "without antennas" sensing at their centre of mass. The B_τ term can be derived as the limit of a sensing mechanism for two antennas spread apart by some angle [8,29]. Typically, the terms B_λ and B_τ lead to either aggregation or lane formation, whereas the term B_0 only leads to aggregation spots. Let us recall our notion of weak solution of equation (1.1).

Definition 2.1 (Weak solution)**.** A function $f\in L^2(0,T;H^1(\Sigma))$ is a weak solution for equation (1.1) if $f(0) = f^0 \in L^1_+(\Sigma)\cap L^\infty_+(\Sigma)$ with mass one and

$$\int_0^T\int_\Sigma [f\partial_t\varphi - (D_T\nabla_{\mathbf{x}}f - \mathrm{Pe}\mathbf{e}(\theta)f)\cdot\nabla_{\mathbf{x}}\varphi - (\partial_\theta f - \gamma B[c]f)\partial_\theta\varphi]\,\mathrm{d}\xi\,\mathrm{d}t + \int_\Sigma f(0)\varphi(0)\,\mathrm{d}\xi = 0, \tag{2.2}$$

for all $\varphi\in C^\infty(\overline{\Sigma_T})$ such that $\varphi(T)=0$ and $c\in L^2(0,T;H^2(\Omega))$ is the unique strong solution of equation (1.1b).

Next, we introduce a discretisation of the domain Σ and a time interval $(0,T)$.

Definition 2.2 (Discretisation of the domain)**.** For some $N_x, N_y, N_\theta\in\mathbb{N}$, we introduce a family of cells

$$C_{i,j,k} = [x_{i-1/2}, x_{i+1/2})\times[y_{j-1/2}, y_{j+1/2})\times[\theta_{k-1/2},\theta_{k+1/2}), \tag{2.3}$$

for

$$(i,j,k)\in\mathcal{I} = \{(i,j,k)\,|\,1\le i\le N_x, 1\le j\le N_y, 1\le k\le N_\theta\}, \tag{2.4}$$

where

$$x_{i-1/2} = -\frac{1}{2} + (i-1)\Delta x, \quad y_{j-1/2} = -\frac{1}{2} + (j-1)\Delta y, \quad \text{and} \quad \theta_{k-1/2} = (k-1)\Delta\theta,$$

with $\Delta x = 1/N_x, \Delta y = 1/N_y$ and $\Delta\theta = 2\pi/N_\theta$. We note that $\bigcup_{\mathcal{I}} \overline{C_{i,j,k}} = \Sigma$. We also define the dual mesh cells as

$$C_{i+1/2,j,k} = [x_i, x_{i+1}) \times [y_{j-1/2}, y_{j+1/2}) \times [\theta_{k-1/2}, \theta_{k+1/2}), \tag{2.5}$$

with $x_i = -\frac{1}{2} + (i - \frac{1}{2})\Delta x$ for $1 \leq i \leq N_x$ (and $x_{N_x+1} \equiv x_1$), and analogously for the y- and θ-direction. We use the notation $\mathbf{x}_{i,j} = (x_i, y_j)$.

For the time interval, we introduce the subintervals $[t^n, t^{n+1})$, where $t^n = n\Delta t$, such that

$$\bigcup_{n=0,1,\ldots,N_T-1} [t^n, t^{n+1}) = [0, T), \tag{2.6}$$

and $N_T = T/\Delta t$ is a positive integer.

Throughout, we call a family of meshes admissibly regular if there exist some constants $\varepsilon_1, \varepsilon_2$ satisfying $0 < \varepsilon_1 < \varepsilon_2 < \infty$ such that, for all meshes

$$\frac{\Delta\theta}{\Delta x}, \frac{\Delta\theta}{\Delta y}, \frac{\Delta x}{\Delta y} \in (\varepsilon_1, \varepsilon_2). \tag{2.7}$$

We are now ready to introduce the numerical scheme using the above notation.

Definition 2.3 (Definition of the scheme)**.** We construct a discretised version of the initial data f^0 *via* the cell averages

$$f^0_{i,j,k} = \frac{1}{\Delta\xi} \int_{C_{i,j,k}} f^0(\xi)\,\mathrm{d}\xi, \tag{2.8}$$

for $(i,j,k) \in \mathcal{I}$, where $\Delta\xi = \Delta\mathbf{x}\Delta\theta$, and $\Delta\mathbf{x} = \Delta x\Delta y$. To discretise equation (1.1a), we rewrite it in divergence form $\partial_t f = -\nabla_\xi \cdot \mathbf{F}$, with flux $\mathbf{F} = (F^x, F^y, F^\theta)$

$$\begin{aligned} F^x(t,\xi) &= -(D_T\partial_x f - \mathrm{Pe}\cos(\theta) f), \\ F^y(t,\xi) &= -(D_T\partial_y f - \mathrm{Pe}\sin(\theta) f), \\ F^\theta(t,\xi) &= -(\partial_\theta f - \gamma B[c] f). \end{aligned} \tag{2.9}$$

Integrating equation (1.1a) over a test cell, $[t^n, t^{n+1}) \times C_{i,j,k}$, yields

$$\begin{aligned} \int_{C_{i,j,k}} f\big(t^{n+1},\xi\big)\,\mathrm{d}\xi - \int_{C_{i,j,k}} f(t^n,\xi)\,\mathrm{d}\xi = &- \int_{t^n}^{t^{n+1}} \int_{C_{i,j,k}} \partial_x F^x(t,\xi)\,\mathrm{d}\xi\,\mathrm{d}t - \int_{t^n}^{t^{n+1}} \int_{C_{i,j,k}} \partial_y F^y(t,\xi)\,\mathrm{d}\xi\,\mathrm{d}t \\ &- \int_{t^n}^{t^{n+1}} \int_{C_{i,j,k}} \partial_\theta F^\theta(t,\xi)\,\mathrm{d}\xi\,\mathrm{d}t. \end{aligned} \tag{2.10}$$

Using the fundamental theorem turning the derivatives on the right-hand side into finite differences, we can approximate equation (2.10) by the backward finite volume scheme

$$\begin{aligned} \frac{f^{n+1}_{i,j,k} - f^n_{i,j,k}}{\Delta t} = &- \frac{1}{\Delta x}\Big(F^{x,n+1}_{i+1/2,j,k} - F^{x,n+1}_{i-1/2,j,k}\Big) - \frac{1}{\Delta y}\Big(F^{y,n+1}_{i,j+1/2,k} - F^{y,n+1}_{i,j-1/2,k}\Big) \\ &- \frac{1}{\Delta\theta}\Big(F^{\theta,n+1}_{i,j,k+1/2} - F^{\theta,n+1}_{i,j,k-1/2}\Big), \end{aligned} \tag{2.11a}$$

where the discrete fluxes are defined as

$$\begin{aligned} F^{x,n}_{i+1/2,j,k} &= -\Big(D_T \mathrm{d}_x f^n_{i+1/2,j,k} - \mathrm{Pe} U^n_{i+1/2,j,k}\Big), \\ F^{y,n}_{i,j+1/2,k} &= -\Big(D_T \mathrm{d}_y f^n_{i,j+1/2,k} - \mathrm{Pe} V^n_{i,j+1/2,k}\Big), \\ F^{\theta,n}_{i,j,k+1/2} &= -\Big(\mathrm{d}_\theta f^n_{i,j,k+1/2} - \gamma W^n_{i,j,k+1/2}\Big), \end{aligned} \tag{2.11b}$$

and discrete partial derivatives defined as

$$\begin{aligned} \mathrm{d}_x f_{i+1/2,j,k} &= \frac{1}{\Delta x}(f_{i+1,j,k} - f_{i,j,k}), \\ \mathrm{d}_y f_{i,j+1/2,k} &= \frac{1}{\Delta y}(f_{i,j+1,k} - f_{i,j,k}), \\ \mathrm{d}_\theta f_{i,j,k+1/2} &= \frac{1}{\Delta \theta}(f_{i,j,k+1} - f_{i,j,k}). \end{aligned} \tag{2.11c}$$

The upwind velocities for the drift terms are, using the notation $(\cdot)^+ = \max(\cdot, 0), (\cdot)^- = \min(\cdot, 0)$,

$$\begin{aligned} U^n_{i+1/2,j,k} &= (\cos\theta_k)^+ f^n_{i,j,k} + (\cos\theta_k)^- f^n_{i+1,j,k}, \\ V^n_{i,j+1/2,k} &= (\sin\theta_k)^+ f^n_{i,j,k} + (\sin\theta_k)^- f^n_{i,j+1,k}, \\ W^n_{i,j,k+1/2} &= \Big(B[c^n]_{i,j,k+1/2}\Big)^+ f^n_{i,j,k} + \Big(B[c^n]_{i,j,k+1/2}\Big)^- f^n_{i,j,k+1}. \end{aligned} \tag{2.11d}$$

Next, we discretise equation (1.1b) using finite differences

$$0 = \big(\mathrm{d}_x^2 + \mathrm{d}_y^2\big) c^n_{i,j} - \alpha c^n_{i,j} + \rho^n_{i,j}, \tag{2.11e}$$

where the Laplacian terms are defined as

$$\mathrm{d}_x^2 c^n_{i,j} = \frac{c^n_{i+1,j} - 2c^n_{i,j} + c^n_{i-1,j}}{\Delta x^2}, \tag{2.11f}$$

and analogously in the y-direction. Here, $\rho^n_{i,j} = \sum_k f^n_{i,j,k} \Delta\theta$ denotes the discretised spatial density.

Lastly, we use the following discretisations for the three different modelling choices for the interaction term $B[c]$.

– For $B_0[c]$, we define

$$B_0[c]_{i,j,k+1/2} = \mathbf{n}(\theta_{k+1/2}) \cdot D c_{i,j}, \tag{2.11g}$$

and the gradient is discretised as a centred finite difference

$$D c_{i,j} = \frac{1}{2}\begin{pmatrix} \dfrac{1}{\Delta x}(c_{i+1,j} - c_{i-1,j}) \\ \dfrac{1}{\Delta y}(c_{i,j+1} - c_{i,j-1}) \end{pmatrix} = \begin{pmatrix} D_x c_{i,j} \\ D_y c_{i,j} \end{pmatrix} = \frac{1}{2}\begin{pmatrix} \mathrm{d}_x c_{i+1/2,j} + \mathrm{d}_x c_{i-1/2,j} \\ \mathrm{d}_y c_{i,j+1/2} + \mathrm{d}_y c_{i,j-1/2} \end{pmatrix}. \tag{2.11h}$$

– For $B_\lambda[c]$, we define

$$B_\lambda[c]_{i,j,k+1/2} = \mathbf{n}(\theta_{k+1/2}) \cdot D c_{i,j,k+1/2}, \tag{2.11i}$$

where $c_{i,j,k+1/2}$ approximates $c(\mathbf{x}_{i,j} + \lambda \mathbf{e}(\theta_{k+1/2}))$ and it is defined using nearest-neighbour interpolation to access c at the points $\mathbf{x}_{i,j} + \lambda \mathbf{e}(\theta)$, see Definition B.1 for a full definition.

– For $B_\tau[c]$, we define

$$B_\tau[c]_{i,j,k+1/2} = \mathbf{n}(\theta_{k+1/2}) \cdot Dc_{i,j} + \tau \mathbf{n}(\theta_{k+1/2}) \cdot Hc_{i,j} \mathbf{e}(\theta_{k+1/2}). \tag{2.11j}$$

where the discrete Hessian is defined as

$$Hc_{i,j} = \begin{pmatrix} \mathrm{d}_x^2 c_{i,j} & D_x D_y c_{i,j} \\ D_x D_y c_{i,j} & \mathrm{d}_y^2 c_{i,j} \end{pmatrix}, \tag{2.11k}$$

where the off-diagonal terms read

$$D_x D_y c_{i,j} = \frac{1}{4\Delta\mathbf{x}}(c_{i+1,j+1} - c_{i-1,j+1} - c_{i+1,j-1} + c_{i-1,j-1}).$$

Definition 2.4 (Piecewise constant interpolations)**.** Given $(f^n_{i,j,k})_{(i,j,k,n)}$ a discrete solution to the scheme, we define the piecewise constant interpolation as

$$f_h(t,x,y,\theta) = f^n_{i,j,k}, \quad \text{for} \quad (t,x,y,\theta) \in [t^{n-1}, t^n) \times C_{i,j,k}, \tag{2.12}$$

where the subscript h corresponds to the mesh size

$$h = \max\{\Delta t, \Delta x, \Delta y, \Delta\theta\}. \tag{2.13}$$

We also use straightforward analogous definitions of piecewise constant interpolations for functions with only (t,x,y) as variables.

We define the discrete gradients on the dual mesh as

$$\mathrm{d}_x f_h(t,x,y,\theta) = \mathrm{d}_x f^n_{i+1/2,j,k} \quad \text{for} \quad (t,x,y,\theta) \in [t^{n-1}, t^n) \times C_{i+1/2,j,k}, \tag{2.14}$$

and analogously for the gradients in y and θ.

We show the following well-posedness result for finite-volume scheme of Definition 2.3. We postpone its proof to the next section.

Proposition 2.1 (Existence, uniqueness and nonnegativity)**.** *Let $f^0 \in L^1_+(\Sigma) \cap L^\infty_+(\Sigma)$ be a nonnegative initial datum with mass one. Then, for any $N_x, N_y, N_\theta, N_T \in \mathbb{N}$, provided $\Delta t < C_{\mathrm{CFL}}$, for some constant $C_{\mathrm{CFL}} > 0$, there is a unique nonnegative solution $(f^n_{i,j,k})$ to the scheme of equation (2.11) for $(n,i,j,k) \in \{0,1,\dots,N_T\} \times \mathcal{I}$. Moreover, the mass is conserved and there holds $\|f_h(t)\|_{L^1} = \|f^0\|_{L^1}$, for all $t \in [0,T)$.*

A priori, the constant C_{CFL} depends on the phase-space mesh, as shown in Appendix A. However, the condition for the well-posedness of the scheme can be made independent of the mesh using the L^∞-regularity estimates of Section 5. This observation is similar to the one made in Remark 2.1 of [24].

We conclude this section by stating the main convergence result, which we will prove later.

Theorem 2.1 (Convergence of the scheme)**.** *Let $f^0 \in L^1_+(\Sigma) \cap L^\infty_+(\Sigma)$ be a nonnegative initial datum with mass one. Then, given an admissibly regular family of meshes as in Definition 2.2, a family of solutions $(f_h)_h$ as in Definition 2.4, and provided $\Delta t < \max\{\frac{D_T}{2\mathrm{Pe}^2}, 1/(C' C_T)\}$ (see Prop. 3.3) the sequence of numerical solutions converges strongly in $L^2(\Sigma_T)$, as $h \to 0$, and its limit is the unique weak solution of equation (1.1) as in Definition 2.1.*

3. *A priori* ESTIMATES

In this section, we prove the existence result of Proposition 2.1 and derive a set of discrete *a priori* estimates necessary for Theorem 2.1. To this end, we recall a discrete Grönwall-type inequality mimicking the standard continuum Grönwall inequality. Next, we define discrete analogues of Sobolev space norms. We then prove an L^2-estimate for the discrete spatial density ρ_h, an H^1-type estimate for the discrete chemical field c_h and the discrete interaction term $\mathrm{d}_\theta B[c_h]$. Finally, we obtain a uniform-in-time L^2-estimate for f_h and $\mathrm{d}_\xi f_h$ on bounded time intervals, independent of the mesh size. All these estimates will be used to prove the convergence of the scheme in Section 4.

Lemma 3.1 (Discrete Grönwall inequality)**.** *Suppose* $(F^n)_n \subset (0,\infty)$ *is a nonnegative real-valued sequence and* $C \in [0,1)$*. If*

$$F^{n+1} - F^n \le CF^{n+1}, \quad \text{for} \quad n = 0,1,2,\dots, \tag{3.1}$$

then

$$F^n \le F^0 \frac{1}{(1-C)^n}. \tag{3.2}$$

Proof. The proof follows by induction. Let $C \in [0,1)$. For $n = 0$, we clearly have

$$F^1 \le F^0 \frac{1}{1-C}.$$

We then assume that for $n = 0,\dots,n'-1$ the inequality $F^n \le F^0 \frac{1}{(1-C)^n}$ holds. Then, for n' we have

$$F^{n'} - F^{n'-1} \le CF^{n'},$$

so that by assumption

$$F^{n'} \le \frac{1}{1-C} F^{n'-1} \le F^0 \frac{1}{1-C} \frac{1}{(1-C)^{n'-1}} = F^0 \frac{1}{(1-C)^{n'}}.$$

This concludes the proof. □

Next, let us define the discrete Sobolev spaces we will use in our analysis.

Definition 3.1 (Function space norms)**.** For $1 \le p < +\infty$, we define the L^p-norm

$$\|f_h\|_{L^p} = \left(\int_\Sigma |f_h(\xi)|^p \mathrm{d}\xi \right)^{1/p} = \left(\sum_{i,j,k} |f_{i,j,k}|^p \Delta\xi \right)^{1/p}. \tag{3.3}$$

We define the L^∞-norm as

$$\|f_h\|_{L^\infty} = \sup_{\xi\in\Sigma} |f_h(\xi)| = \sup_{i,j,k} |f_{i,j,k}|. \tag{3.4}$$

For $1 \le p < +\infty$ we define the Sobolev seminorm

$$\begin{aligned} |f_h|_{1,p} &= \left(\sum_{i,j,k} \left(\left|\mathrm{d}_x f_{i+1/2,j,k}\right|^p + \left|\mathrm{d}_y f_{i,j+1/2,k}\right|^p + \left|\mathrm{d}_\theta f_{i,j,k+1/2}\right|^p \right) \Delta\xi \right)^{1/p} \\ &= \|\mathrm{d}_\xi f_h\|_{L^p}, \end{aligned} \tag{3.5}$$

where

$$\mathrm{d}_\xi f_h(x,y,\theta) = (\mathrm{d}_x f_h, \mathrm{d}_y f_h, \mathrm{d}_\theta f_h)^\mathsf{T}. \tag{3.6}$$

The discrete Sobolev norm is defined as

$$\|f_h\|_{1,p} = |f_h|_{1,p} + \|f_h\|_{L^p}. \tag{3.7}$$

We also use the notation $\|f_h\|_{1,p,\Sigma}$ if it is relevant what domain, here Σ, is used for the integration or summation.

Lemma 3.2 (Summation by parts)**.** *For $(a_{i,j,k})_{(i,j,k)\in\mathcal{I}}$ and $(b_{i,j,k})_{(i,j,k)\in\mathcal{I}}$ on the discretised periodic domain we have*

$$\sum_{i,j,k}(a_{i+1,j,k} - a_{i,j,k})b_{i,j,k} = -\sum_{i,j,k}(b_{i+1,j,k} - b_{i,j,k})a_{i+1,j,k}. \tag{3.8}$$

Analogous relations hold in the j- and k-directions.

We can now prove Proposition 2.1.

Proof of Proposition 2.1. We first establish the existence and uniqueness of the finite difference equation (2.11e). For a given nonnegative $(\rho^n_{i,j})_{i,j}$, we note that we have a linear system of N_xN_y unknowns $c^n_{i,j}$ and the same number of coupled equations

$$-\mathrm{d}_x^2 c^n_{i,j} - \mathrm{d}_y^2 c^n_{i,j} + \alpha c^n_{i,j} = \rho^n_{i,j}.$$

We next show that the kernel of this linear system is the trivial solution $c^n_{i,j} = 0$. Indeed, for $\rho^n_{i,j} = 0$, multiplying equation (2.11e) by $c^n_{i,j}$ and summing over all pairs (i,j), summation by parts yields

$$\sum_{i,j}\left(\left|\mathrm{d}_x c^n_{i+1/2,j}\right|^2 + \left|\mathrm{d}_y c^n_{i,j+1/2}\right|^2 + \alpha\left|c^n_{i,j}\right|^2\right) = 0. \tag{3.9}$$

Hence, we find $c^n_{i,j} = 0$. This implies that for any $\rho^n_{i,j}$, there is a unique solution to the linear system equation (2.11e). Further, since $\rho^n_{i,j}$ is nonnegative by assumption, $c^n_{i,j}$ is nonnegative by the same argument as in the proof of Theorem 2.1 from [24].

Now, we construct the fixed-point operator. First, for a given prescribed chemical field $c^n_{i,j}$, possibly varying in time, and an initial datum $f^0_{i,j,k}$ satisfying the assumptions, there exists a unique solution $f^n_{i,j,k}$ to the scheme (2.11) (without Eq. (2.11e)), see Theorem 17.1 of [23]. This solution is also nonnegative and mass-preserving by the same arguments as in the proof of Theorem 3.1 from [4]. Upon setting

$$\mathcal{X} = \left\{ \left(c^n_{i,j}\right) \in \mathbb{R}^{N_T N_x N_y} \;\middle|\; \sup_{n=0,1,\ldots,N_T} \sum_{i,j}\left|c^n_{i,j}\right|\Delta\mathbf{x} \le 1/\alpha \right\},$$

which is convex and compact, we define the operator

$$\mathcal{S} : \mathcal{X} \to \mathcal{X}, c^n_{i,j} \mapsto f^n_{i,j,k} \mapsto \tilde{c}^n_{i,j},$$

where $\tilde{c}^n_{i,j}$ solves equation (2.11e) with $\rho^n_{i,j} = \sum_k f^n_{i,j,k}\Delta\theta$ as its source term. This operator is continuous, provided $\frac{\Delta t}{\Delta \mathbf{x}^2}$ is small enough. For a proof, see Appendix A.

Also, we have that $\sum_{i,j}\tilde{c}^n_{i,j}\Delta\mathbf{x} = 1/\alpha$ for all $n = 0, 1, \ldots, N_T$, using the mass-preserving property and equation (2.11e). Hence, by the nonnegativity of $\tilde{c}^n_{i,j}$, the operator $\mathcal{S}$ is a fixed-point operator. Applying Brouwer's fixed point theorem, we obtain a solution $(f^n_{i,j,k}, c^n_{i,j})$ to the scheme as in Definition 2.3. Furthermore, the solution is unique by as similar approach as showing the uniqueness. This requires $\Delta t/\Delta\xi^2$ to be small enough, see again Appendix A. This concludes the proof. □

Remark 3.1 (Scheme for $\rho_{i,j}^n$)**.** Multiplying equation (2.11a) by $\Delta\theta$ and summing over k, results in the following scheme for the spatial density,

$$\frac{\rho_{i,j}^{n+1}-\rho_{i,j}^n}{\Delta t} = -\frac{1}{\Delta x}\Big(\bar{F}^{x,n+1}_{i+1/2,j}-\bar{F}^{x,n+1}_{i-1/2,j}\Big) - \frac{1}{\Delta y}\Big(\bar{F}^{y,n+1}_{i,j+1/2}-\bar{F}^{y,n+1}_{i,j-1/2}\Big), \tag{3.10a}$$

with

$$\begin{aligned} \bar{F}^{x,n+1}_{i+1/2,j} &= -\Big(D_T \mathrm{d}_x\rho^{n+1}_{i+1/2,j} - \mathrm{Pe}\bar{U}^{n+1}_{i+1/2,j}\Big), \quad \text{where} \quad \bar{U}^{n+1}_{i+1/2,j} = \sum_k U^{n+1}_{i+1/2,j,k}\Delta\theta,\\ \bar{F}^{y,n+1}_{i,j+1/2} &= -\Big(D_T \mathrm{d}_y\rho^{n+1}_{i,j+1/2} - \mathrm{Pe}\bar{V}^{n+1}_{i,j+1/2}\Big), \quad \text{where} \quad \bar{V}^{n+1}_{i,j+1/2} = \sum_k V^{n+1}_{i,j+1/2,k}\Delta\theta, \end{aligned} \tag{3.10b}$$

where $U^n_{i+1/2,j,k}, V^n_{i+1/2,j,k}$ are given in equation (2.11d). As expected, it corresponds to the finite-volume discretisation of equation (1.2). We define the piecewise constant interpolation $\rho_h(t,\mathbf{x})$ from $\rho^n_{i,j}$ analogously to Definition 2.4.

Proposition 3.1 (L^2-estimate for ρ_h)**.** *Provided that $\Delta t < \frac{D_T}{2\mathrm{Pe}^2}$, the meshes are an admissibly regular family as in Definition 2.2 and $\rho^0 \in L^1_+(\Omega)\cap L^\infty_+(\Omega)$, then the spatial density of the solution for the scheme of equation (2.11) satisfies*

$$\sup_{t\in[0,T]} \|\rho_h(t)\|_{L^2} \le C_T, \tag{3.11}$$

where $C_T>0$ is independent of $h>0$ but may depend on $T>0$ and $\|\rho^0\|_{L^2}$.

Proof. We multiply equation (3.10a) by $\rho^{n+1}_{i,j}$ and sum over all i,j to get

$$\begin{aligned} \sum_{i,j}\frac{\rho_{i,j}^{n+1}-\rho_{i,j}^n}{\Delta t}\rho^{n+1}_{i,j} &= D_T\sum_{i,j}(\mathrm{d}_x^2\rho^{n+1}_{i,j}+\mathrm{d}_y^2\rho^{n+1}_{i,j})\rho^{n+1}_{i,j}\\ &\quad - \mathrm{Pe}\sum_{i,j}\left(\frac{\bar{U}^{n+1}_{i+1/2,j}-\bar{U}^{n+1}_{i-1/2,j}}{\Delta x}\rho^{n+1}_{i,j} + \frac{\bar{V}^{n+1}_{i,j+1/2}-\bar{V}^{n+1}_{i,j-1/2}}{\Delta y}\rho^{n+1}_{i,j}\right). \end{aligned} \tag{3.12}$$

Then, multiplying by $\Delta\mathbf{x}$ and summing by parts as in Lemma 3.2, we get

$$\begin{aligned} \sum_{i,j}\frac{\rho_{i,j}^{n+1}-\rho_{i,j}^n}{\Delta t}\rho^{n+1}_{i,j}\Delta\mathbf{x} &= -D_T\sum_{i,j}\left(\left|\mathrm{d}_x\rho^{n+1}_{i+1/2,j}\right|^2+\left|\mathrm{d}_y\rho^{n+1}_{i,j+1/2}\right|^2\right)\Delta\mathbf{x}\\ &\quad + \mathrm{Pe}\sum_{i,j}\Big(\bar{U}^{n+1}_{i+1/2,j}\,\mathrm{d}_x\rho^{n+1}_{i+1/2,j}+\bar{V}^{n+1}_{i,j+1/2}\,\mathrm{d}_y\rho^{n+1}_{i,j+1/2}\Big)\Delta\mathbf{x}\\ &\le -(D_T-\varepsilon)\big\|\mathrm{d}_{\mathbf{x}}\rho_h(t^{n+1})\big\|^2_{L^2(\Omega)} + \frac{\mathrm{Pe}^2}{\varepsilon}\big\|\rho_h(t^{n+1})\big\|^2_{L^2(\Omega)}, \end{aligned} \tag{3.13}$$

using $|\bar{U}^{n+1}_{i+1/2,j}| \le \max\{\rho^{n+1}_{i,j},\rho^{n+1}_{i+1,j}\}$, the weighted Young's inequality, and the notation

$$\mathrm{d}_{\mathbf{x}}\rho_h = (\mathrm{d}_x\rho_h,\mathrm{d}_y\rho_h)^{\mathsf{T}}. \tag{3.14}$$

Finally, using the inequality $b(b-a)\ge\frac{1}{2}(b^2-a^2)$ on the left-hand side of equation (3.13), we obtain

$$\frac{1}{2\Delta t}\Big(\big\|\rho_h(t^{n+1})\big\|^2_{L^2(\Omega)} - \|\rho_h(t^n)\|^2_{L^2(\Omega)}\Big) \le -(D_T-\varepsilon)\big\|\mathrm{d}_{\mathbf{x}}\rho_h(t^{n+1})\big\|^2_{L^2(\Omega)} + \frac{\mathrm{Pe}^2}{\varepsilon}\big\|\rho_h(t^{n+1})\big\|^2_{L^2(\Omega)}. \tag{3.15}$$

Thus, after setting $\varepsilon = D_T$

$$\|\rho_h(t^{n+1})\|^2_{L^2(\Omega)} - \|\rho_h(t^n)\|^2_{L^2(\Omega)} \leq \frac{2\Delta t \mathrm{Pe}^2}{D_T}\|\rho_h(t^{n+1})\|^2_{L^2(\Omega)}. \tag{3.16}$$

Applying Grönwall's estimate (Lem. 3.1) we get, provided $\frac{2\Delta t\mathrm{Pe}^2}{D_T} < 1$ (strictly, uniformly over the sequence of meshes)

$$\|\rho_h(t^n)\|^2_{L^2(\Omega)} \leq \frac{1}{\left(1 - \frac{2\Delta t\mathrm{Pe}^2}{D_T}\right)^n}\|\rho_h^0\|^2_{L^2(\Omega)}. \tag{3.17}$$

We note that for $u = \frac{2\Delta t\mathrm{Pe}^2}{D_T}$, there is some $A \in (0,1)$ such that $u < A < 1$ uniformly in h. Hence, we estimate, using $(1+\frac{s}{n})^n \leq e^s$ for $s > 0$,

$$\begin{aligned}\frac{1}{(1-u)^n} &= \left(1 + \frac{u}{1-u}\right)^n \\ &\leq \left(1 + \frac{1}{A}u\right)^n \leq \left(1 + \frac{1}{A}\frac{2T\mathrm{Pe}^2}{D_T N_T}\right)^{N_T} \\ &\leq \exp\left(\frac{2T\mathrm{Pe}^2}{AD_T}\right).\end{aligned}$$

This concludes the proof. □

Remark 3.2 (Hessian of c_h)**.** For the chemical field c, we define the piecewise constant Hessian

$$\mathrm{d}^2_{\mathbf{x}} c_h = \begin{pmatrix} \mathrm{d}^2_x c_h & \mathrm{d}_{xy} c_h \\ \mathrm{d}_{xy} c_h & \mathrm{d}^2_y c_h \end{pmatrix}, \tag{3.18}$$

such that

$$\mathrm{d}^2_x c_h = \mathrm{d}^2_x c_{i,j} \quad \text{for } (x,y) \in C_{i,j}, \qquad \mathrm{d}_{xy} c_h = \mathrm{d}_y \mathrm{d}_x c_{i+1/2,j+1/2} \quad \text{for } (x,y) \in C_{i+1/2,j+1/2}. \tag{3.19}$$

Proposition 3.2. *There exist constants $C, C_B > 0$ such that, for any choice of $B \in \{B_0, B_\lambda, B_\tau\}$, provided the meshes are an admissibly regular family as in Definition 2.2, the following inequalities hold for ρ_h and c_h numerical solutions of equation* (2.11e)*:*

$$\|c_h\|_{1,2} \leq C\|\rho_h\|_{L^2(\Omega)}, \qquad \|\mathrm{d}^2_{\mathbf{x}} c_h\|_{L^2} \leq C\|\rho_h\|_{L^2(\Omega)} \tag{3.20}$$

and

$$\|\mathrm{d}_\theta B[c_h]\|_{L^2} \leq C_B\|\rho_h\|_{L^2(\Omega)}. \tag{3.21}$$

The constants $C, C_B > 0$ do not depend on h.

Proof. First, we derive pointwise estimates on $\mathrm{d}_\theta B[c_h]$ in terms of the discrete gradient and discrete Hessian of c_h. Then, we derive estimates for the L^2-norms of the discrete gradient and Hessian of c_h in terms of the L^2-norm of ρ_h.

We note that for all the modelling choices for the interaction terms for $B[c]$ we have

$$\mathrm{d}_\theta B[c]_{i,j,k} = \frac{1}{\Delta\theta}\big(B[c]_{i,j,k+1/2} - B[c]_{i,j,k-1/2}\big). \tag{3.22}$$

– B_0 interaction: using the mean-value theorem, we have

$$|\mathrm{d}_\theta B_0[c]_{i,j,k}| = \left| \frac{1}{\Delta\theta} \big[\mathbf{n}\big(\theta_{k+1/2}\big) - \mathbf{n}\big(\theta_{k-1/2}\big)\big] \cdot Dc_{i,j} \right| \le |Dc_{i,j}|, \tag{3.23}$$

where the norm of the vector is $|(a,b)^\mathsf{T}| = \sqrt{a^2+b^2}$.

– B_τ interaction: we have

$$\begin{aligned} \mathrm{d}_\theta B_\tau[c]_{i,j,k} = {} & \frac{1}{\Delta\theta} \big[\mathbf{n}\big(\theta_{k+1/2}\big) - \mathbf{n}\big(\theta_{k-1/2}\big)\big] \cdot Dc_{i,j} \\ & + \frac{\tau}{\Delta\theta} \big[\mathbf{n}\big(\theta_{k+1/2}\big) Hc_{i,j} \mathbf{e}\big(\theta_{k+1/2}\big) - \mathbf{n}\big(\theta_{k-1/2}\big) Hc_{i,j} \mathbf{e}\big(\theta_{k-1/2}\big)\big]. \end{aligned} \tag{3.24}$$

Using the mean value theorem, we obtain

$$|\mathrm{d}_\theta B_\tau[c]_{i,j,k}| \le C(|Dc_{i,j}| + |Hc_{i,j}|), \tag{3.25}$$

for some constant $C > 0$ independent of h, and where the norm of the matrix is

$$\left| \begin{pmatrix} a & b \\ c & d \end{pmatrix} \right| = \sqrt{a^2+b^2+c^2+d^2}. \tag{3.26}$$

– B_λ interaction: a pointwise estimate is given in Appendix B, and we prove that there is a constant $C > 0$ independent of the mesh size such that

$$\|\mathrm{d}_\theta B_\lambda[c_h]\|_{L^2}^2 \le C\Big(\|\mathrm{d}_\mathbf{x} c_h\|_{L^2(\Omega)}^2 + \big\|\mathrm{d}_\mathbf{x}^2 c_h\big\|_{L^2(\Omega)}^2 \Big), \tag{3.27}$$

for the details of the proof see Appendix B.

Now we derive the L^2-estimates on c_h in terms of the L^2-norm of ρ_h. Multiplying equation (2.11e) by $c_{i,j}\Delta\mathbf{x}$ and summing over all doubles (i,j) and using integration by parts gives

$$\|\mathrm{d}_\mathbf{x} c_h\|_{L^2(\Omega)}^2 + \alpha \|c_h\|_{L^2(\Omega)}^2 = \|\rho_h c_h\|_{L^1(\Omega)} \le \frac{1}{2\alpha} \|\rho_h\|_{L^2(\Omega)}^2 + \frac{\alpha}{2} \|c_h\|_{L^2(\Omega)}^2. \tag{3.28}$$

This gives the inequality to be obtained, $\|\mathrm{d}_\theta B[c_h]\|_{L^2} \le C_B \|\rho_h\|_{L^2(\Omega)}$, for the B_0 interaction. For the Hessian appearing in the B_τ and B_λ interaction terms, we want to derive a discrete equivalent of the L^2-elliptic regularity estimate, *i.e.*, $\|\nabla^2 c\|_{L^2(\Omega)} = \|\Delta c\|_{L^2(\Omega)}$. By integration by parts, we find

$$\sum_{i,j} \mathrm{d}_x^2 c_{i,j}\, \mathrm{d}_y^2 c_{i,j} \Delta\mathbf{x} = \sum_{i,j} |\mathrm{d}_x\, \mathrm{d}_y c_{i+1/2,j+1/2}|^2 \Delta\mathbf{x}. \tag{3.29}$$

Hence, the L^2-norm of the mixed derivatives is controlled by the diagonal terms of $\mathrm{d}_\mathbf{x}^2 c$, just as with the continuum PDE level. As a result, squaring equation (2.11e) and using $ab \le \frac{1}{2}(a^2+b^2)$ gives

$$\big(\mathrm{d}_x^2 c_{i,j} + \mathrm{d}_y^2 c_{i,j}\big)^2 = \alpha^2 (c_{i,j})^2 - 2\alpha c_{i,j} \rho_{i,j}^n + \big(\rho_{i,j}^n\big)^2 \le 2\alpha^2 (c_{i,j})^2 + 2\big(\rho_{i,j}^n\big)^2. \tag{3.30}$$

Summing over doubles (i,j) and multiplying by $\Delta\mathbf{x}$ and using equation (3.28) to control $\|c_h\|_{L^2(\Omega)}$ by $\|\rho_h\|_{L^2(\Omega)}$, gives the result for B_τ and B_λ as well. ☐

Proposition 3.3 (L^2-estimate for f_h)**.** *Provided that $\Delta t < \frac{D_T}{2\mathrm{Pe}^2}$, $\Delta t < C' C_T$, where $C' > 0$ is given in the proof below and C_T is as in Proposition 3.1, the meshes are uniformly regular as in Definition 2.2, and $\|f^0\|_{L^2} < +\infty$, there exists a constant $C > 0$ independent of the mesh size but depending on the final time $T > 0$ such that the solutions to the numerical scheme* (2.11) *satisfy*

$$\sup_{t\in[0,T]} \|f_h(t)\|_{L^2} + \int_0^T \|\mathrm{d}_\xi f_h(t)\|_{L^2}^2 \,\mathrm{d}t \le C. \tag{3.31}$$

Proof. The idea for the proof is a discrete analogue of Lemma 3.3 from [13]. We multiply equation (2.11) by $f^{n+1}_{i,j,k}\Delta\xi$ and sum over all triples (i,j,k). This gives

$$\begin{aligned}\sum_{i,j,k}\frac{\Delta\xi}{\Delta t}\left(f^{n+1}_{i,j,k}-f^{n}_{i,j,k}\right)f^{n+1}_{i,j,k} &= \sum_{i,j,k}\Delta\xi\Big[D_T\big(\mathrm{d}_x^2+\mathrm{d}_y^2\big)f^{n+1}_{i,j,k}f^{n+1}_{i,j,k}+\mathrm{d}_\theta^2 f^{n+1}_{i,j,k}f^{n+1}_{i,j,k}\Big]\\ &\quad-\sum_{i,j,k}\mathrm{Pe}\Delta\xi\left[\frac{1}{\Delta x}\left(U^{n+1}_{i+1/2,j,k}-U^{n+1}_{i-1/2,j,k}\right)f^{n+1}_{i,j,k}+\frac{1}{\Delta y}\left(V^{n+1}_{i,j+1/2,k}-V^{n+1}_{i,j-1/2,k}\right)f^{n+1}_{i,j,k}\right]\\ &\quad-\sum_{i,j,k}\frac{\gamma\Delta\xi}{\Delta\theta}\left(W^{n+1}_{i,j,k+1/2}-W^{n+1}_{i,j,k-1/2}\right)f^{n+1}_{i,j,k}.\end{aligned}\tag{3.32}$$

We now use summation by parts Lemma 3.2 for the discrete Laplacians, the Pe and γ drift terms. This gives

$$\begin{aligned}\sum_{i,j,k}\frac{\Delta\xi}{\Delta t}\left(f^{n+1}_{i,j,k}-f^{n}_{i,j,k}\right)f^{n+1}_{i,j,k} &= -\sum_{i,j,k}\Delta\xi\Big[D_T\Big(|\mathrm{d}_x f^{n+1}_{i+1/2,j,k}|^2+|\mathrm{d}_y f^{n+1}_{i,j+1/2,k}|^2\Big)+|\mathrm{d}_\theta f^{n+1}_{i,j,k+1/2}|^2\Big]\\ &\quad+\mathrm{Pe}\sum_{i,j,k}\Delta\xi\Big[U^{n+1}_{i+1/2,j,k}\,\mathrm{d}_x f^{n+1}_{i+1/2,j,k}+V^{n+1}_{i,j+1/2,k}\,\mathrm{d}_y f^{n+1}_{i,j+1/2,k}\Big]\\ &\quad+\gamma\sum_{i,j,k}\Delta\xi W^{n+1}_{i,j,k+1/2}\,\mathrm{d}_\theta f^{n+1}_{i,j,k+1/2}.\end{aligned}\tag{3.33}$$

Then, we have the inequality, for $a, u_i, u_{i+1}\in\mathbb{R}$ and $\Delta x>0$,

$$(a_+u_i+a_-u_{i+1})\tfrac{1}{\Delta x}(u_{i+1}-u_i)\le\tfrac{1}{2}a\tfrac{1}{\Delta x}(u_{i+1}^2-u_i^2).\tag{3.34}$$

Indeed, we calculate

$$\begin{aligned}(a_+u_i+a_-u_{i+1})\tfrac{1}{\Delta x}(u_{i+1}-u_i) &= \frac{1}{\Delta x}\left[au_iu_{i+1}-a_+u_i^2+a_-u_{i+1}^2\right]\\ &\le\frac{1}{\Delta x}\left[\tfrac{1}{2}|a|(u_i^2+u_{i+1}^2)-a_+u_i^2+a_-u_{i+1}^2\right]\\ &=\frac{a}{2\Delta x}(u_{i+1}^2-u_i^2).\end{aligned}$$

Hence, by equation (2.11d) and inequality (3.34), we have

$$U^{n+1}_{i+1/2,j,k}\,\mathrm{d}_x f^{n+1}_{i+1/2,j,k}\le\tfrac{1}{2}\cos(\theta_k)\,\mathrm{d}_x\left(f^{n+1}_{i+1/2,j,k}\right)^2,\tag{3.35}$$

and analogously for the y-direction. For the θ-direction we have

$$\left|W^{n+1}_{i,j,k+1/2}\right|\le\left|B\left[c^{n+1}\right]_{i,j,k+1/2}\right|f^{n+1}_{i,j,k}.\tag{3.36}$$

Therefore, from equation (3.33) we obtain

$$\begin{aligned}\sum_{i,j,k}\frac{\Delta\xi}{\Delta t}\left(f^{n+1}_{i,j,k}-f^{n}_{i,j,k}\right)f^{n+1}_{i,j,k} &\le-\sum_{i,j,k}\Delta\xi\left[D_T\left(\left|\mathrm{d}_x f_{i+1/2,j,k}\right|^2+\left|\mathrm{d}_y f_{i,j+1/2,k}\right|^2\right)+\left|\mathrm{d}_\theta f^{n+1}_{i,j,k+1/2}\right|^2\right]\\ &\quad+\frac{1}{2}\mathrm{Pe}\sum_{i,j,k}\Delta\xi\left[\cos(\theta_k)\,\mathrm{d}_x\left(f^{n+1}_{i+1/2,j,k}\right)^2+\sin(\theta_k)\,\mathrm{d}_y\left(f^{n+1}_{i,j+1/2,k}\right)^2\right]\\ &\quad+\frac{1}{2}\gamma\sum_{i,j,k}\Delta\xi\left|B[c^{n+1}]_{i,j,k+1/2}\right|\mathrm{d}_\theta\left(f^{n+1}_{i,j,k+1/2}\right)^2.\end{aligned}\tag{3.37}$$

We observe that the sum over the terms premultiplied by Pe in equation (3.37) equals zero. Concerning the terms multiplied by γ, integrating by parts yields

$$\begin{aligned}\frac{\gamma}{2}\sum_{i,j,k}\Delta\xi\left|B[c^{n+1}]_{i,j,k+1/2}\right|\mathrm{d}_\theta\left|f^{n+1}_{i,j,k+1/2}\right|^2 &= -\frac{\gamma}{2}\sum_{i,j,k}\Delta\xi\,\mathrm{d}_\theta\left|B\left[c^{n+1}\right]_{i,j,k}\right|\left|f^{n+1}_{i,j,k}\right|^2\\ &\le \frac{\gamma}{2}\sum_{i,j,k}\Delta\xi\left|\mathrm{d}_\theta\left|B\left[c^{n+1}\right]_{i,j,k}\right|\right|\left|f^{n+1}_{i,j,k}\right|^2\\ &= \int\left|\mathrm{d}_\theta B\left[c_h\left(t^{n+1}\right)\right]\right|\left|f_h\left(t^{n+1}\right)\right|^2\mathrm{d}\xi,\end{aligned} \tag{3.38}$$

having used the reverse triangle inequality, $||a|-|b||\le|a-b|$ for $a,b\in\mathbb{R}$. Then, the second sum in equation (3.37) can be estimated as

$$\begin{aligned}&\frac{1}{2}\mathrm{Pe}\sum_{i,j,k}\Delta\xi\left[\cos(\theta_k)\,\mathrm{d}_x\left(f^{n+1}_{i+1/2,j,k}\right)^2+\sin(\theta_k)\,\mathrm{d}_y\left(f^{n+1}_{i,j+1/2,k}\right)^2\right]\\ &\quad+\frac{1}{2}\gamma\sum_{i,j,k}\Delta\xi\left|B\left[c^{n+1}\right]_{i,j,k+1/2}\right|\mathrm{d}_\theta\left(f^{n+1}_{i,j,k+1/2}\right)^2\\ &\quad\le\frac{\gamma}{2}\int\left|\mathrm{d}_\theta B\left[c_h\left(t^{n+1}\right)\right]\right|\left|f_h\left(t^{n+1}\right)\right|^2\mathrm{d}\xi.\end{aligned} \tag{3.39}$$

Combining (3.39) and (3.37) and using $(b-a)b\ge\frac{1}{2}(b^2-a^2)$ on the left-hand of (3.37) for the time-derivative, we obtain

$$\begin{aligned}\frac{1}{2\Delta t}\left(\left\|f_h\left(t^{n+1}\right)\right\|^2_{L^2}-\left\|f_h\left(t^{n}\right)\right\|^2_{L^2}\right)&\le -D_T\left\|\mathrm{d}_{\mathbf{x}}f_h\left(t^{n+1}\right)\right\|^2_{L^2}-\left\|\mathrm{d}_\theta f_h\left(t^{n+1}\right)\right\|^2_{L^2}\\ &\quad+\frac{\gamma}{2}\int\left|\mathrm{d}_\theta B\left[c_h\left(t^{n+1}\right)\right]\right|\left|f_h\left(t^{n+1}\right)\right|^2\mathrm{d}\xi.\end{aligned} \tag{3.40}$$

Applying the Cauchy–Schwarz and the Hölder interpolation inequality, we can estimate the last term in equation (3.40) as

$$\begin{aligned}\frac{\gamma}{2}\int\left|\mathrm{d}_\theta B\left[c_h\left(t^{n+1}\right)\right]\right|\left|f_h\left(t^{n+1}\right)\right|^2\mathrm{d}\xi&\le\left\|\mathrm{d}_\theta B\left[c_h\left(t^{n+1}\right)\right]\right\|_{L^2}\left\|f_h\left(t^{n+1}\right)\right\|^2_{L^4}\\ &\le\left\|\mathrm{d}_\theta B\left[c_h\left(t^{n+1}\right)\right]\right\|_{L^2}\left\|f_h\left(t^{n+1}\right)\right\|^{\frac{3}{2}}_{L^2}\left\|f_h\left(t^{n+1}\right)\right\|^{\frac{1}{2}}_{L^6}.\end{aligned} \tag{3.41}$$

By the discrete Poincaré–Sobolev inequality ([9], Thm. 3.2), we have $\|f_h\|_{L^6}\le C_{\mathrm{PS}}\|f_h\|_{1,2}=C_{\mathrm{PS}}(\|f_h\|_{L^2}+\|\mathrm{d}_\xi f_h\|_{L^2})$, for some constant $C_{\mathrm{PS}}>0$, which is independent of the mesh size. Therefore, we arrive at the inequality

$$\frac{1}{2\Delta t}\left(\left\|f_h\left(t^{n+1}\right)\right\|^2_{L^2}-\left\|f_h\left(t^{n}\right)\right\|^2_{L^2}\right)\le -D(\varepsilon)\left\|\mathrm{d}_\xi f_h\left(t^{n+1}\right)\right\|^2_{L^2}+C(\varepsilon)\left\|\mathrm{d}_\theta B\left[c_h\left(t^{n+1}\right)\right]\right\|^2_{L^2}\left\|f_h\left(t^{n+1}\right)\right\|^2_{L^2}, \tag{3.42}$$

where $\varepsilon>0$ can be made arbitrarily small at the cost of increasing $C(\varepsilon)>0$ and decreasing $D(\varepsilon)>0$. Hence, multiplying by Δt and summing in time gives the result, provided $\Delta t<C'C_T$ for the application of the discrete Grönwall inequality (Lem. 3.1), with

$$C'=\tfrac{1}{2}C(\varepsilon)C_B. \tag{3.43}$$

□

Corollary 3.1. *Provided the conditions for Proposition 3.3 hold, there exists a constant $C>0$, independent of the mesh size, such that the solutions to the numerical scheme* (2.11) *satisfy*

$$\sup_{t\in[0,T]}\|\rho_h(t)\|_{L^2(\Omega)}\le C. \tag{3.44}$$

Proof. The result follows by applying Jensen's inequality to $\|\rho_h(t)\|^2_{L^2(\Omega)}$, such that, if $t \in [t^{n-1}, t^n)$,

$$\begin{aligned}\|\rho_h(t)\|^2_{L^2(\Omega)} &= \sum_{i,j} \left(\rho^n_{i,j}\right)^2 \Delta\mathbf{x} \\ &\le \sum_{i,j} \left(\sum_k f^n_{i,j,k} \Delta\theta\right)^2 \Delta\mathbf{x} \\ &\le \sum_{i,j,k} \left(f^n_{i,j,k}\right)^2 \Delta\xi. \end{aligned} \tag{3.45}$$

This concludes the proof. □

4. Convergence of the scheme

The estimates from Section 3 allow us to derive a convergence result on a finite time interval $[0, T]$ as stated in Theorem 2.1, using a compactness method, similar to [24].

Lemma 4.1. *For all sequences $(f_h)_h$ of numerical solutions to the scheme as in Definition 2.3, provided the meshes are an admissibly regular family as in Definition 2.2, there exists a $C > 0$ such that for all $s > 0$ and $z \in \Sigma$,*

$$\int_0^{T-s} \int_\Sigma [f_h(t+s,\xi) - f_h(t,\xi)]\varphi(t,\xi)\, \mathrm{d}\xi\, \mathrm{d}t \le Cs\|\varphi\|_{L^2(0,T;H^4(\Sigma))}, \tag{4.1}$$

for all $\varphi \in L^2(0,T;H^4(\Sigma))$ and

$$\int_0^T \int_{\Sigma - z} |f_h(t,\xi+z) - f_h(t,\xi)|^2\, \mathrm{d}\xi\, \mathrm{d}t \le C|z|, \tag{4.2}$$

where $\Sigma - z = \{\xi - z : \xi \in \Sigma\} \cap \Sigma$.

Proof. Following Proof of Proposition 4.1 from [24], we begin by showing that there exists $C > 0$, such that

$$\|f_h\|_{L^3(\Sigma_T)} \le C. \tag{4.3}$$

To this end, we use the discrete TV inequality ([9], Thm. 2.4.1), so that for $u = |f_h|^2$ and a constant $C_{\mathrm{TV}} > 0$, we have, using the notation $\bar{u} = \frac{1}{|\Sigma|}\int_\Sigma u\, \mathrm{d}\xi$,

$$\|f_h\|^3_{L^3(\Sigma)} = \|u\|^{\frac{3}{2}}_{L^{\frac{3}{2}}(\Sigma)} \le \left(\|u - \bar{u}\|_{L^{\frac{3}{2}}(\Sigma)} + \|\bar{u}\|_{L^{\frac{3}{2}}(\Sigma)}\right)^{\frac{3}{2}} \le \left(C_{\mathrm{TV}}|u|_{1,1,\Sigma} + \|\bar{u}\|_{L^{\frac{3}{2}}(\Sigma)}\right)^{\frac{3}{2}}. \tag{4.4}$$

Hence, using $(x+y)^{\frac{3}{2}} \le \frac{3}{2}(x^{\frac{3}{2}} + y^{\frac{3}{2}})$ for $x, y \in \mathbb{R}_{\ge 0}$, we get

$$\|f_h\|^3_{L^3} \le \tfrac{3}{2}\left(C_{\mathrm{TV}}^{\frac{3}{2}}\left\|\mathrm{d}_\xi|f_h|^2\right\|^{\frac{3}{2}}_{L^1} + |\Sigma|\|f_h\|^3_{L^2}\right) \le C\left(\left\|\mathrm{d}_\xi|f_h|^2\right\|^{\frac{3}{2}}_{L^1} + \|f_h\|^3_{L^2}\right). \tag{4.5}$$

We used here that $\|\bar{u}\|^{3/2}_{L^{3/2}} = \int(\int |f_h|^2\, \mathrm{d}\xi)^{3/2}\, \mathrm{d}\xi = |\Sigma|\|f_h\|^3_{L^2}$.

Using $a^2 - b^2 = (a+b)(a-b)$, we estimate $\mathrm{d}_\xi|f_h|^2$ and get

$$\int_0^T \|f_h(t)\|^3_{L^3}\, \mathrm{d}t \le C\int_0^T \left(\left(\|f_h(t)\|_{L^2}\|\mathrm{d}_\xi f_h(t)\|_{L^2}\right)^{\frac{3}{2}} + \|f_h(t)\|^3_{L^2}\right) \mathrm{d}t. \tag{4.6}$$

Finally, we use the Hölder inequality and the uniform L^2-bounds on f_h and $\mathrm{d}_\xi f_h$ from Proposition 3.3 so that there is a constant $C > 0$ independent of h such that $\|f_h\|_{L^3(\Sigma_T)} \le C$.

Now turning to the time translate, for $\varphi \in L^2(0,T;H^4(\Sigma))$, we get, defining $M\Delta t \leq s < (M+1)\Delta t$ for $M \in \mathbb{N}$, and for $\varphi^n_{i,j,k} = \int_{t^{n-1}}^{t^n} \int_{C_{i,j,k}} \varphi(t,\xi)\,\mathrm{d}\xi\,\mathrm{d}t$,

$$\begin{aligned} I_s &= \int_0^{T-s} \int_\Sigma (f_h(t+s,\xi) - f_h(t,\xi))\varphi(t,\xi)\,\mathrm{d}\xi\,\mathrm{d}t \\ &= \sum_{n=1}^{N_T-M} \sum_{i,j,k} \Big(f^{n+M}_{i,j,k} - f^n_{i,j,k} \Big) \varphi^n_{i,j,k} \Delta\xi \Delta t \\ &= \sum_{n=1}^{N_T-M} \sum_{i,j,k} \sum_{m=1}^{M} \frac{1}{\Delta t} \Big(f^{n+m}_{i,j,k} - f^{n+m-1}_{i,j,k} \Big) \varphi^n_{i,j,k} \Delta\xi \Delta t^2. \end{aligned} \tag{4.7}$$

We insert the scheme as in equation (2.11) into I_s to obtain, after summation by parts,

$$\begin{aligned} I_s = -D_T &\sum_{n=1}^{N_T-M} \sum_{i,j,k} \sum_{m=1}^{M} \Big(\mathrm{d}_x f^{n+m}_{i+1/2,j,k}\, \mathrm{d}_x \varphi^n_{i+1/2,j,k} + \mathrm{d}_y f^{n+m}_{i,j+1/2,k}\, \mathrm{d}_y \varphi^n_{i,j+1/2,k} \Big) \Delta\xi \Delta t^2 \\ &+ \mathrm{Pe} \sum_{n=1}^{N_T-M} \sum_{i,j,k} \sum_{m=1}^{M} \Big(U^{n+m}_{i+1/2,j,k}\, \mathrm{d}_x \varphi^n_{i+1/2,j,k} + V^{n+m}_{i,j+1/2,k}\, \mathrm{d}_y \varphi^n_{i,j+1/2,k} \Big) \Delta\xi \Delta t^2 \\ &- \sum_{n=1}^{N_T-M} \sum_{i,j,k} \sum_{m=1}^{M} \mathrm{d}_\theta f^{n+m}_{i,j,k+1/2}\, \mathrm{d}_\theta \varphi^n_{i,j,k+1/2} \Delta\xi \Delta t^2 + I_s^\gamma, \end{aligned} \tag{4.8}$$

where I_s^γ is defined as

$$I_s^\gamma = \gamma \sum_{n=1}^{N_T-M} \sum_{i,j,k} \sum_{m=1}^{M} W^{n+m}_{i,j,k+1/2}\, \mathrm{d}_\theta \varphi^n_{i,j,k+1/2} \Delta\xi \Delta t^2. \tag{4.9}$$

It is clear that, except for the I_s^γ term, using Cauchy–Schwarz, we get, for some constant $C > 0$, using the uniform L^2-bounds on f_h and $\mathrm{d}_\xi f_h$ from Proposition 3.3,

$$|I_s| \leq Cs\|\varphi\|_{L^2(0,T;H^4(\Sigma))} + |I_s^\gamma|. \tag{4.10}$$

For the I_s^γ term we get

$$|I_s^\gamma| \leq \gamma \sum_{n=1}^{N_T-M} \sum_{i,j,k} \sum_{m=1}^{M} \Big| B\left[c^{n+m}\right]_{i,j,k+1/2} \Big| \Big(\Big| f^{n+m}_{i,j,k} \Big| + \Big| f^{n+m}_{i,j,k+1} \Big| \Big) \Big| \mathrm{d}_\theta \varphi^n_{i,j,k+1/2} \Big| \Delta\xi \Delta t^2. \tag{4.11}$$

By the definitions for $B \in \{B_0, B_\lambda, B_\tau\}$ and the discrete elliptic estimates on c_h in terms of $L^2(\Omega)$-norms of ρ_h (cf. Prop. 3.2), we obtain, by Corollary 3.1,

$$\|B[c_h(t)]\|_{L^2} \leq C_B \|\rho_h(t)\|_{L^2(\Omega)} \leq C, \tag{4.12}$$

for some $C > 0$. Hence, by using Hölder's inequality, we can estimate, for some constant $C > 0$,

$$\begin{aligned} |I_s^\gamma| &\leq 2\gamma \sum_{m=1}^{M} \|B[c_h]\|_{L^6(0,T;L^2(\Sigma))} \|f_h\|_{L^3(\Sigma_T)} \|\mathrm{d}_\theta \varphi_h\|_{L^2(0,T;L^6(\Sigma))} \Delta t \\ &\leq Cs\|\mathrm{d}_\theta \varphi_h\|_{L^2(0,T;L^6(\Sigma))} \\ &\leq Cs\|\varphi\|_{L^2(0,T;H^4(\Sigma))}, \end{aligned} \tag{4.13}$$

where we used ([22], Thm. 3 in Sect. 5.8.2) for the difference quotient estimate $\|\mathrm{d}_\theta \varphi_h\|_{L^6(\Sigma)} \le C\|\partial_\theta \varphi\|_{L^6(\Sigma)}$ for some constant $C > 0$, and we used the Sobolev embedding in the last line.

Finally, the phase-space translate result can be obtained in a similar fashion as in a classical theorem for finite volume schemes ([23], Lem. 9.3). Indeed, let (i_z, j_z, k_z) be such that $z \in C_{i_z,j_z,k_z}$, then

$$\int_0^T \int_{\Sigma - z} |f_h(t,\xi+z) - f_h(t,\xi)|^2 \,\mathrm{d}\xi\,\mathrm{d}t = \sum_{n=1}^{N_T} \sum_{\substack{i,j,k:\\ C_{i,j,k}\subseteq \Sigma - z}} \left| f^n_{i+i_z,j+j_z,k+k_z} - f^n_{i,j,k}\right|^2 \Delta\xi\Delta t$$

$$= \sum_{n=1}^{N_T} \sum_{\substack{i,j,k:\\ C_{i,j,k}\subseteq \Sigma - z}} \left(f^n_{i+i_z,j+j_z,k+k_z} - f^n_{i,j,k}\right)$$

$$\times \left[\sum_{i'=i}^{i+i_z-1} \mathrm{d}_x f^n_{i'+1/2,j,k}\Delta x + \sum_{j'=j}^{j+j_z-1} \mathrm{d}_y f^n_{i+i_z,j'+1/2,k}\Delta y + \sum_{k'=k}^{k+k_z-1} \mathrm{d}_\theta f^n_{i+i_z,j+j_z,k'+1/2}\Delta\theta \right] \Delta\xi\Delta t. \quad (4.14)$$

Hence, by Cauchy–Schwarz, we get

$$\int_0^T \int_{\Sigma - z} |f_h(t,\xi+z) - f_h(t,\xi)|^2 \mathrm{d}\xi\,\mathrm{d}t \le C|z| \|f_h\|_{L^2(\Sigma_T)} \|\mathrm{d}_\xi f_h\|_{L^2(\Sigma_T)} \le C|z|, \quad (4.15)$$

having used Proposition 3.3. This concludes the proof. □

Lemma 4.2. *For all sequences $(f_h)_h$ of numerical solutions to the scheme as in Definition 2.3, provided the meshes are an admissibly regular family as in Definition 2.2, there is a subsequence that converges strongly to a function $f \in L^2(\Sigma_T)$, the discrete gradient $(\mathrm{d}_\xi f_h)_h$ converges weakly in $L^2(\Sigma_T)$ to $\nabla_\xi f \in L^2(\Sigma_T)$ and $B[c_h]$ converges weakly-∗ in $L^\infty(0,T;L^2)$ to $B[c]$ such that c is the unique strong solution for the elliptic equation with $\rho = \int_0^{2\pi} f \mathrm{d}\theta$ (possibly up to another subsequence).*

Proof. By the Banach–Alaoglu theorem and the uniform bounds of Proposition 3.3 on f_h in $L^2(\Sigma_T)$ we can extract a subsequence that converges weakly in $L^2(\Sigma_T)$ to some $f \in L^2(\Sigma_T)$. We first show that this convergence is strong in $L^2(\Sigma_T)$.

By Lemma 4.1 we have that $(f_h)_h$ is bounded in $L^2(0,T;X_3(\Sigma))$, where $X_3(\Sigma)$ is as in Section 4 from [24], and that

$$\|f_h(t+s) - f_h(t)\|_{L^2(0,T;(H^4(\Sigma))')} \le Cs. \quad (4.16)$$

We also have the continuous embeddings

$$X_3(\Sigma) \subset L^2(\Sigma) \subset (H^4(0\Sigma))', \quad (4.17)$$

and $X_3(\Sigma) \subset L^2(\Sigma)$ is a compact embedding by the Riesz–Kolmogorov–Fréchet theorem ([10], Thm. 4.26). Hence, by the Aubin–Lions–Simon theorem ([31], Thm. 5), we have that $f_h \to f$ strongly in $L^2(\Sigma_T)$ (up to a subsequence).

From the Banach–Alaoglu theorem, we derive that $\mathrm{d}_\xi f_h \rightharpoonup \nabla_\xi f$, weakly in $L^2(\Sigma_T)$.

By the uniform bounds in $L^\infty(0,T;L^2(\Omega))$ for $c_h, \mathrm{d}_{\mathbf{x}} c_h$ and $\mathrm{d}^2_{\mathbf{x}} c_h$ of Proposition 3.2, it follows that there are functions $c, \mathbf{g}$ and $\mathbf{M}$ in $L^\infty(0,T;L^2(\Omega))$ such that there is a subsequence and there is component-wise convergence

$$c_h \overset{*}{\rightharpoonup} c, \quad \mathrm{d}_{\mathbf{x}} c_h \overset{*}{\rightharpoonup} \mathbf{g}, \quad \mathrm{d}^2_{\mathbf{x}} c_h \overset{*}{\rightharpoonup} \mathbf{M}, \quad (4.18)$$

weakly-∗ in $L^\infty(0,T;L^2(\Omega))$. Similar as in Lemma 4.4 from [19], it follows that $\mathbf{g} = \nabla_{\mathbf{x}} c$ and $\mathbf{M} = \nabla^2_{\mathbf{x}} c$ in distribution.

Following the argument in Proposition 4.2 of [24], we next show that the limit c satisfies the elliptic equation $0 = \Delta c - \alpha c + \rho$, where ρ is the limit of ρ_h, and $\rho = \int_0^{2\pi} f(t,\xi)\,\mathrm{d}\theta$ by the convergence of f_h. To this end, let $\varphi \in C^\infty(\overline{\Omega_T})$ be given and define

$$I_h = \int_{\Omega_T} (-\mathrm{d}_{\mathbf{x}} c_h \cdot \nabla_{\mathbf{x}}\varphi + (\rho_h - \alpha c_h)\varphi)\,\mathrm{d}\mathbf{x}\,\mathrm{d}t, \tag{4.19}$$

and show that

$$I_h \to \int_{\Omega_T} (-\nabla_{\mathbf{x}} c \cdot \nabla_{\mathbf{x}}\varphi + (\rho - \alpha c)\varphi)\,\mathrm{d}\mathbf{x}\,\mathrm{d}t, \tag{4.20}$$

as well as

$$I_h = \int_{\Omega_T} (-\mathrm{d}_{\mathbf{x}} c_h \cdot \nabla_{\mathbf{x}}\varphi + (\rho_h - \alpha c_h)\varphi)\,\mathrm{d}\mathbf{x}\,\mathrm{d}t \to 0, \tag{4.21}$$

as $h \to 0$.

The first convergence in equation (4.20) follows from the weak-$*$ convergence. For the second convergence, equation (4.21), we multiply the equation for c_h, equation (2.11e), by $\varphi^n_{i,j} = \int_{t^{n-1}}^{t^n} \int_{C_{i,j}} \varphi(t,\mathbf{x})\,\mathrm{d}\mathbf{x}\,\mathrm{d}t$ and $\Delta\mathbf{x}$, Δt, and sum over all i,j,n to get

$$I_h^S = \sum_{n=0}^{N_T-1}\sum_{i,j} \big(\mathrm{d}_x^2 c^n_{i,j} + \mathrm{d}_y^2 c^n_{i,j}\big)\varphi^n_{i,j}\Delta\mathbf{x}\Delta t + \sum_{n=0}^{N_T-1}\sum_{i,j}(-\alpha c^n_{i,j} + \rho^n_{i,j})\varphi^n_{i,j}\Delta\mathbf{x}\Delta t = 0. \tag{4.22}$$

Then, a discrete integration by parts yields

$$\begin{aligned} I_h^S = &\sum_{n=0}^{N_T-1}\sum_{i,j}\Big(\mathrm{d}_x c^n_{i+1/2,j}\,\mathrm{d}_x\varphi^n_{i+1/2,j} + \mathrm{d}_y c^n_{i,j+1/2}\,\mathrm{d}_y\varphi^n_{i,j+1/2}\Big)\Delta\mathbf{x}\Delta t \\ &- \sum_{n=0}^{N_T-1}\sum_{i,j}\big(\rho^n_{i,j} - \alpha c^n_{i,j}\big)\varphi^n_{i,j}\Delta\mathbf{x}\Delta t. \end{aligned} \tag{4.23}$$

We now show that $I_h - I_h^S \to 0$, as $h \to 0$. Using the mean value theorem, Corollary 3.1 and Proposition 3.2 and writing $\varphi^n_{i,j} = \varphi(t^n, \mathbf{x}_{i,j})$,

$$\begin{aligned} \int_{\Omega_T} (\rho_h - \alpha c_h)\varphi\,\mathrm{d}\mathbf{x}\,\mathrm{d}t &- \sum_{n=0}^{N_T-1}\sum_{i,j}\big(\rho^n_{i,j} - \alpha c^n_{i,j}\big)\varphi^n_{i,j}\Delta\mathbf{x}\Delta t \\ &= \sum_{n=0}^{N_T-1}\sum_{i,j}\int_{t^n}^{t^n}\int_{C_{i,j}} \big(\rho^n_{i,j} - \alpha c^n_{i,j}\big)\big(\varphi(t,\mathbf{x}) - \varphi^n_{i,j}\big)\,\mathrm{d}\mathbf{x}\,\mathrm{d}t \\ &\to 0. \end{aligned} \tag{4.24}$$

Similarly, for the gradient terms, we have

$$\begin{aligned} \int_{\Omega_T} \mathrm{d}_x c_h \partial_x\varphi\,\mathrm{d}\mathbf{x}\,\mathrm{d}t &- \sum_{n=0}^{N_T-1}\sum_{i,j}\mathrm{d}_x c^n_{i+1/2,j}\,\mathrm{d}_x\varphi^n_{i+1/2,j}\Delta\mathbf{x}\Delta t \\ &= \sum_{n=0}^{N_T-1}\sum_{i,j}\mathrm{d}_x c^n_{i+1/2,j}\left[\int_{t^{n-1}}^{t^n}\int_{C_{i+1/2,j}}\Big(\partial_x\varphi - \mathrm{d}_x\varphi^n_{i+1/2,j}\Big)\,\mathrm{d}\mathbf{x}\,\mathrm{d}t\right] \\ &\to 0. \end{aligned} \tag{4.25}$$

In conclusion, we have shown that

$$I_h - I_h^S \to 0. \tag{4.26}$$

This implies that c solves the elliptic equation on Ω_T. Similarly, using the $L^\infty(0,T;L^2(\Omega))$ bounds on c_h and $\mathrm{d}_{\mathbf{x}}c_h$ we can also show that c solves the elliptic equation pointwise in time and in the strong sense.

Finally, for each $B \in \{B_0, B_\lambda, B_\tau\}$ we have $B[c_h] \overset{*}{\rightharpoonup} B[c]$ weakly-$*$ in $L^\infty(0,T;L^2(\Sigma))$. Indeed, we have the following computations.

First we rewrite the difference $B[c_h] - B[c]$ as follows. For $(x,y,\theta) \in C_{i,j,k+1/2}$ and $t \in (t^{n-1}, t^n]$, we observe that

$$\begin{aligned} B_0[c_h] - B_0[c] &= \mathbf{n}\big(\theta_{k+1/2}\big) \cdot Dc_{i,j}^n - \mathbf{n}(\theta) \cdot \nabla_{\mathbf{x}} c \\ &= \big(\mathbf{n}\big(\theta_{k+1/2}\big) - \mathbf{n}(\theta)\big) \cdot Dc_{i,j}^n - \mathbf{n}(\theta) \cdot \big(\nabla_{\mathbf{x}} c - Dc_{i,j}^n\big). \end{aligned} \tag{4.27}$$

as well as

$$\begin{aligned} B_\lambda[c_h] - B_\lambda[c] &= \mathbf{n}\big(\theta_{k+1/2}\big) \cdot Dc_{i,j,k+1/2}^n - \mathbf{n}(\theta) \cdot \nabla_{\mathbf{x}} c_\lambda \\ &= \big(\mathbf{n}(\theta_{k+1/2}) - \mathbf{n}(\theta)\big) \cdot Dc_{i,j,k+1/2}^n - \mathbf{n}(\theta) \cdot \Big(\nabla_{\mathbf{x}} c_\lambda - Dc_{i,j,k+1/2}^n\Big) \end{aligned} \tag{4.28}$$

and

$$\begin{aligned} B_\tau[c_h] - B_\tau[c] &= \mathbf{n}\big(\theta_{k+1/2}\big) \cdot Dc_{i,j}^n + \tau \mathbf{n}\big(\theta_{k+1/2}\big) \cdot Hc_{i,j}^n \mathbf{e}\big(\theta_{k+1/2}\big) - \mathbf{n}(\theta) \cdot \nabla c - \tau \mathbf{n}(\theta) \nabla_{\mathbf{x}}^2 c \mathbf{e}(\theta) \\ &= \big(\mathbf{n}\big(\theta_{k+1/2}\big) - \mathbf{n}(\theta)\big) \cdot Dc_{i,j}^n - \mathbf{n}(\theta) \cdot \big(\nabla_{\mathbf{x}} c - Dc_{i,j}^n\big) \\ &\quad + \tau\big(\mathbf{e}\big(2\theta_{k+1/2}\big) - \mathbf{e}(2\theta)\big) \cdot \tilde{H}c_{i,j}^n - \tau \mathbf{e}(2\theta) \cdot \Big(\tilde{\nabla}_{\mathbf{x}}^2 c - \tilde{H}c_{i,j}^n\Big), \end{aligned} \tag{4.29}$$

and where we used the notation

$$\tilde{H}c_{i,j} = \big(D_x D_y c_{i,j}, \tfrac{1}{2}\big(\mathrm{d}_y^2 c_{i,j} - \mathrm{d}_x^2 c_{i,j}\big)\big)^\mathsf{T}, \quad \tilde{\nabla}_{\mathbf{x}}^2 c = \big(\partial_{xy} c, \tfrac{1}{2}(\partial_y^2 c - \partial_x^2 c)\big)^\mathsf{T}, \tag{4.30}$$

to rewrite the Hessian term, and the identity

$$\mathbf{n}(\theta) \cdot \begin{pmatrix} a & c \\ c & d \end{pmatrix} \mathbf{e}(\theta) = \mathbf{e}(2\theta) \cdot \begin{pmatrix} c \\ \frac{1}{2}(d-a) \end{pmatrix}. \tag{4.31}$$

Now, let us begin by addressing the term B_0. To this end, we let $\varphi \in L^1(0,T;L^2(\Sigma))$ and observe that

$$\begin{aligned} \mathcal{I}_h &= \int_{\Sigma_T} (B_0[c_h] - B_0[c]) \varphi \,\mathrm{d}\xi \,\mathrm{d}t \\ &= \sum_{n=1}^{N_T} \int_{t^{n-1}}^{t^n} \sum_{i,j,k} \int_{C_{i,j,k+1/2}} \big(\mathbf{n}(\theta_{k+1/2}) \cdot Dc_{i,j}^n - \mathbf{n}(\theta) \cdot \nabla_{\mathbf{x}} c\big) \varphi \,\mathrm{d}\xi \,\mathrm{d}t \\ &= \sum_{n=1}^{N_T} \int_{t^{n-1}}^{t^n} \sum_{i,j,k} \int_{C_{i,j,k+1/2}} \big(\big(\mathbf{n}\big(\theta_{k+1/2}\big) - \mathbf{n}(\theta)\big) \cdot Dc_{i,j}^n - \mathbf{n}(\theta) \cdot \big(\nabla_{\mathbf{x}} c - Dc_{i,j}^n\big)\big) \varphi \,\mathrm{d}\xi \,\mathrm{d}t. \end{aligned} \tag{4.32}$$

Passing to the modulus, we get, for some $C > 0$,

$$\begin{aligned} |\mathcal{I}_h| \le C \Bigg(& \|\mathbf{n}_h - \mathbf{n}\|_{L^\infty} \|\mathrm{d}_{\mathbf{x}} c_h\|_{L^\infty(0,T;L^2(\Omega))} \|\varphi\|_{L^1(0,T;L^2)} \\ & + \left| \int_{\Sigma_T} (\partial_x c - \mathrm{d}_x c_h) \sin(\theta) \varphi \,\mathrm{d}\xi \,\mathrm{d}t \right| + \left| \int_{\Sigma_T} (\partial_y c - \mathrm{d}_y c_h) \cos(\theta) \varphi \,\mathrm{d}\xi \,\mathrm{d}t \right| \Bigg), \end{aligned} \tag{4.33}$$

where we use the notation $\mathbf{n}_h$ to denote the piecewise-constant interpolation, such that $\mathbf{n}_h(\xi) = \mathbf{n}(\theta_{k+1/2})$ when $\xi \in C_{i,j,k+1/2}$.

By the mean-value theorem, we have

$$\|\mathbf{n}_h - \mathbf{n}\|_{L^\infty} \le h. \tag{4.34}$$

We also have that $\sin(\theta)\varphi, \cos(\theta)\varphi \in L^1(0,T;L^2)$, and $\mathrm{d}_x c_h \overset{*}{\rightharpoonup} \partial_x c$ and $\mathrm{d}_y c_h \overset{*}{\rightharpoonup} \partial_y c$ weakly-$*$ in $L^\infty(0,T;L^2(\Omega))$. Hence, as $h \to 0$, we have $|\mathcal{I}_h| \to 0$, *i.e.*,

$$B_0[c_h] \overset{*}{\rightharpoonup} B[c], \quad \text{in } L^\infty(0,T;L^2). \tag{4.35}$$

Similarly, we can show that

$$B_\lambda[c_h] \overset{*}{\rightharpoonup} B_\lambda[c], \quad B_\tau[c_h] \overset{*}{\rightharpoonup} B_\tau[c], \quad \text{in } L^\infty(0,T;L^2). \tag{4.36}$$

This concludes the proof. □

Proof of Theorem 2.1. Let $\varphi \in C^\infty(\overline{\Sigma_T})$ such that $\varphi(T) = 0$ and define the error term $\epsilon(h)$ as

$$\begin{aligned}\epsilon(h) = &-\int_0^T \int_\Sigma f_h \partial_t \varphi \,\mathrm{d}\xi\,\mathrm{d}t - \int_0^T \int_\Sigma [(D_T \mathrm{d}_{\mathbf{x}} f_h - \mathrm{Pe}\mathbf{e}(\theta) f_h)\cdot\nabla_{\mathbf{x}}\varphi + (\mathrm{d}_\theta f_h - \gamma B[c_h] f_h)\,\partial_\theta\varphi]\,\mathrm{d}\xi\,\mathrm{d}t \\ &- \int_\Sigma f_h(0)\varphi(0)\,\mathrm{d}\xi.\end{aligned} \tag{4.37}$$

By Lemma 4.2, we have that $\epsilon(h)$ converges to the left-hand side of equation (2.2), that is,

$$\begin{aligned}\epsilon(h) \to &-\int_0^T \int_\Sigma [f\partial_t\varphi - (D_T\nabla_{\mathbf{x}} f - \mathrm{Pe}\mathbf{e}(\theta) f)\cdot\nabla_{\mathbf{x}}\varphi - (\partial_\theta f - \gamma B[c] f)\partial_\theta\varphi]\,\mathrm{d}\xi\,\mathrm{d}t \\ &- \int_\Sigma f(0)\varphi(0)\,\mathrm{d}\xi.\end{aligned} \tag{4.38}$$

It remains to prove that $\epsilon(h) \to 0$ for the subsequence of Lemma 4.2, so that the limit f is indeed a weak solution as in Definition 2.1.

To do this, we compare $\epsilon(h)$ with the scheme. We multiply the scheme as in equation (2.11) by $\varphi^{n+1}_{i,j,k}$, where

$$\varphi^n_{i,j,k} = \frac{1}{\Delta\xi}\int_{C_{i,j,k}} \varphi(t^n,\xi)\,\mathrm{d}\xi, \tag{4.39}$$

and sum from $n = 0$ to N_T and over all cells so that

$$0 = \sum_{n=0}^{N_T}\sum_{i,j,k}\left[\frac{f^{n+1}_{i,j,k} - f^n_{i,j,k}}{\Delta t} + \mathrm{d}_x F^{x,n+1}_{i,j,k} + \mathrm{d}_y F^{y,n+1}_{i,j,k} + \mathrm{d}_\theta F^{\theta,n+1}_{i,j,k}\right]\varphi^{n+1}_{i,j,k}\Delta\xi\Delta t. \tag{4.40}$$

By summation by parts and $\varphi(T) = 0 = \varphi^{N_T+1}_{i,j,k}$ we obtain

$$\begin{aligned}0 = &-\sum_{i,j,k}\left[\sum_{n=1}^{N_T} f^{n+1}_{i,j,k}(\varphi^{n+1}_{i,j,k} - \varphi^n_{i,j,k}) + \varphi^1_{i,j,k} f^0_{i,j,k}\right]\Delta\xi \\ &- \sum_{n=0}^{N_T}\sum_{i,j,k}\Big[\mathrm{d}_x\varphi^{n+1}_{i+1/2,j,k} F^{x,n+1}_{i+1/2,j,k} + \mathrm{d}_y\varphi^{n+1}_{i,j+1/2,k} F^{y,n+1}_{i,j+1/2,k} + \mathrm{d}_\theta\varphi^{n+1}_{i,j,k+1/2} F^{\theta,n+1}_{i,j,k+1/2}\Big]\Delta\xi\Delta t.\end{aligned} \tag{4.41}$$

We compare each component of the scheme with its continuum counterpart. Similar as in [4], we write

$$0 = \hat{\mathcal{T}}(h) + \hat{\mathcal{D}}(h) + \hat{\mathcal{P}}(h) + \hat{\mathcal{U}}(h) + \hat{\mathcal{E}}(h), \tag{4.42}$$

where

$$\hat{\mathcal{T}}(h) = -\sum_{i,j,k}\left[\sum_{n=1}^{N_T} f_{i,j,k}^{n}\left(\varphi_{i,j,k}^{n+1} - \varphi_{i,j,k}^{n}\right) + \varphi_{i,j,k}^{1} f_{i,j,k}^{0}\right]\Delta\xi, \tag{4.43}$$

$$\hat{\mathcal{D}}(h) = \sum_{n=0}^{N_T}\sum_{i,j,k}\Big[D_T\Big(\mathrm{d}_x\varphi_{i+1/2,j,k}^{n+1}\,\mathrm{d}_x f_{i+1/2,j,k}^{n+1} + \mathrm{d}_y\varphi_{i,j+1/2,k}^{n+1}\,\mathrm{d}_y f_{i,j+1/2,k}^{n+1}\Big) + \mathrm{d}_\theta\varphi_{i,j,k+1/2}^{n+1}\,\mathrm{d}_\theta f_{i,j,k+1/2}^{n+1}\Big]\Delta\xi\Delta t, \tag{4.44}$$

$$\hat{\mathcal{P}}(h) = -\sum_{n=0}^{N_T}\sum_{i,j,k}\mathrm{Pe}\Big[\mathrm{d}_x\varphi_{i+1/2,j,k}^{n+1}\cos(\theta_k) f_{i,j,k}^{n+1} + \mathrm{d}_y\varphi_{i,j+1/2,k}^{n+1}\sin(\theta_k) f_{i,j,k}^{n+1}\Big]\Delta\xi\Delta t, \tag{4.45}$$

$$\hat{\mathcal{U}}(h) = -\sum_{n=0}^{N_T}\sum_{i,j,k}\gamma\Big[\mathrm{d}_\theta\varphi_{i,j,k+1/2}^{n+1} B\big[c^{n+1}\big]_{i,j,k+1/2} f_{i,j,k}^{n+1}\Big]\Delta\xi\Delta t, \tag{4.46}$$

$$\begin{aligned}\hat{\mathcal{E}}(h) = {} & \sum_{n=0}^{N_T}\sum_{i,j,k}\mathrm{Pe}\Big[\Delta x(\cos(\theta_k))^{-}\,\mathrm{d}_x f_{i+1/2,j,k}^{n+1}\,\mathrm{d}_x\varphi_{i+1/2,j,k}^{n+1} \\ & + \Delta y(\sin(\theta_k))^{-}\,\mathrm{d}_y f_{i,j+1/2,k}^{n+1}\,\mathrm{d}_y\varphi_{i,j+1/2,k}^{n+1}\Big]\Delta\xi\Delta t \\ & + \sum_{n=0}^{N_T}\sum_{i,j,k}\gamma\Big[\Delta\theta\Big(B\big[c^{n+1}\big]_{i,j,k+1/2}\Big)^{-}\,\mathrm{d}_\theta f_{i,j,k+1/2}^{n+1}\,\mathrm{d}_\theta\varphi_{i,j,k+1/2}^{n+1}\Big]\Delta\xi\Delta t.\end{aligned} \tag{4.47}$$

Similarly, we write equation (4.37) as

$$\epsilon(h) = \mathcal{T}(h) + \mathcal{D}(h) + \mathcal{P}(h) + \mathcal{U}(h), \tag{4.48}$$

where

$$\mathcal{T}(h) = -\int_\Sigma\left[\int_0^T f_h\partial_t\varphi\,\mathrm{d}t + f_h(0,\xi)\varphi(0,\xi)\right]\mathrm{d}\xi, \tag{4.49}$$

$$\mathcal{D}(h) = \int_0^T\int_\Sigma [D_T\mathrm{d}_{\mathbf{x}} f_h\cdot\nabla_{\mathbf{x}}\varphi + \mathrm{d}_\theta f_h\partial_\theta\varphi]\,\mathrm{d}\xi\,\mathrm{d}t, \tag{4.50}$$

$$\mathcal{P}(h) = -\int_0^T\int_\Sigma \mathrm{Pe} f_h\mathbf{e}_h(\theta)\cdot\nabla_{\mathbf{x}}\varphi\,\mathrm{d}\xi\,\mathrm{d}t, \tag{4.51}$$

$$\mathcal{U}(h) = -\int_0^T\int_\Sigma \gamma B[c_h] f_h\partial_\theta\varphi\,\mathrm{d}\xi\,\mathrm{d}t. \tag{4.52}$$

We now compare the terms $\hat{\mathcal{T}}, \hat{\mathcal{D}}, \hat{\mathcal{P}}$ and $\hat{\mathcal{U}}$ with their continuum counterpart $\mathcal{T}, \mathcal{D}, \mathcal{P}$ and $\mathcal{U}$, respectively, and show that their difference converges to zero. We also show that $\hat{\mathcal{E}}(h) \to 0$, so that, in conclusion, $\epsilon(h) \to 0$.

Clearly, as in Proof of Theorem 2.1 from [4], $\mathcal{T}(h) - \hat{\mathcal{T}}(h) = 0$ for all h. Similarly, the term $\hat{\mathcal{E}}(h)$ goes to zero by using Cauchy–Schwarz, the uniform $L^2(\Sigma_T)$-bound on $\mathrm{d}_\xi f_h$ and $B[c_h]$ and $\|\varphi\|_{C^2(\Sigma_T)}$.

Indeed, using equation (4.39), we find

$$\begin{aligned}\mathcal{T}(h)-\hat{\mathcal{T}}(h) &= \sum_{i,j,k}\left(\sum_{n=1}^{N_T} f_{i,j,k}^n\left[\left(\varphi_{i,j,k}^{n+1}-\varphi_{i,j,k}^n\right)\Delta\xi-\int_{C_{i,j,k}}\int_{t^n}^{t^{n+1}}\partial_t\varphi\,\mathrm{d}t\,\mathrm{d}\xi\right]\right)\\ &\quad+\sum_{i,j,k} f_{i,j,k}^0\left[\varphi_{i,j,k}^1-\frac{1}{\Delta\xi}\int_{C_{i,j,k}}\left(\int_{t^0}^{t^1}\partial_t\varphi\,\mathrm{d}t+\varphi(0,\xi)\right)\mathrm{d}\xi\right]\Delta\xi\\ &= 0.\end{aligned}\tag{4.53}$$

Next, let us address $\hat{\mathcal{E}}(h)$. Focusing on the term premultiplied by γ, and application of the Cauchy–Schwarz inequality, we have

$$\begin{aligned}&\sum_{n=0}^{N_T}\sum_{i,j,k}\gamma\left[\Delta\theta\left(B\left[c^{n+1}\right]_{i,j,k+1/2}\right)^{-}\mathrm{d}_\theta f_{i,j,k+1/2}^{n+1}\,\mathrm{d}_\theta\varphi_{i,j,k+1/2}^{n+1}\right]\Delta\xi\Delta t\\ &\quad\le \Delta\theta\gamma\|\varphi\|_{C^1(\overline{\Sigma_T})}\|B[c_h]\|_{L^2(\Sigma_T)}\|\mathrm{d}_\theta f_h\|_{L^2(\Sigma_T)}\\ &\quad\to 0.\end{aligned}\tag{4.54}$$

The other terms in $\hat{\mathcal{E}}(h)$ go to zero by a similar argument.

Concerning the difference $\mathcal{D}(h)-\hat{\mathcal{D}}(h)$, we focus on the x-terms noting that the terms corresponding to y and θ are treated in the same manner. The x-terms in $\mathcal{D}(h)-\hat{\mathcal{D}}(h)$ are

$$\begin{aligned}&D_T\left[\int_0^T\int_\Sigma \mathrm{d}_x f_h\partial_x\varphi\,\mathrm{d}\xi\,\mathrm{d}t-\sum_{n=0}^{N_T}\sum_{i,j,k}\mathrm{d}_x\varphi_{i+1/2,j,k}^{n+1}\,\mathrm{d}_x f_{i+1/2,j,k}^{n+1}\Delta\xi\Delta t\right]\\ &\quad= D_T\sum_{n=1}^{N_T}\sum_{i,j,k}\left[\int_{t^{n-1}}^{t^n}\int_{C_{i+1/2,j,k}}\mathrm{d}_x f_h\partial_x\varphi\,\mathrm{d}\xi\,\mathrm{d}t-\mathrm{d}_x\varphi_{i+1/2,j,k}^n\,\mathrm{d}_x f_{i+1/2,j,k}^n\Delta\xi\Delta t\right]\\ &\quad= D_T\sum_{n=1}^{N_T}\sum_{i,j,k}\mathrm{d}_x f_{i+1/2,j,k}^n\left[\int_{t^{n-1}}^{t^n}\int_{C_{i+1/2,j,k}}\partial_x\varphi\,\mathrm{d}\xi\,\mathrm{d}t-\mathrm{d}_x\varphi_{i+1/2,j,k}^n\Delta\xi\Delta t\right]\end{aligned}\tag{4.55}$$

where we use again $\varphi_{i,j,k}^{N+1}=0$. Furthermore, we can write the continuum integrals for each cell as

$$\begin{aligned}&\int_{t^{n-1}}^{t^n}\int_{C_{i+1/2,j,k}} D_T\,\mathrm{d}_x f_h\partial_x\varphi\,\mathrm{d}\xi\,\mathrm{d}t\\ &\quad= D_T\,\mathrm{d}_x f_{i+1/2,j,k}^n\int_{t^{n-1}}^{t^n}\int_{y_{j-1/2}}^{y_{j+1/2}}\int_{\theta_{k-1/2}}^{\theta_{k+1/2}}[\varphi(t,x_{i+1},y,\theta)-\varphi(t,x_i,y,\theta)]\,\mathrm{d}y\,\mathrm{d}\theta\,\mathrm{d}t.\end{aligned}\tag{4.56}$$

Hence, comparing using the mean value theorem, we get

$$\begin{aligned}&\int_{t^{n-1}}^{t^n}\int_{y_{j-1/2}}^{y_{j+1/2}}\int_{\theta_{k-1/2}}^{\theta_{k+1/2}}[\varphi(t,x_{i+1},y,\theta)-\varphi(t,x_i,y,\theta)]\,\mathrm{d}y\,\mathrm{d}\,\theta\,\mathrm{d}t-\mathrm{d}_x\varphi_{i+1/2,j,k}^n\Delta\xi\Delta t\\ &\quad=\int_{y_{j-1/2}}^{y_{j+1/2}}\int_{\theta_{k-1/2}}^{\theta_{k+1/2}}\left[\int_{t^{n-1}}^{t^n}\frac{\varphi(t,x_{i+1},y,\theta)-\varphi(t,x_i,y,\theta)}{\Delta x}\,\mathrm{d}t\right.\\ &\qquad\left.-\frac{\Delta t}{\Delta x}\int_{x_{i-1/2}}^{x_{i+1/2}}\frac{\varphi(t^n,\Delta x+s,y,\theta)-\varphi(t^n,s,y,\theta)}{\Delta x}\,\mathrm{d}s\right]\Delta x\,\mathrm{d}y\,\mathrm{d}\theta\end{aligned}\tag{4.57}$$

$$= \int_{y_{j-1/2}}^{y_{j+1/2}} \int_{\theta_{k-1/2}}^{\theta_{k+1/2}} \left[\partial_x \varphi\big(\tilde{t}, \tilde{x}, y, \theta\big) - \frac{\varphi(t^n, \Delta x + \tilde{s}, y, \theta) - \varphi(t^n, \tilde{s}, y, \theta)}{\Delta x} \right] \Delta t \Delta x \,\mathrm{d}y \,\mathrm{d}\theta$$
$$= \int_{y_{j-1/2}}^{y_{j+1/2}} \int_{\theta_{k-1/2}}^{\theta_{k+1/2}} \big[\partial_x \varphi\big(\tilde{t}, \tilde{x}, y, \theta\big) - \partial_x \varphi(t^n, \hat{x}, y, \theta) \big] \Delta t \Delta x \,\mathrm{d}y \,\mathrm{d}\theta.$$

Then, taking absolute value and using the multi-variate mean value theorem we obtain

$$\begin{aligned} &\int_{y_{j-1/2}}^{y_{j+1/2}} \int_{\theta_{k-1/2}}^{\theta_{k+1/2}} \big| \partial_x \varphi\big(\tilde{t}, \tilde{x}, y, \theta\big) - \partial_x \varphi(t^n, \hat{x}, y, \theta) \big| \Delta t \Delta x \,\mathrm{d}y \,\mathrm{d}\theta \\ &\quad \le \int_{y_{j-1/2}}^{y_{j+1/2}} \int_{\theta_{k-1/2}}^{\theta_{k+1/2}} \Big(\|\partial_{tx}\varphi\|_{L^\infty(\Sigma_T)} + \big\|\partial_x^2 \varphi\big\|_{L^\infty(\Sigma_T)} \Big) (\Delta t + \Delta x) \Delta t \Delta x \,\mathrm{d}y \,\mathrm{d}\theta \\ &\quad \le \|\varphi\|_{C^2(\Sigma_T)} (\Delta t + \Delta x) \Delta \xi \Delta t. \end{aligned} \tag{4.58}$$

Hence, using this estimate and the uniform bound on $\|\mathrm{d}_\xi f_h\|_{L^2(\Sigma_T)}$ from Proposition 3.3 in equation (4.55), we obtain $\mathcal{D}(h) - \hat{\mathcal{D}}(h) \to 0$.

Then, for $\mathcal{P}(h) - \hat{\mathcal{P}}(h)$, the x-terms are treated as follows. The y-term can be treated in the same manner. The x-terms in $\mathcal{P}(h) - \hat{\mathcal{P}}(h)$ are

$$\begin{aligned} &- \int_0^T \int_\Sigma \mathrm{Pe} \cos_h(\theta) f_h \partial_x \varphi \,\mathrm{d}\xi \,\mathrm{d}t + \sum_{n=0}^{N_T} \sum_{i,j,k} \mathrm{Pe} \cos(\theta_k) f_{i,j,k}^{n+1} \,\mathrm{d}_x \varphi_{i+1/2,j,k}^{n+1} \Delta\xi \Delta t \\ &\quad = -\mathrm{Pe} \sum_{n=1}^{N_T} \sum_{i,j,k} \left[\int_{t^{n-1}}^{t^n} \int_{C_{i+1/2,j,k}} \cos(\theta_k) f_h \partial_x \varphi \,\mathrm{d}\xi \,\mathrm{d}t - \cos(\theta_k) f_{i,j,k}^n \,\mathrm{d}_x \varphi_{i+1/2,j,k}^n \Delta\xi \Delta t \right]. \end{aligned} \tag{4.59}$$

We can split the integral over $C_{i+1/2,j,k}$ in two terms

$$\begin{aligned} I_h &= \int_{C_{i+1/2,j,k}} \cos(\theta_k) f_h \partial_x \varphi \,\mathrm{d}\xi \\ &= \left[\int_{C_{i,j,k} \cap C_{i+1/2,j,k}} \cos(\theta_k) f_{i,j,k}^n \partial_x \varphi \,\mathrm{d}\xi + \int_{C_{i+1/2,j,k} \cap C_{i+1,j,k}} \cos(\theta_k) f_{i+1,j,k}^n \partial_x \varphi \,\mathrm{d}\xi \right], \end{aligned} \tag{4.60}$$

and putting it back together, we get

$$\begin{aligned} I_h = \int_{\theta_{k-1/2}}^{\theta_{k+1/2}} \int_{y_{j-1/2}}^{y_{j+1/2}} \cos(\theta_k) \Big[& f_{i,j,k}^n \big(\varphi\big(t, x_{i+1/2}, y, \theta\big) - \varphi(t, x_i, y, \theta) \big) \\ & + f_{i+1,j,k}^n \big(\varphi(t, x_{i+1}, y, \theta) - \varphi\big(t, x_{i+1/2}, y, \theta\big) \big) \Big] \,\mathrm{d}y \,\mathrm{d}\theta. \end{aligned} \tag{4.61}$$

Using these identities for equation (4.59) and rewriting, while adding and subtracting the term

$$\sum_{n=1}^{N_T} \sum_{i,j,k} \int_{t^{n-1}}^{t^n} \int_{\theta_{k-1/2}}^{\theta_{k+1/2}} \int_{y_{j-1/2}}^{y_{j+1/2}} \mathrm{Pe} \cos(\theta_k) f_{i,j,k}^n \big(\varphi(t, x_{i+1}, y, \theta) - \varphi\big(t, x_{i+1/2}, y, \theta\big) \big) \,\mathrm{d}y \,\mathrm{d}\theta \,\mathrm{d}t,$$

we get for the x-terms in $\mathcal{P}(h) - \hat{\mathcal{P}}(h)$

$$\begin{aligned}
&\sum_{n=1}^{N_T}\sum_{i,j,k} \mathrm{Pe}\cos(\theta_k) f_{i,j,k}^n \iiint_{C_{j,k}^n} -\big[\big(\varphi\big(t,x_{i+1/2},y,\theta\big)-\varphi(t,x_i,y,\theta)\big)-\big(\varphi_{i+1,j,k}^n-\varphi_{i,j,k}^n\big)\big]\,\mathrm{d}y\,\mathrm{d}\theta\,\mathrm{d}t\\
&\quad+\sum_{n=1}^{N_T}\sum_{i,j,k} \mathrm{Pe}\cos(\theta_k)\big(-f_{i+1,j,k}^n\big)\iiint_{C_{j,k}^n}\big(\varphi(t,x_{i+1},y,\theta)-\varphi\big(t,x_{i+1/2},y,\theta\big)\big)\,\mathrm{d}\theta\,\mathrm{d}y\,\mathrm{d}t\\
&\quad=-\sum_{n=1}^{N_T}\sum_{i,j,k} \mathrm{Pe}\cos(\theta_k) f_{i,j,k}^n \iiint_{C_{j,k}^n}\big[(\varphi(t,x_{i+1},y,\theta)-\varphi(t,x_i,y,\theta))-\big(\varphi_{i+1,j,k}^n-\varphi_{i,j,k}^n\big)\big]\,\mathrm{d}y\,\mathrm{d}\theta\,\mathrm{d}t\\
&\qquad-\sum_{n=1}^{N_T}\sum_{i,j,k} \mathrm{Pe}\cos(\theta_k)\big(f_{i+1,j,k}^n-f_{i,j,k}^n\big)\iiint_{C_{j,k}^n}\big(\varphi(t,x_{i+1},y,\theta)-\varphi\big(t,x_{i+1/2},y,\theta\big)\big)\,\mathrm{d}\theta\,\mathrm{d}y\,\mathrm{d}t\\
&\quad=\mathcal{I}_{\mathcal{P}}^1+\mathcal{I}_{\mathcal{P}}^2,
\end{aligned} \tag{4.62}$$

where $C_{j,k}^n = (t^{n-1}, t^n) \times (y_{j-1/2}, y_{j+1/2}) \times (\theta_{k-1/2}, \theta_{k+1/2})$.

The $\mathcal{I}_{\mathcal{P}}^1$ term converges to zero by using again the mean-value theorem and the finite-ness of the $C^2(\Sigma_T)$-norm of φ, as in equations (4.57) and (4.58).

The $\mathcal{I}_{\mathcal{P}}^2$ term can be upper bounded as, using the mean-value theorem and Cauchy–Schwarz,

$$\begin{aligned}
\mathcal{I}_{\mathcal{P}}^2 &\le \|\varphi\|_{C^1(\Sigma_T)}\sum_{n=1}^{N_T}\sum_{i,j,k}\mathrm{Pe}\cos(\theta_k)\big(f_{i+1,j,k}^n-f_{i,j,k}^n\big)\Delta\xi\Delta t\\
&\le C\|\varphi\|_{C^1(\Sigma_T)}\|\mathrm{d}_x f_h\|_{L^2(\Sigma_T)}\Delta x,
\end{aligned} \tag{4.63}$$

which goes to zero as $h \to 0$.

Finally, the convergence of $\mathcal{U}(h) - \hat{\mathcal{U}}(h)$ follows by a similar splitting argument as used for the $\mathcal{P}$-term. Indeed, by writing

$$\begin{aligned}
&\mathcal{U}(h)-\hat{\mathcal{U}}(h)\\
&\quad=\sum_{n=1}^{N_T}\sum_{i,j,k}\gamma B[c^n]_{i,j,k+1/2} f_{i,j,k}^n\iiint_{C_{i,j}^n}\big[(\varphi(t,x,y,\theta_{k+1})-\varphi(t,x,y,\theta_k))-\big(\varphi_{i,j,k+1}^n-\varphi_{i,j,k}^n\big)\big]\,\mathrm{d}x\,\mathrm{d}y\,\mathrm{d}t\\
&\qquad+\sum_{n=1}^{N_T}\sum_{i,j,k}\gamma B[c^n]_{i,j,k+1/2}\big(f_{i,j,k+1}^n-f_{i,j,k}^n\big)\iiint_{C_{i,j}^n}\big(\varphi(t,x,y,\theta_{k+1})-\varphi\big(t,x,y,\theta_{k+1/2}\big)\big)\,\mathrm{d}x\,\mathrm{d}y\,\mathrm{d}t,
\end{aligned} \tag{4.64}$$

where $C_{i,j}^n = (t^{n-1}, t^n) \times (x_{i-1/2}, x_{i+1/2}) \times (y_{j-1/2}, y_{j+1/2})$. Using the uniform control on $\|B[c_h]\|_{L^2(\Sigma_T)}$, the terms also go to zero by the same argument as before. This concludes the proof.

We note that by the uniqueness of the PDE solutions to equation (1.1a) for all B-terms, as shown in [8, 13], we get that the full sequence converges to the PDE solution. □

5. Higher regularity estimates and long-time estimates

In addition to the convergence result of Section 4, we establish a uniform-in-time estimate for the L^2 and L^∞ norms of the discrete solution to the scheme (2.11). For the L^2 norm, we can use the discrete analogue of the strategy used for the continuum problem in [13] for $B \in \{B_0, B_\lambda, B_\tau\}$. The key is to convert the negative gradient terms into negative L^2 terms using the Gagliardo–Nirenberg inequality. In contrast, we require a different strategy for the L^∞ norm since it is not clear what the analogue of the $W^{2,p}$-regularity theory for discrete finite volume functions is. Instead, we show a higher regularity result for the discrete H^1-norm of ρ_h, which allows us to use a discrete Morrey inequality ([30], Thm. 4.1) for the discrete gradient of c_h. We can cover the cases $B \in \{B_0, B_\lambda\}$.

5.1. L^∞ estimate for ρ_h and the discrete Alikakos lemma

Lemma 5.1 (Discrete Alikakos)**.** *Let $(F_k^n)_{n\in\mathbb{N}_0}$ be a family of nonnegative sequences for $k = 1, 2, \ldots,$ such that*

(i) $F_1^n = 1$, *for all $n \in \mathbb{N}_0$,*
(ii) $\lim_{k\to+\infty}(F_k^0)^{1/(\lambda_k+1)} = F_\infty^0 < \infty$ *exists,*
(iii) $(F_k^0)^{1/(\lambda_k+1)} \leq F_\infty^0$, *for all $k = 1, 2, \ldots,$*

where $\lambda_k = 2^k - 1$. Moreover, assume that there are constants $C > 0$ and $q, \beta \geq 1$, independent of k such that, for $k = 1, 2, \ldots,$

$$a_k = C2^k\big(2^k - 1\big), \qquad \epsilon_k = 2^{-qk}, \qquad c_k = 2^{\beta k}, \tag{5.1}$$

and

$$\frac{1}{\Delta t}\big(F_k^{n+1} - F_k^n\big) \leq -\epsilon_k F_k^{n+1} + (a_k + \epsilon_k)c_k\left(\sup_{n\in\mathbb{N}_0} F_{k-1}^n\right)^2. \tag{5.2}$$

Then, there exists a constant $A = A(\beta, q, C) \geq 1$ such that

$$\left(\sup_{n\in\mathbb{N}_0} F_k^n\right)^{1/(\lambda_k+1)} \leq AK, \tag{5.3}$$

for all $k = 1, 2, \ldots,$ where $K = \max\{1, F_\infty^0\}$.

Proof. The proof is a discrete analogue of the proof of Lemma 5.1 from [26], based on an induction argument over $k = 1, 2, \ldots$. First, by assumption, we have $\sup_{n\in\mathbb{N}_0} F_1^n \leq K$. For the induction step, let us assume that for some $M \in \mathbb{N}$, we have

$$\left(\sup_{n\in\mathbb{N}_0} F_k^n\right)^{1/(\lambda_k+1)} \leq A_k K, \qquad \text{for } k = 1, \ldots, M-1,$$

where $A_k > 0$ is to be determined. We note that $A_1 = 1$. Now, we can rewrite equation (5.2) as

$$F_k^{n+1} \leq \frac{F_k^n + \Delta t\epsilon_k\delta_k\big(\sup_n F_{k-1}^n\big)^2}{1 + \epsilon_k\Delta t}, \tag{5.4}$$

where we defined $\delta_k = \frac{a_k+\epsilon_k}{\epsilon_k}c_k$. By induction on n, this implies that

$$\sup_{n\in\mathbb{N}_0} F_k^n \leq \max\left\{F_k^0, \delta_k\left(\sup_{n\in\mathbb{N}_0} F_{k-1}^n\right)^2\right\}. \tag{5.5}$$

By assumption (iii) and the definition of K, we have $F_k^0 \leq K^{\lambda_k+1}$. Therefore, we can conclude that

$$\begin{aligned}\sup_{n\in\mathbb{N}_0} F_k^n &\leq \max\left\{K^{\lambda_k+1}, \delta_k\Big(\big(A_{k-1}K\big)^{\lambda_{k-1}+1}\Big)^2\right\} \leq \max\left\{K^{\lambda_k+1}, \delta_k A_{k-1}^{\lambda_k+1}K^{\lambda_k+1}\right\}\\ &\leq \left(\max\left\{1, \delta_k^{1/(\lambda_k+1)}A_{k-1}\right\}K\right)^{\lambda_k+1},\end{aligned} \tag{5.6}$$

having used the definition of λ_k. Therefore, it follows that, for $k \geq 2$,

$$A_k^{\lambda_k+1} = \delta_k\delta_{k-1}^2\ldots(\delta_2)^{2^{k-2}} = \prod_{\ell=2}^{k}\delta_\ell^{2^{k-\ell}}. \tag{5.7}$$

By the definition of δ_k and equation (5.1), we also note that

$$\delta_k \leq (1+C)2^{(2+q+\beta)k},$$

such that we can write for equation (5.7)

$$A_k^{\lambda_k+1} \leq \prod_{\ell=2}^{k}(1+C)^{2^{k-\ell}}2^{(2+q+\beta)\ell 2^{k-\ell}} = \prod_{\ell=0}^{k-2}(1+C)^{2^\ell}2^{(2+q+\beta)(k-\ell)2^\ell},$$

such that equation (5.6) becomes,

$$\sup_{n\in\mathbb{N}_0} F_k^n \leq K^{\lambda_k+1}\prod_{\ell=0}^{k-2}(1+C)^{2^\ell}2^{(2+q+\beta)(k-\ell)2^\ell}. \tag{5.8}$$

We now use the identities

$$\sum_{\ell=0}^{k-2}2^\ell = 2^{k-1}-1, \qquad \text{and} \qquad \sum_{\ell=0}^{k-2}\ell 2^\ell = 2+2^{k-1}(k-3),$$

to conclude that

$$\sup_{n\in\mathbb{N}_0} F_k^n \leq K^{2^k}(1+C)^{2^{k-1}-1}2^{(2+q+\beta)(3\cdot 2^{k-1}-k-2)} \leq K^{2^k}(1+C)^{2^k}2^{(2+q+\beta)2^{k+1}}. \tag{5.9}$$

Therefore

$$\left(\sup_{n\in\mathbb{N}_0} F_k^n\right)^{1/(\lambda_k+1)} \leq (1+C)2^{2(2+q+\beta)}K. \tag{5.10}$$

Hence, we have obtained the inequality of equation (5.3) with $A=(1+C)2^{2(2+q+\beta)}$. $\square$

Corollary 5.1. *Assume the initial data satisfies $\|\rho^0\|_{L^\infty} < +\infty$, and that the meshes are an admissibly regular family as in Definition 2.2. Then, the piecewise-constant approximation of the spatial density of the solution of the numerical scheme, equation (3.10a), satisfies*

$$\sup_{t\geq 0}\|\rho_h(t)\|_{L^{\lambda_k+1}} \leq A(\beta,q)\max\{1,\|\rho^0\|_{L^\infty}\}, \tag{5.11}$$

for $\lambda_k = 2^k-1$ and $k=1,2,\ldots$, and for some $C,\beta,q>0$, independent of the mesh size.

Proof. The idea of the proof is a discrete analogue of Proposition 3.5 from [13]. We first multiply equation (3.10a) by $(p+1)(\rho_{i,j}^{n+1})^p\mathbf{\Delta x}$ and sum over all (i,j) to obtain

$$\begin{aligned}(p+1)\sum_{i,j}\frac{\mathbf{\Delta x}}{\Delta t}\big(\rho_{i,j}^{n+1}-\rho_{i,j}^n\big)\big(\rho_{i,j}^{n+1}\big)^p &= (p+1)\sum_{i,j}D_T\mathbf{\Delta x}\big(\mathrm{d}_x^2\rho_{i,j}^{n+1}+\mathrm{d}_y^2\rho_{i,j}^{n+1}\big)\big(\rho_{i,j}^{n+1}\big)^p\\ &\quad -(p+1)\sum_{i,j}\mathrm{Pe}\mathbf{\Delta x}\Big[\mathrm{d}_x\bar{U}_{i,j}^{n+1}\big(\rho_{i,j}^{n+1}\big)^p+\mathrm{d}_y\bar{V}_{i,j}^{n+1}\big(\rho_{i,j}^{n+1}\big)^p\Big]\\ &=: I_{D_T}+I_{\mathrm{Pe}}.\end{aligned} \tag{5.12}$$

By summation by parts, we have

$$I_{\mathrm{Pe}} = \mathrm{Pe}(p+1)\sum_{i,j}\mathbf{\Delta x}\Big[\bar{U}_{i+1/2,j}^{n+1}\,\mathrm{d}_x\Big(\rho_{i+1/2,j}^{n+1}\Big)^p+\bar{V}_{i,j+1/2}^{n+1}\,\mathrm{d}_y\Big(\rho_{i,j+1/2}^{n+1}\Big)^p\Big]. \tag{5.13}$$

Using the elementary inequality

$$|b-a| \le \frac{2p}{p+1}\max\Big\{b^{\frac{p-1}{2p}}, a^{\frac{p-1}{2p}}\Big\}\Big|b^{\frac{p+1}{2p}} - a^{\frac{p+1}{2p}}\Big|,$$

for $a, b \in \mathbb{R}_{\ge 0}$, with $b = (\rho^{n+1}_{i+1,j})^p$, $a = (\rho^{n+1}_{i,j})^p$ in conjunction with $|\bar{U}^{n+1}_{i+1/2,j}| \le \max\{\rho^{n+1}_{i,j}, \rho^{n+1}_{i+1,j}\}$, we obtain

$$\Big|\bar{U}^{n+1}_{i+1/2,j}\,\mathrm{d}_x\Big(\rho^{n+1}_{i+1/2,j}\Big)^p\Big| \le \frac{2p}{p+1}\max\{\rho^{n+1}_{i,j}, \rho^{n+1}_{i+1,j}\}^{\frac{p+1}{2}}\Big|\mathrm{d}_x\Big(\rho^{n+1}_{i+1/2,j}\Big)^{\frac{p+1}{2}}\Big|.$$

Similarly, we get

$$\Big|\bar{V}^{n+1}_{i,j+1/2}\,\mathrm{d}_y\Big(\rho^{n+1}_{i,j+1/2}\Big)^p\Big| \le \frac{2p}{p+1}\max\{\rho^{n+1}_{i,j}, \rho^{n+1}_{i,j+1}\}^{\frac{p+1}{2}}\Big|\mathrm{d}_y\Big(\rho^{n+1}_{i,j+1/2}\Big)^{\frac{p+1}{2}}\Big|.$$

Returning to equation (5.12), we get for the x-part of I_{Pe}, denoted by $I_{\mathrm{Pe},x}$,

$$I_{\mathrm{Pe},x} \le \mathrm{Pe}(p+1)\sum_{i,j}\frac{2p}{p+1}\max\{\rho^{n+1}_{i,j}, \rho^{n+1}_{i+1,j}\}^{\frac{p+1}{2}}\Big|\mathrm{d}_x\Big(\rho^{n+1}_{i+1/2,j}\Big)^{\frac{p+1}{2}}\Big|\Delta\mathbf{x}. \tag{5.14}$$

Now, using Young–Hölder we have

$$\begin{aligned}\max\{\rho^{n+1}_{i,j}, \rho^{n+1}_{i+1,j}\}^{\frac{p+1}{2}}\frac{2\mathrm{Pe}}{p+1}\Big|\mathrm{d}_x\Big(\rho^{n+1}_{i+1/2,j}\Big)^{\frac{p+1}{2}}\Big| \\ \le \frac{1}{2\varepsilon}\max\{\rho^{n+1}_{i,j}, \rho^{n+1}_{i+1,j}\}^{p+1} + \frac{4\mathrm{Pe}^2}{(p+1)^2}\frac{\varepsilon}{2}\Big|\mathrm{d}_x\Big(\rho^{n+1}_{i+1/2,j}\Big)^{\frac{p+1}{2}}\Big|^2.\end{aligned}$$

Applying the above inequalities analogously to $I_{\mathrm{Pe},y}$, and using $(p+1)b^p(b-a) \ge (b^{p+1} - a^{p+1})$ for the time derivative,

$$\begin{aligned}&\frac{1}{\Delta t}\Big(\|\rho_h(t^{n+1})\|^{p+1}_{L^{p+1}} - \|\rho_h(t^n)\|^{p+1}_{L^{p+1}}\Big) \\ &\quad\le -(p+1)D_T\sum_{i,j}\Delta\mathbf{x}\Big(\mathrm{d}_x\rho^{n+1}_{i+1/2,j}\,\mathrm{d}_x\Big(\rho^{n+1}_{i+1/2,j}\Big)^p + \mathrm{d}_y\rho^{n+1}_{i,j+1/2}\,\mathrm{d}_y\Big(\rho^{n+1}_{i,j+1/2}\Big)^p\Big) \\ &\qquad + p(p+1)\sum_{i,j}\Delta\mathbf{x}\Bigg(\tfrac{1}{2\varepsilon}\max\{\rho^{n+1}_{i,j}, \rho^{n+1}_{i+1,j}\}^{p+1} + \frac{4\mathrm{Pe}^2\varepsilon}{2(p+1)^2}\Big|\mathrm{d}_x\Big(\rho^{n+1}_{i+1/2,j}\Big)^{\frac{p+1}{2}}\Big|^2\Bigg) \\ &\qquad + p(p+1)\sum_{i,j}\Delta\mathbf{x}\Bigg(\tfrac{1}{2\varepsilon}\max\{\rho^{n+1}_{i,j}, \rho^{n+1}_{i,j+1}\}^{p+1} + \frac{4\mathrm{Pe}^2\varepsilon}{2(p+1)^2}\Big|\mathrm{d}_y\Big(\rho^{n+1}_{i,j+1/2}\Big)^{\frac{p+1}{2}}\Big|^2\Bigg).\end{aligned} \tag{5.15}$$

Using the elementary inequality $-(p+1)(a-b)(a^p - b^p) \le -\frac{4p}{p+1}(a^{\frac{p+1}{2}} - b^{\frac{p+1}{2}})^2$, the first sum can be bounded as

$$I_{D_T} \le -\frac{4pD_T}{p+1}\Bigg(\Big\|\mathrm{d}_x(\rho_h(t^{n+1}))^{\frac{p+1}{2}}\Big\|^2_{L^2} + \Big\|\mathrm{d}_y(\rho_h(t^{n+1}))^{\frac{p+1}{2}}\Big\|^2_{L^2}\Bigg). \tag{5.16}$$

Combining the above and noting $\max\{a,b\}^p \le a^p + b^p$ for $a, b \in \mathbb{R}_{\ge 0}, p \ge 1$, we can write

$$\begin{aligned}&\frac{1}{\Delta t}\Big(\|\rho_h(t^{n+1})\|^{p+1}_{L^{p+1}} - \|\rho_h(t^n)\|^{p+1}_{L^{p+1}}\Big) \\ &\quad\le -\frac{4p}{p+1}\big(D_T - \tfrac{\varepsilon}{2}\mathrm{Pe}^2\big)\Big\|\mathrm{d}_{\mathbf{x}}(\rho_h)^{\frac{p+1}{2}}\Big\|^2_{L^2} + \tfrac{2}{\varepsilon}p(p+1)\|\rho_h(t^{n+1})\|^{p+1}_{L^{p+1}}.\end{aligned} \tag{5.17}$$

We now use the discrete Gagliardo–Nirenberg inequality ([9], Thm. 3.4) so that, for $u = \rho_h^{\frac{p+1}{2}}$,

$$\|u\|_{L^2}^2 \leq C_{GN}\|u\|_{L^1}\|u\|_{1,2}. \tag{5.18}$$

After some algebra (see Appendix C) for constants a_k, ϵ_k and c_k as in Lemma 5.1, we obtain

$$\frac{1}{\Delta t}\Big(\big\|\rho_h\big(t^{n+1}\big)\big\|_{L^{p+1}}^{p+1} - \|\rho_h(t^n)\|_{L^{p+1}}^{p+1}\Big) \leq -\epsilon_k\big\|\rho_h\big(t^{n+1}\big)\big\|_{L^{p+1}}^{p+1} + (a_k+\epsilon_k)c_k\bigg(\big\|\rho_h\big(t^{n+1}\big)\big\|_{L^{\frac{p+1}{2}}}^{\frac{p+1}{2}}\bigg)^2. \tag{5.19}$$

Therefore, by inserting the notation $\lambda_k = p = 2^k - 1$ we get the result by Lemma 5.1. □

5.2. L^∞ estimate for f_h

Proposition 5.1 (Uniform L^2-estimate for f_h)**.** *Assume the conditions in Corollary 5.1 are met. Then, any solution to the numerical scheme as in equation* (2.11) *satisfies*

$$\sup_{t\geq 0}\|f_h(t)\|_{L^2} < C, \tag{5.20}$$

for some constant $C>0$ that does not depend on the mesh size.

Proof. The idea is to revisit equation (3.40), *i.e.*,

$$\begin{aligned}\frac{1}{2\Delta t}\Big(\big\|f_h\big(t^{n+1}\big)\big\|_{L^2}^2 - \|f_h(t^n)\|_{L^2}^2\Big) &\leq -D_T\big\|\mathrm{d}_{\mathbf{x}} f_h\big(t^{n+1}\big)\big\|_{L^2}^2 - \big\|\mathrm{d}_\theta f_h\big(t^{n+1}\big)\big\|_{L^2}^2 \\ &\quad + \frac{\gamma}{2}\int\big|\mathrm{d}_\theta B\big[c_h\big(t^{n+1}\big)\big]\big|\,\big|f_h\big(t^{n+1}\big)\big|^2\,\mathrm{d}\xi,\end{aligned} \tag{5.21}$$

and use an improved version of the estimate of equation (3.41) to estimate the last term. Applying the discrete Gagliardo–Nirenberg inequality ([9], Thm. 3.4) we get for some constants $\epsilon, A(\epsilon) > 0$

$$\|f_h\|_{L^2}^2 \leq C_{\mathrm{GN}}^2\|f_h\|_{1,2}^{\frac{6}{5}}\|f_h\|_{L^1}^{\frac{4}{5}} \leq C_{\mathrm{GN}}^2\epsilon\|f_h\|_{1,2}^2 + A(\epsilon)\|f_h\|_{L^1}^2, \tag{5.22}$$

using Hölder conjugates $p=\frac{5}{3}$ and $q=\frac{5}{2}$. We rewrite this as

$$-C_{\mathrm{GN}}^2\epsilon\|\mathrm{d}_\xi f_h\|_{L^2}^2 \leq \big(C_{\mathrm{GN}}^2\epsilon - 1\big)\|f_h\|_{L^2}^2 + C_{\mathrm{GN}}^2 A(\epsilon)\|f_h\|_{L^1}^2. \tag{5.23}$$

Choosing ϵ such that $C_{\mathrm{GN}}^2\epsilon = \frac{1}{2}$ we get, using the conservation of unit mass,

$$-\|\mathrm{d}_\xi f_h\|_{L^2}^2 \leq -\|f_h\|_{L^2}^2 + C. \tag{5.24}$$

for some constant $C = 2C_{\mathrm{GN}}^2 A(\epsilon) > 0$. Therefore, we estimate setting $\kappa = \min(1, D_T)/2$

$$\begin{aligned}-D_T\|\mathrm{d}_{\mathbf{x}} f_h\|_{L^2}^2 - \|\mathrm{d}_\theta f_h\|_{L^2}^2 &\leq -\kappa\|\mathrm{d}_\xi f_h\|_{L^2}^2 \\ &\leq -\tfrac{1}{2}\kappa\|f_h\|_{L^2}^2 + \tfrac{1}{2}C\kappa - \tfrac{1}{2}\kappa\|\mathrm{d}_\xi f_h\|_{L^2}^2.\end{aligned}$$

Applying this to the first two terms in equation (5.21) we obtain

$$\begin{aligned}\frac{1}{2\Delta t}\Big(\big\|f_h\big(t^{n+1}\big)\big\|_{L^2}^2 - \|f_h(t^n)\|_{L^2}^2\Big) &\leq C - \tfrac{1}{2}\kappa\big\|f_h\big(t^{n+1}\big)\big\|_{1,2}^2 \\ &\quad + \frac{\gamma}{2}\int\big|\mathrm{d}_\theta B\big[c_h\big(t^{n+1}\big)\big]\big|\,\big|f_h\big(t^{n+1}\big)\big|^2\,\mathrm{d}\xi.\end{aligned} \tag{5.25}$$

Next, using the Hölder inequality and then an interpolation of Lebesgue spaces, we observe

$$
\begin{aligned}
\frac{\gamma}{2}\int |\mathrm{d}_\theta B[c_h]|\,\big|f_h(t^{n+1})\big|^2\,\mathrm{d}\xi &\le C\|\mathrm{d}_\theta B[c_h]\|_{L^2}\big\|f_h(t^{n+1})\big\|_{L^4}^2\\
&\le C\big\|f_h(t^{n+1})\big\|_{L^1}^{\frac{2}{10}}\big\|f_h(t^{n+1})\big\|_{L^6}^{\frac{18}{10}}\\
&\le C\big\|f_h(t^{n+1})\big\|_{L^6}^{\frac{9}{5}},
\end{aligned}
\tag{5.26}
$$

having also used mass conservation, Proposition 3.2 and Corollary 5.1. Here, the constant C changes from line to line but remains independent of the mesh size. Then, equation (5.25) becomes

$$
\frac{1}{2\Delta t}\Big(\big\|f_h(t^{n+1})\big\|_{L^2}^2-\|f_h(t^n)\|_{L^2}^2\Big)\le C-\tfrac{1}{2}\kappa\big\|f_h(t^{n+1})\big\|_{1,2}^2+C\big\|f_h(t^{n+1})\big\|_{L^6}^{\frac{9}{5}}. \tag{5.27}
$$

Next, we can apply the discrete Poincaré–Sobolev inequality

$$
\|f_h\|_{L^6}\le C_{\mathrm{PS}}\|f_h\|_{1,2} \tag{5.28}
$$

cf. Theorem 3.2 of [9], to estimate the L^6-norm in equation (5.27). Next, we use product inequalities to obtain for $\epsilon>0$ and some constant $A(\epsilon)>0$,

$$
C\|f_h\|_{L^6}^{\frac{9}{5}}\le CC_{\mathrm{PS}}^{\frac{9}{5}}\|f_h\|_{1,2}^{\frac{9}{5}}\le\epsilon\|f_h\|_{1,2}^2+A(\epsilon). \tag{5.29}
$$

This way, choosing $\epsilon<\kappa/2$, we obtain

$$
\frac{1}{\Delta t}\Big(\big\|f_h(t^{n+1})\big\|_{L^2}^2-\|f_h(t^n)\|_{L^2}^2\Big)\le-\big(\tfrac{\kappa}{2}-\epsilon\big)\big\|f_h(t^{n+1})\big\|_{L^2}^2+C(\kappa,\epsilon), \tag{5.30}
$$

for positive constant $C(\kappa,\epsilon)>0$ that does not depend on the mesh size. Hence, we find

$$
\sup_{n\in\mathbb{N}_0}\|f_h(t^n)\|_{L^2}^2\le C, \tag{5.31}
$$

where $C>0$ only depends on α and $\|f_0\|_{L^2}$ but is independent of the mesh size. Then, the result follows. □

Proposition 5.2 (L^2-estimate for $\mathrm{d}_{\mathbf{x}}\rho_h$)**.** *Given $\|\nabla_{\mathbf{x}}\rho^0\|_{L^2(\Omega)}<+\infty$ and the same assumptions of Proposition 5.1, the solution to the numerical scheme* (2.11) *satisfies*

$$
\sup_{t\ge 0}\|\mathrm{d}_{\mathbf{x}}\rho_h(t)\|_{L^2(\Omega)}\le C, \tag{5.32}
$$

where $C>0$ is independent of the mesh size.

Proof. The idea is to obtain a discrete analogue of the regularity result for linear parabolic equations in Section 7.1.3 of [22]. Recalling equation (3.10a), we know that $\rho_{i,j}^n$ satisfies

$$
\frac{\rho_{i,j}^{n+1}-\rho_{i,j}^n}{\Delta t}=D_T\big(\mathrm{d}_x^2\rho_{i,j}^{n+1}+\mathrm{d}_y^2\rho_{i,j}^{n+1}\big)-\mathrm{Pe}\,\mathrm{d}_x\bar{U}_{i,j}^{n+1}-\mathrm{Pe}\,\mathrm{d}_y\bar{V}_{i,j}^{n+1}, \tag{5.33}
$$

where the $\bar{U}$ and $\bar{V}$ terms are as before, cf. equation (3.10b).

Next, we move the discrete Laplacian on the left-hand side, square both sides, multiply by $\Delta\xi$, and sum over all i,j. This yields

$$
\sum_{i,j}\Delta\xi\left(\left|\frac{\rho_{i,j}^{n+1}-\rho_{i,j}^n}{\Delta t}\right|^2-2D_T\frac{\rho_{i,j}^{n+1}-\rho_{i,j}^n}{\Delta t}\big(\mathrm{d}_x^2\rho_{i,j}^{n+1}+\mathrm{d}_y^2\rho_{i,j}^{n+1}\big)+D_T^2\big|\mathrm{d}_x^2\rho_{i,j}^{n+1}+\mathrm{d}_y^2\rho_{i,j}^{n+1}\big|^2\right)
$$

$$= \mathrm{Pe}^2 \sum_{i,j} \Delta\xi \Big(\big|\mathrm{d}_x \bar{U}_{i,j}^{n+1}\big|^2 + 2\mathrm{d}_x \bar{U}_{i,j}^{n+1}\, \mathrm{d}_y \bar{V}_{i,j}^{n+1} + \big|\mathrm{d}_y \bar{V}_{i,j}^{n+1}\big|^2 \Big). \tag{5.34}$$

We observe that all but one term on the left-hand side are squares. Concerning the cross term, we obtain by a summation by parts

$$\begin{aligned} -2D_T \sum_{i,j} \Delta\xi \frac{\rho_{i,j}^{n+1} - \rho_{i,j}^n}{\Delta t} \big(\mathrm{d}_x^2 \rho_{i,j}^{n+1} + \mathrm{d}_y^2 \rho_{i,j}^{n+1}\big) \\ = 2D_T \sum_{i,j} \Delta\xi \left(\frac{\mathrm{d}_x \rho_{i+1/2,j}^{n+1} - \mathrm{d}_x \rho_{i+1/2,j}^n}{\Delta t} \mathrm{d}_x \rho_{i+1/2,j}^{n+1} + \frac{\mathrm{d}_y \rho_{i,j+1/2}^{n+1} - \mathrm{d}_y \rho_{i,j+1/2}^n}{\Delta t} \mathrm{d}_y \rho_{i,j+1/2}^{n+1} \right) \\ \geq D_T \sum_{i,j} \Delta\xi \left(\frac{\big|\mathrm{d}_x \rho_{i+1/2,j}^{n+1}\big|^2 - \big|\mathrm{d}_x \rho_{i+1/2,j}^n\big|^2}{\Delta t} + \frac{\big|\mathrm{d}_y \rho_{i,j+1/2}^{n+1}\big|^2 - \big|\mathrm{d}_y \rho_{i,j+1/2}^n\big|^2}{\Delta t} \right), \end{aligned} \tag{5.35}$$

where we used the inequality $(a-b)a \geq \frac{1}{2}(a^2 - b^2)$ in the last line. Substituting equation (5.35) into equation (5.34), we obtain

$$\begin{aligned} D_T \sum_{i,j} \Delta\xi \left(\frac{\big|\mathrm{d}_x \rho_{i+1/2,j}^{n+1}\big|^2 - \big|\mathrm{d}_x \rho_{i+1/2,j}^n\big|^2}{\Delta t} + \frac{\big|\mathrm{d}_y \rho_{i,j+1/2}^{n+1}\big|^2 - \big|\mathrm{d}_y \rho_{i,j+1/2}^n\big|^2}{\Delta t} \right) \\ \leq \mathrm{Pe}^2 \sum_{i,j} \Delta\xi \Big(\big|\mathrm{d}_x \bar{U}_{i,j}^{n+1}\big|^2 + 2\mathrm{d}_x \bar{U}_{i,j}^{n+1}\, \mathrm{d}_y \bar{V}_{i,j}^{n+1} + \big|\mathrm{d}_y \bar{V}_{i,j}^{n+1}\big|^2 \Big) \\ \leq C \sum_{i,j,k} \Delta\xi |\mathrm{d}_{\mathbf{x}} f_{i,j,k}^{n+1}|^2, \end{aligned} \tag{5.36}$$

where we used Jensen's inequality and Cauchy–Schwarz, to estimate $|\mathrm{d}_x \bar{U}_{i,j}^{n+1}| \leq |\mathrm{d}_x f_{i-1/2,j}^{n+1}| + |\mathrm{d}_x f_{i+1/2,j}^{n+1}|$ and $\mathrm{d}_y \bar{V}$, respectively. In summary, we have, for some $C > 0$,

$$\sum_{i,j} \Delta\xi \frac{\big|\mathrm{d}_{\mathbf{x}} \rho_{i,j}^{n+1}\big|^2 - \big|\mathrm{d}_{\mathbf{x}} \rho_{i,j}^n\big|^2}{\Delta t} \leq C \sum_{i,j,k} \Delta\xi \Big|\mathrm{d}_{\mathbf{x}} f_{i,j,k}^{n+1}\Big|^2. \tag{5.37}$$

Fix $t > 0$ and let $m = \lfloor t/\Delta t \rfloor$, $M = \lfloor (t+1)/\Delta t \rfloor$ and let $m \leq s \leq M$. Then, multiplying (5.37) by Δt and summing n from s to M, we get

$$\begin{aligned} \big\|\mathrm{d}_{\mathbf{x}} \rho\big(t^{m+1}\big)\big\|_{L^2(\Omega)}^2 &\leq \|\mathrm{d}_{\mathbf{x}} \rho(t^s)\|_{L^2(\Omega)}^2 + C \sum_{n=s}^{M} \Delta t \big\|\mathrm{d}_{\mathbf{x}} f\big(t^{n+1}\big)\big\|_{L^2(\Omega)}^2 \\ &\leq \|\mathrm{d}_{\mathbf{x}} \rho(t^s)\|_{L^2(\Omega)}^2 + C \sum_{n=m}^{M} \Delta t \big\|\mathrm{d}_{\mathbf{x}} f\big(t^{n+1}\big)\big\|_{L^2(\Omega)}^2. \end{aligned} \tag{5.38}$$

We note that by Inequality (3.42) in the proof of Proposition 3.3, we have

$$\begin{aligned} \big\|\mathrm{d}_\xi f_h\big(t^{n+1}\big)\big\|_{L^2}^2 &\leq \frac{C(\varepsilon)}{D(\varepsilon)} \big\|\mathrm{d}_\theta B\big[c_h\big(t^{n+1}\big)\big]\big\|_{L^2}^2 \big\|f_h\big(t^{n+1}\big)\big\|_{L^2}^2 + \frac{1}{2\Delta t D(\varepsilon)} \Big(\|f_h(t^n)\|_{L^2}^2 - \big\|f_h\big(t^{n+1}\big)\big\|_{L^2}^2 \Big) \\ &\leq C \bigg(\big\|f_h\big(t^{n+1}\big)\big\|_{L^2}^2 + \frac{1}{\Delta t} \Big(\|f_h(t^n)\|_{L^2}^2 - \big\|f_h\big(t^{n+1}\big)\big\|_{L^2}^2 \Big) \bigg), \end{aligned} \tag{5.39}$$

for some $C > 0$. Here, we used the uniform bound $\sup_{t\geq 0}\|\mathrm{d}_\theta B[c_h](t)\|_{L^2} \leq C$ for some $C > 0$, which follows from Propositions 3.2 and 5.1. Then, multiplying (5.39) by Δt and summing from $n = m$ to M, we have, for some $C > 0$ independent of the time and mesh size,

$$\sum_{n=m}^{M} \Delta t \big\|\mathrm{d}_{\mathbf{x}} f_h\big(t^{n+1}\big)\big\|_{L^2}^2 \leq C\left(\sum_{n=m}^{M} \big\|f_h\big(t^{n+1}\big)\big\|_{L^2}^2 \Delta t + \|f_h(t^m)\|_{L^2}^2\right) \leq C. \tag{5.40}$$

Here, we made use of the uniform bound $\sup_{t\geq 0}\|f_h(t)\|_{L^2}^2$ from Proposition 5.1 and the inequality $(M+1-m)\Delta t \leq 2$ (which we have by construction and for $\Delta t < 1$).

Hence, multiplying (5.38) by Δt and summing s from m to M, we get, using $\frac{1}{2} \leq \sum_{n=m}^{M}\Delta t \leq 2$, for some $C > 0$

$$\big\|\mathrm{d}_{\mathbf{x}}\rho\big(t^{m+1}\big)\big\|_{L^2(\Omega)}^2 \leq C\left(1 + \sum_{n=m}^{M} \Delta t \|\mathrm{d}_{\mathbf{x}}\rho(t^n)\|_{L^2(\Omega)}^2\right). \tag{5.41}$$

Finally, using $\|\mathrm{d}_{\mathbf{x}}\rho(t^n)\|_{L^2(\Omega)} \leq C\|\mathrm{d}_{\mathbf{x}} f(t^n)\|_{L^2}$, for some constant $C > 0$, we can bound the right hand side of (5.41) *via* (5.40). So, we may conclude, for some $C > 0$ independent of m and the mesh size,

$$\big\|\mathrm{d}_{\mathbf{x}}\rho\big(t^{m+1}\big)\big\|_{L^2(\Omega)}^2 \leq C. \tag{5.42}$$

Hence, the result follows for any $t \geq 0$. □

Proposition 5.3 (L^∞-estimate on $\mathrm{d}_{\mathbf{x}}c_h$)**.** *With the same assumptions as in Corollary 5.1 and Proposition 5.2, the solution to the numerical scheme as in equation* (2.11) *satisfies*

$$\sup_{t\geq 0}\|\mathrm{d}_{\mathbf{x}}c_h(t)\|_{L^\infty(\Omega)} \leq C, \tag{5.43}$$

for a constant $C > 0$ that does not depend on the mesh size.

Proof. We first repeat the same elliptic L^2-estimate as in equation (3.28) for the gradient of c_h, by applying d_x and d_y to (2.11e). First, applying d_x to (2.11e) gives

$$-\mathrm{d}_x\big(\mathrm{d}_x^2 c + \mathrm{d}_y^2 c\big)_{i+1/2,j} = -\alpha \mathrm{d}_x c_{i+1/2,j} + \mathrm{d}_x\rho_{i+1/2,j}, \tag{5.44}$$

such that

$$\begin{aligned}\mathrm{d}_x\,\mathrm{d}_x^2 c_{i+1/2,j} &= \frac{1}{\Delta x}\big(\mathrm{d}_x^2 c_{i+1,j} - \mathrm{d}_x^2 c_{i,j}\big)\\ &= \frac{1}{\Delta x^3}\big(c_{i+2,j} - 2c_{i+1,j} + c_{i,j} - c_{i+1,j} + 2c_{i,j} - c_{i-1,j}\big)\\ &= \frac{1}{\Delta x^2}\big(\mathrm{d}_x c_{i+3/2,j} - 2\mathrm{d}_x c_{i+1/2,j} + \mathrm{d}_x c_{i-1/2,j}\big)\\ &= \mathrm{d}_x^2\mathrm{d}_x c_{i+1/2,j}.\end{aligned} \tag{5.45}$$

Squaring both sides of (5.44), multiplying by $\Delta\mathbf{x}$ and summing over (i,j), then leads to

$$\begin{aligned}\sum_{i,j}\Big|\mathrm{d}_x\big(\mathrm{d}_x^2 c + \mathrm{d}_y^2 c\big)_{i+1/2,j}\Big|^2 \Delta\mathbf{x} &\leq C\sum_{i,j}\Big[\alpha^2\big|\mathrm{d}_x c_{i+1/2,j}\big|^2 + \big|\mathrm{d}_x\rho_{i+1/2,j}\big|^2\Big]\Delta\mathbf{x}\\ &\leq C\Big(\|\mathrm{d}_{\mathbf{x}}c_h\|_{L^2(\Omega)}^2 + \|\mathrm{d}_{\mathbf{x}}\rho_h\|_{L^2(\Omega)}^2\Big)\\ &\leq C,\end{aligned} \tag{5.46}$$

by Propositions 5.2 and 3.2, for some $C > 0$ independent of the mesh size. Similarly, we obtain

$$\sum_{i,j} \left| \mathrm{d}_y \left(\mathrm{d}_x^2 c + \mathrm{d}_y^2 c \right)_{i,j+1/2} \right|^2 \Delta\mathbf{x} \leq C. \tag{5.47}$$

Expanding the left-hand side of equation (5.46), an integration by parts yields

$$\sum_{i,j} \Big(\left| \mathrm{d}_x\, \mathrm{d}_x^2 c_{i+1/2,j} \right|^2 + 2 \left| \mathrm{d}_x\, \mathrm{d}_{xy} c_{i,j+1/2} \right|^2 + \left| \mathrm{d}_x\, \mathrm{d}_y^2 c_{i+1/2,j} \right|^2 \Big) \Delta\mathbf{x} \leq C, \tag{5.48a}$$

and similarly

$$\sum_{i,j} \Big(\left| \mathrm{d}_y\, \mathrm{d}_x^2 c_{i,j+1/2} \right|^2 + 2 \left| \mathrm{d}_y\, \mathrm{d}_{xy} c_{i+1/2,j} \right|^2 + \left| \mathrm{d}_y\, \mathrm{d}_y^2 c_{i,j+1/2} \right|^2 \Big) \Delta\mathbf{x} \leq C. \tag{5.48b}$$

This implies that we have componentwise control over the $\|\cdot\|_{1,2}$-norm of the Hessian $\mathrm{d}_{\mathbf{x}}^2 c_h$. Now, because of the discrete Poincaré–Sobolev inequality ([9], Thm. 3.2), we have, in dimension two,

$$\begin{aligned} &\left\| \mathrm{d}_x^2 c_h \right\|_{L^3(\Omega)} + \left\| \mathrm{d}_{xy} c_h \right\|_{L^3(\Omega)} + \left\| \mathrm{d}_y^2 c_h \right\|_{L^3(\Omega)} \\ &\qquad \leq C_{\mathrm{PS}} \Big(\left\| \mathrm{d}_x^2 c_h \right\|_{1,2,\Omega} + \left\| \mathrm{d}_{xy} c_h \right\|_{1,2,\Omega} + \left\| \mathrm{d}_y^2 c_h \right\|_{1,2,\Omega} \Big) \\ &\qquad \leq C, \end{aligned} \tag{5.49}$$

where we used equation (5.48), that is

$$|\mathrm{d}_{\mathbf{x}} c_h|_{1,3,\Omega} < C,$$

for some constant $C > 0$ independent of the mesh size. Therefore, by the discrete Morrey inequality ([30], Thm. 4.1)

$$\|\mathrm{d}_{\mathbf{x}} c_h\|_{L^\infty(\Omega)} \leq C_M |\mathrm{d}_{\mathbf{x}} c_h|_{1,3}^{\frac{3}{4}} \|\mathrm{d}_{\mathbf{x}} c_h\|_{L^2(\Omega)}^{\frac{1}{4}}. \tag{5.50}$$

Now, the right-hand side can be uniformly bounded using equation (5.48) as well as Proposition 5.2, which concludes the proof. ☐

Corollary 5.2 (L^∞-estimate for f_h)**.** *Provided the same assumptions as for Proposition 5.3 hold, the solution to the numerical scheme as in equation* (2.11)*, for the interaction terms $B \in \{B_0, B_\lambda\}$, satisfies*

$$\sup_{t \geq 0} \|f_h(t)\|_{L^\infty} \leq C, \tag{5.51}$$

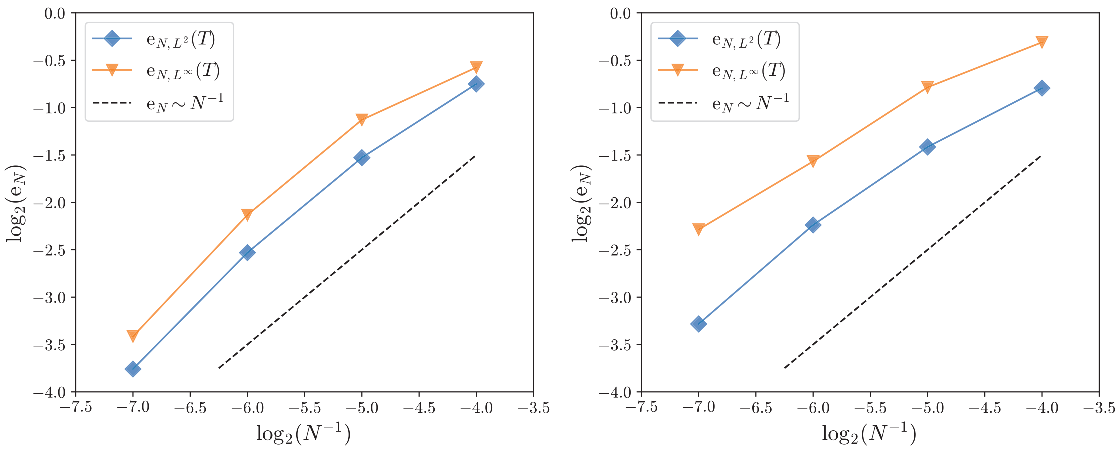
for some constant C independent of the mesh size.

Proof. The idea of the proof is a discrete analogue of Proposition 3.5 from [13]. We multiply equation (2.11) by $(p+1)(f_{i,j,k}^{n+1})^p$. After summation by parts and using Young's product inequality, we get, for some $\varepsilon > 0$,

$$\begin{aligned} &\frac{1}{\Delta t} \Big(\|f_h(t^{n+1})\|_{L^{p+1}}^{p+1} - \|f_h(t^n)\|_{L^{p+1}}^{p+1} \Big) \\ &\quad \leq -\frac{4p}{p+1} \Big(\min\{D_T, 1\} - \gamma^2 C_c \frac{\varepsilon}{2} \Big) \left\| \mathrm{d}_\xi \left[f_h \left(t^{n+1}\right)^{\frac{p+1}{2}} \right] \right\|_{L^2}^2 + \frac{1}{\varepsilon} p(p+1) \left\| f_h \left(t^{n+1}\right) \right\|_{L^{p+1}}^{p+1} \\ &\quad \leq -\frac{4p}{p+1} C \left\| \mathrm{d}_\xi \left[f_h \left(t^{n+1}\right)^{\frac{p+1}{2}} \right] \right\|_{L^2}^2 + \frac{1}{\varepsilon} p(p+1) \left\| f_h \left(t^{n+1}\right) \right\|_{L^{p+1}}^{p+1}, \end{aligned} \tag{5.52}$$

where C_c comes from Proposition 5.3, and $C > 0$ is some constant. Then, using the discrete Gagliardo–Nirenberg inequality ([9], Thm. 3.4) for dimension three

$$\|u\|_{L^2} \leq C_{GN} \|u\|_{L^1}^{\frac{2}{5}} \|u\|_{1,2}^{\frac{3}{5}}, \tag{5.53}$$



(a) B_0 (b) B_λ

FIGURE 1. Relative error (6.1) for numerical solutions to (2.11) with varying mesh size, for (a) B_0 interaction (2.11g) and (b) B_λ (2.11i) with $\lambda = 0.1$. The initial condition is $f^0(x,\theta) = C\mathbf{1}_{|x\pm\frac{1}{8}|\leq\frac{1}{8}}$ such that $\int f^0 \mathrm{d}x\, \mathrm{d}\theta = 1$. The other parameters are set as $T = 1.0, D_T = 10^{-1}, \gamma = 500, \mathrm{Pe} = 2, \alpha = 1, \Delta t = 10^{-2}$.

we can estimate

$$\frac{1}{\Delta t}\left(\left\|f_h\left(t^{n+1}\right)\right\|_{L^{p+1}}^{p+1} - \|f_h(t^n)\|_{L^{p+1}}^{p+1}\right) \leq -\epsilon_k\left\|f_h\left(t^{n+1}\right)\right\|_{L^{p+1}}^{p+1} + (a_k+\epsilon_k)c_k\left(\left\|f_h\left(t^{n+1}\right)\right\|_{L^{\frac{p+1}{2}}}^{\frac{p+1}{2}}\right)^2, \tag{5.54}$$

where ϵ_k, a_k, c_k are as in Lemma 5.1 (see the appendix for the algebra). Hence, the result follows by an application of Lemma 5.1. □

We note that this L^∞-estimate works for $B \in \{B_0, B_\lambda\}$. For $B = B_\tau$, this would require a uniform estimate on Hc. This would require an L^3-elliptic estimate for $\nabla^3 c$, and similarly an L^2-parabolic estimate on $\nabla^2\rho$.

6. NUMERICAL RESULTS

In this section, we test the numerical scheme analysed in the previous sections. Since we do not have access to explicit solutions, we test the convergence of the numerical scheme (2.11) by comparing solutions with increasingly finer meshes. We choose y-invariant initial data so that, given the form of the equations and boundary conditions of System (1.1), solutions remain y-invariant for $t > 0$, and we may use a y-invariant version of the numerical scheme (corresponding to setting $N_y = 1$ in (2.3)). We also take $N_x = N_\theta = N$. The numerical scheme is implemented in `Julia` and can be accessed in [20].

We numerically test the order of convergence of the scheme (2.11) by plotting the relative error of the approximation f^N with N grid points in space at time t as

$$e_{N,L}(t) = \|f_N(t) - f_{N^*}(t)\|_L / \|f_{N^*}(t)\|_L, \tag{6.1}$$

for any norm L, with respect to the finest mesh with $N^* = 256$. Figure 1a shows the relative errors in L^2 and L^∞ at the final time T for the interaction term B_0. This shows that the scheme is of order one in space, as can be expected for a linear finite volume discretisation. Since we are using an order one discretisation in time, we can expect that the full scheme is also of order one. In Figure 1b, the relative error for the interaction term B_λ in equation (2.11i) with $\lambda = 0.1$ is shown, again with order one convergence as expected.

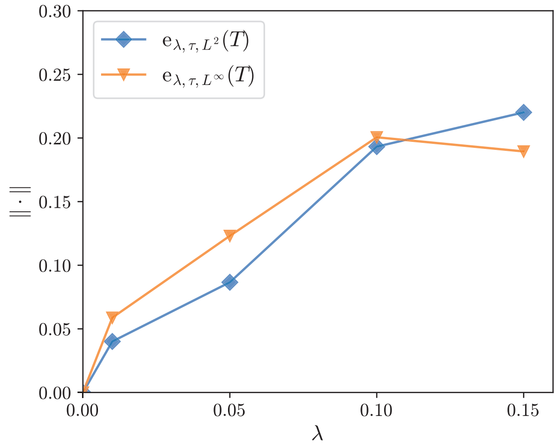


FIGURE 2. Relative norm of the difference between numerical solutions for B_λ and B_τ, $\lambda = \tau$, with $D_T = 10^{-1}, \gamma = 500, \mathrm{Pe} = 2, N_x = N_\theta = N = 64, \Delta t = 10^{-2}, T = 1.0$. The initial condition is $f^0(x,\theta) = C\mathbf{1}_{|x\pm\frac{1}{8}|\leq\frac{1}{8}}$, with $C > 0$ such that $\int f \mathrm{d}x\, \mathrm{d}\theta = 1$.

Finally, in Figure 2, for a fixed mesh, we compare the numerical solutions for B_λ and B_τ such that $\tau = \lambda$ varies. We use the notation

$$\mathrm{e}_{\lambda,\tau,L}(t) = \|f_{N,\lambda}(t) - f_{N,\tau}(t)\|_L / \|f_{N,\lambda}(t)\|_L. \tag{6.2}$$

FUNDING

The first author is supported by a Royal Society University Research Fellowship (URF/R1/180040). The third author is supported by the Cambridge Trust.

DATA AVAILABILITY STATEMENT

The research code associated with this article are available in numerics_ants, under the reference https://github.com/odewit8/numerics_ants [20].

INFORMED CONSENT

REFERENCES

[1] D. Albritton and L. Ohm, On the stabilizing effect of swimming in an active suspension. *SIAM J. Math. Anal.* **55** (2023) 6093–6132.
[2] R. Alert, J. Casademunt and J.-F. Joanny, Active turbulence. *Ann. Rev. Condens. Matter Phys.* **13** (2022) 143–170.
[3] K.A. Bacik, B.S. Bacik and T. Rogers, Lane nucleation in complex active flows. *Science* **379** (2023) 923–928.
[4] R. Bailo, J.A. Carrillo, H. Murakawa and M. Schmidtchen, Convergence of a fully discrete and energy-dissipating finite-volume scheme for aggregation-diffusion equations. *Math. Models Methods Appl. Sci.* **30** (2020) 2487–2522.
[5] N. Bellomo, P. Degond and E. Tadmor, editors, Active Particles, Volume 1. Advances in Theory, Models, and Applications. *Model. Simul. Sci. Eng. Technol.* Birkhäuser/Springer, Basel (2017).
[6] N. Bellomo, P. Degond and E. Tadmor, editors, Active Particles, Volume 2. *Model. Simul. Sci. Eng. Technol.* Birkhäuser, Cham (2019).

[7] N. Bellomo, J.A. Carrillo and E. Tadmor, editors, Active Particles. Volume 3. Advances in Theory, Models, and Applications. *Model. Simul. Sci. Eng. Technol.* Birkhäuser, Cham (2022).

[8] C. Bertucci, M. Rakotomalala and M. Tomasevic, Curvature in chemotaxis: a model for ant trail pattern formation. Preprint arXiv:2408.13363 (2024).

[9] M. Bessemoulin-Chatard, C. Chainais-Hillairet and F. Filbet, On discrete functional inequalities for some finite volume schemes. *IMA J. Numer. Anal.* **35** (2015) 1125–1149.

[10] H. Brezis, Functional Analysis, Sobolev Spaces and Partial Differential Equations, 1 edition. Springer, New York, NY, USA (2010).

[11] M. Briant, A. Diez and S. Merino-Aceituno, Cauchy theory for general kinetic Vicsek models in collective dynamics and mean-field limit approximations. *SIAM J. Math. Anal.* **54** (2022) 1131–1168.

[12] M. Bruna, M. Burger, A. Esposito and S.M. Schulz, Phase separation in systems of interacting active Brownian particles. *SIAM J. Appl. Math.* **82** (2022) 1635–1660.

[13] M. Bruna, M. Burger and O. de Wit, Lane formation and aggregation spots in a model of ants. *SIAM J. Appl. Dyn. Syst.* **24** (2025) 675–709.

[14] M. Burger and S. Schulz, Well-posedness and stationary states for a crowded active Brownian system with size-exclusion. Preprint arXiv:2309.17326 (2023).

[15] M. Burger, S. Hittmeir, H. Ranetbauer and M.-T. Wolfram, Lane formation by side-stepping. *SIAM J. Math. Anal.* **48** (2016) 981–1005.

[16] J.A. Carrillo, Y. Huang and M. Schmidtchen, Zoology of a nonlocal cross-diffusion model for two species. *SIAM J. Appl. Math.* **78** (2018) 1078–1104.

[17] J.A. Carrillo, R.S. Gvalani and J.S.-H. Wu, An invariance principle for gradient flows in the space of probability measures. *J. Differ. Equ.* **345** (2023) 233–284.

[18] C. Chainais-Hillairet and F. Filbet, Asymptotic behaviour of a finite-volume scheme for the transient drift-diffusion model. *IMA J. Numer. Anal.* **27** (2007) 689–716.

[19] C. Chainais-Hillairet, J.-G. Liu and Y.-J. Peng, Finite volume scheme for multi-dimensional drift-diffusion equations and convergence analysis. *ESAIM: Math. Model. Numer. Anal.* **37** (2003) 319–338.

[20] O. de Wit, Numerics ants, Last accessed 25 March 2025. https://github.com/odewit8/numerics_ants (2025).

[21] P. Degond, A. Frouvelle and J.-G. Liu, Macroscopic limits and phase transition in a system of self-propelled particles. *J. Nonlinear Sci.* **23** (2013) 427–456.

[22] L.C. Evans, Partial Differential Equations, 2 edition. American Mathematical Society (2010).

[23] R. Eymard, T. Gallouët and R. Herbin, Finite volume methods. *Handb. Numer. Anal.* **7** (2000) 713–1018.

[24] F. Filbet, A finite volume scheme for the Patlak–Keller–Segel chemotaxis model. *Numer. Math.* **104** (2006) 457–488.

[25] C.K. Hemelrijk and H. Hildenbrandt, Schools of fish and flocks of birds: their shape and internal structure by self-organization. *Interface Focus* **2** (2012) 726–737.

[26] R. Kowalczyk, Preventing blow-up in a chemotaxis model. *J. Math. Anal. App.* **305** (2005) 566–588.

[27] L. Ohm and M.J. Shelley, Weakly nonlinear analysis of pattern formation in active suspensions. *J. Fluid Mech.* **942** (2022) A53.

[28] K.J. Painter, P.K. Maini and H.G. Othmer, Stripe formation in juvenile pomacanthus explained by a generalized turing mechanism with chemotaxis. *Proc. Nat. Acad. Sci.* **96** (1999) 5549–5554.

[29] A. Perna, B. Granovskiy, S. Garnier, S.C. Nicolis, M. Labédan, G. Theraulaz, V. Fourcassié and D.J. Sumpter, Individual rules for trail pattern formation in argentine ants (*Linepithema humile*). *PLoS Comput. Biol.* **8** (2012) e1002592.

[30] A. Porretta, A note on the Sobolev and Gagliardo–Nirenberg inequality when $p > n$. *Adv. Nonlinear Stud.* **20** (2020) 361–371.

[31] J. Simon, Compact sets in the space $l^p(o, t; b)$. *Annali di Matematica pura ed applicata* **146** (1986) 65–96.

[32] G. Zhou and N. Saito, Finite volume methods for a Keller–Segel system: discrete energy, error estimates and numerical blow-up analysis. *Numer. Math.* **135** (2017) 265–311.

Appendix A. Continuity of the fixed point operator and uniqueness of numerical solutions

First, we introduce the notation

$$W[c^n,f^n]_{i,j,k+1/2} = \Big(B[c^n]_{i,j,k+1/2}\Big)_+ f^n_{i,j,k} + \Big(B[c^n]_{i,j,k+1/2}\Big)_- f^n_{i,j,k+1}. \tag{A.1}$$

Then, we consider a sequence $((c_m)^n_{i,j})_m$ that converges to $c^n_{i,j}$ in $\mathbb{R}^{N_x\times N_y\times N_T}$. Subtracting the scheme for $(f_m)^n_{i,j,k}$ by $f^n_{i,j,k}$ (obtained from $(c_m)^n_{i,j}$ and $c^n_{i,j}$, respectively) and multiplying by $(f_m-f)^{n+1}_{i,j,k} = g^{n+1}_{i,j,k}$, we obtain, after summation by parts,

$$\begin{aligned}
\sum_{i,j,k} \frac{g^{n+1}_{i,j,k}-g^n_{i,j,k}}{\Delta t} g^{n+1}_{i,j,k}\Delta\xi = -D_T\big\|\mathrm{d}_{\mathbf{x}} g_h\big(t^{n+1}\big)\big\|^2_{L^2} - \big\|\mathrm{d}_\theta g_h\big(t^{n+1}\big)\big\|^2_{L^2}& \\
+ \mathrm{Pe}\sum_{i,j,k}\Big(U^{n+1}_{i+1/2,j,k}\,\mathrm{d}_x g^{n+1}_{i+1/2,j,k} + V^{n+1}_{i,j+1/2,k}\,\mathrm{d}_y g^{n+1}_{i,j+1/2,k}\Big)\Delta\xi& \\
+ \gamma\sum_{i,j,k}\Big[W\big[(c_m)^{n+1},(f_m)^{n+1}\big]_{i,j,k+1/2} - W\big[c^{n+1},f^{n+1}\big]_{i,j,k+1/2}\Big]\,\mathrm{d}_\theta g^{n+1}_{i,j,k+1/2}\Delta\xi.&
\end{aligned}$$

Using the same estimates as in the proof of Proposition 3.3, we get

$$\mathrm{Pe}\sum_{i,j,k}\Big(U^{n+1}_{i+1/2,j,k}\,\mathrm{d}_x g^{n+1}_{i+1/2,j,k} + V^{n+1}_{i,j+1/2,k}\,\mathrm{d}_y g^{n+1}_{i,j+1/2,k}\Big)\Delta\xi \le 0. \tag{A.2}$$

We then add and subtract $W[(c_m)^{n+1},f^{n+1}]_{i,j,k+1/2}$ to the W-sum, so that

$$\begin{aligned}
&\sum_{i,j,k}\Big[W\big[(c_m)^{n+1},(f_m)^{n+1}\big]_{i,j,k+1/2} - W\big[c^{n+1},f^{n+1}\big]_{i,j,k+1/2}\Big]\,\mathrm{d}_\theta g^{n+1}_{i,j,k+1/2}\Delta\xi \\
&\quad = \sum_{i,j,k}\Big\{W\big[(c_m)^{n+1},(f_m)^{n+1}\big]_{i,j,k+1/2} - W\big[(c_m)^{n+1},f^{n+1}\big]_{i,j,k+1/2}\Big\}\,\mathrm{d}_\theta g^{n+1}_{i,j,k+1/2}\Delta\xi \\
&\qquad + \sum_{i,j,k}\Big\{W\big[(c_m)^{n+1},f^{n+1}\big]_{i,j,k+1/2} - W\big[c^{n+1},f^{n+1}\big]_{i,j,k+1/2}\Big\}\,\mathrm{d}_\theta g^{n+1}_{i,j,k+1/2}\Delta\xi \\
&\quad = I_1 + I_2.
\end{aligned}$$

For I_1, we note that we have

$$\begin{aligned}
&W\big[(c_m)^{n+1},(f_m)^{n+1}\big]_{i,j,k+1/2} - W\big[(c_m)^{n+1},f^{n+1}\big]_{i,j,k+1/2} \\
&\quad = \Big(B\big[(c_m)^{n+1}\big]_{i,j,k+1/2}\Big)_+\big((f_m)^{n+1}_{i,j,k} - f^{n+1}_{i,j,k}\big) \\
&\qquad + \Big(B\big[(c_m)^{n+1}\big]_{i,j,k+1/2}\Big)_-\big((f_m)^{n+1}_{i,j,k+1} - f^{n+1}_{i,j,k+1}\big) \\
&\quad = \Big(B\big[(c_m)^{n+1}\big]_{i,j,k+1/2}\Big)_+ g^{n+1}_{i,j,k} + \Big(B\big[(c_m)^{n+1}\big]_{i,j,k+1/2}\Big)_- g^{n+1}_{i,j,k+1}.
\end{aligned}$$

Then, we recall the inequality (3.34). Applying this inequality to I_1 gives

$$I_1 = \frac{\gamma}{2}\sum_{i,j,k} B\left[(c_m)^{n+1}\right]_{i,j,k+1/2}\mathrm{d}_\theta\left(g_{i,j,k+1/2}\right)^2\Delta\xi$$
$$= -\frac{\gamma}{2}\sum_{i,j,k}\mathrm{d}_\theta B\left[(c_m)^{n+1}\right]_{i,j,k}g_{i,j,k}^2\Delta\xi.$$

Then, using the discrete Poincaré–Sobolev inequality as in the proof of Proposition 3.3, we get for $\epsilon_1 > 0$ and some constant $A(\epsilon_1) > 0$,

$$|I_1| \le \epsilon_1\left\|\mathrm{d}_\xi g_h\left(t^{n+1}\right)\right\|_{L^2}^2 + A(\epsilon_1)\left\|\mathrm{d}_\theta B\left[(c_m)_h\left(t^{n+1}\right)\right]\right\|_{L^2}^2\left\|g_h\left(t^{n+1}\right)\right\|_{L^2}^2.$$

Using that $(\rho_m)_{i,j}^{n+1} \le \frac{1}{\Delta\mathbf{x}}$, writing $\Delta\mathbf{x} = \Delta x\Delta y$ for the fixed meshsize we get, and using Proposition 3.2,

$$|I_1| \le \epsilon_1\left\|\mathrm{d}_\xi g_h\left(t^{n+1}\right)\right\|_{L^2}^2 + \frac{A(\epsilon_1)}{\Delta\mathbf{x}^2}\left\|g_h\left(t^{n+1}\right)\right\|_{L^2}^2. \tag{A.3}$$

For I_2, we use $f_{i,j,k}^{n+1} \le \frac{1}{\Delta\xi}$, so that for $\epsilon_2 > 0, A(\epsilon_2) > 0$,

$$|I_2| \le \frac{1}{\Delta\xi}\sum_{i,j,k}\left|B\left[(c_m)^{n+1}\right]_{i,j,k+1/2} - B\left[c^{n+1}\right]_{i,j,k+1/2}\right|\left|\mathrm{d}_\theta g_{i,j,k}^{n+1}\right|\Delta\xi$$
$$\le \epsilon_2\left\|\mathrm{d}_\theta g_h\left(t^{n+1}\right)\right\|_{L^2}^2 + \frac{D(\epsilon_2)}{\Delta\xi^2}\left\|B\left[(c_m)_h\left(t^{n+1}\right)\right] - B\left[c_h\left(t^{n+1}\right)\right]\right\|_{L^2}^2.$$

Hence, combining everything we obtain, choosing ϵ_1, ϵ_2 small enough,

$$\frac{1}{2\Delta t}\left[\left\|g_h\left(t^{n+1}\right)\right\|_{L^2}^2 - \left\|g_h(t^n)\right\|_{L^2}^2\right] \le -\epsilon_3\left\|\mathrm{d}_\xi g_h\left(t^{n+1}\right)\right\|_{L^2}^2 + \frac{A(\epsilon_1)}{\Delta\mathbf{x}^2}\left\|g_h\left(t^{n+1}\right)\right\|_{L^2}^2$$
$$+ \frac{D(\epsilon_2)}{\Delta\xi^2}\left\|B\left[(c_m)_h\left(t^{n+1}\right)\right] - B\left[c_h\left(t^{n+1}\right)\right]\right\|_{L^2}^2.$$

Then, using the Gronwall inequality with forcing, requiring $\eta = \Delta t\frac{2A(\epsilon_1)}{\Delta\mathbf{x}^2} < 1$,

$$\left\|g_h(t^n)\right\|_{L^2}^2 \le \frac{1}{(1-\eta)^n}\left\|g_h(0)\right\|_{L^2}^2 + \frac{D(\epsilon_2)}{\Delta\xi^2}\sum_{\ell=1}^{n}\eta^{n-\ell+1}\left\|B\left[(c_m)_h\left(t^{n+1}\right)\right] - B\left[c_h\left(t^{n+1}\right)\right]\right\|_{L^2}^2\Delta t. \tag{A.4}$$

Hence, letting $(c_m) \to c$ and noting $\|g_h(0)\|_{L^2} = 0$ by construction, we have that $(f_m)_{i,j,k}^n = f_{i,j,k}^n$, hence $c \mapsto f$ is continuous.

Finally, for the map $f \mapsto \tilde{c}$ we note that the equation for $\tilde{c}$ is linear, so by the gradient estimate equation (3.28) we also get that $\tilde{c} \mapsto f$ is continuous. This concludes the continuity of the fixed point operator.

For the uniqueness, we take a similar approach but in the end convert the term $\|B[(c^1)_h(t^{n+1})] - B[(c^2)_h(t^{n+1})]\|_{L^2}$ to $\|(f^1)_h(t^{n+1}) - (f^2)_h(t^{n+1})\|_{L^2}$, for two fixed points (f^1, c^1) and (f^2, c^2). This would require $\Delta t/\Delta\xi^2$ to be small enough.

Appendix B. Calculations for look-ahead term

Definition B.1 (B_λ-term *via* interpolation)**.** We define the B_λ-term as

$$B_\lambda[c^n]_{i,j,k+1/2} = \mathbf{n}(\theta_{k+1/2})\cdot Dc_{i,j,k+1/2}^n, \tag{B.1}$$

where the discrete gradient with triple indices is accordingly defined as

$$Dc_{i,j,k+1/2} = \frac{1}{2}\begin{pmatrix}\frac{1}{\Delta x}(c_{i+1,j,k+1/2} - c_{i-1,j,k+1/2})\\ \frac{1}{\Delta y}(c_{i,j+1,k+1/2} - c_{i,j-1,k+1/2})\end{pmatrix}, \tag{B.2}$$

as before and where the term $c_{i,j,k+1/2}$ is defined by nearest-neighbour interpolation to approximate c at the point $\mathbf{x}_{i,j} + \lambda\mathbf{e}(\theta_{k+1/2})$, so that $c_{i,j,k+1/2} \approx c(\mathbf{x}_{i,j} + \lambda\mathbf{e}(\theta_{k+1/2}))$. That is, we define

$$c_{i,j,k+1/2} = c_{i(k+1/2),j(k+1/2)} \text{ if } \mathbf{x}_{i,j} + \lambda\mathbf{e}(\theta_{k+1/2}) \in C_{i(k+1/2),j(k+1/2)}, \tag{B.3}$$

so that $i(k+1/2)$ and $j(k+1/2)$ denote the i- and j-indices of the cell in which the point $\mathbf{x}_{i,j} + \lambda\mathbf{e}(\theta_{k+1/2})$ lies. See also Figure B.1 for an illustration of the points involved in the nearest-neighbour interpolation.

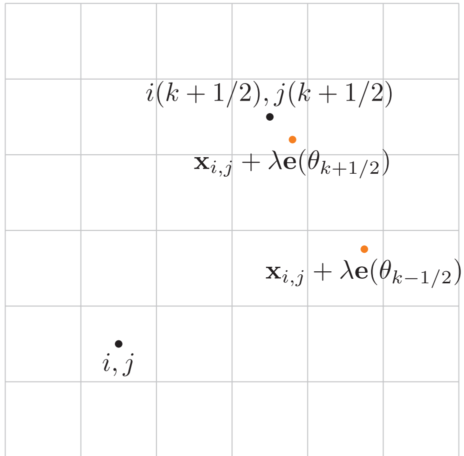


FIGURE B.1. Nearest-neighbour interpolation for B_λ.

Proof of B_λ interaction. We have

$$\mathrm{d}_\theta B_\lambda[c]_{i,j,k} = \frac{1}{\Delta\theta}\big(\mathbf{n}(\theta_{k+1/2}) \cdot Dc_{i,j,k+1/2} - \mathbf{n}(\theta_{k-1/2}) \cdot Dc_{i,j,k-1/2}\big). \tag{B.4}$$

By adding and subtracting a term with mixed signs in the k-index we can also write this as

$$\begin{aligned} \mathrm{d}_\theta B_\lambda[c]_{i,j,k} &= \frac{1}{\Delta\theta}\big(\mathbf{n}(\theta_{k+1/2}) - \mathbf{n}(\theta_{k-1/2})\big) \cdot Dc_{i,j,k+1/2} \\ &\quad + \frac{1}{\Delta\theta}\mathbf{n}(\theta_{k-1/2}) \cdot \big(Dc_{i,j,k+1/2} - Dc_{i,j,k-1/2}\big) \\ &=: T_1 + T_2. \end{aligned} \tag{B.5}$$

The idea of the proof is to bound T_1 and T_2 as the terms that would appear by applying the chain rule and product rule at the continuum level,

$$\partial_\theta(\mathbf{n}(\theta) \cdot \nabla c(\mathbf{x} + \lambda\mathbf{e}(\theta))) = \underbrace{-\mathbf{e}(\theta) \cdot \nabla c(\mathbf{x} + \lambda\mathbf{e}(\theta)))}_{\approx T_1} + \underbrace{\lambda\mathbf{n}(\theta) \cdot D^2 c(\mathbf{x} + \lambda\mathbf{e}(\theta))\mathbf{n}(\theta)}_{\approx T_2}. \tag{B.6}$$

For T_1, we first note that

$$\frac{c_{i+1,j,k+1/2} - c_{i-1,j,k+1/2}}{2\Delta x} = \frac{c_{i(k+1/2)+1,j(k+1/2)} - c_{i(k+1/2)-1,j(k+1/2)}}{2\Delta x}, \tag{B.7}$$

so that we have

$$D_x c_{i,j,k+1/2} = D_x c_{i(k+1/2),j(k+1/2)}, \tag{B.8}$$

and the same applies to the y-component. Therefore, we can bound T_1 as

$$|T_1| \le \max_{\theta\in[0,2\pi]} |\mathbf{e}(\theta)| \big|Dc_{i(k+1/2),j(k+1/2)}\big|. \tag{B.9}$$

For T_2, we split the x- and y-components of the inner product, so that we may write

$$\begin{aligned} T_2 = &- \tfrac{1}{\Delta\theta}\sin\big(\theta_{k-1/2}\big)\big(D_x c_{i(k+1/2),j(k+1/2)} - D_x c_{i(k-1/2),j(k-1/2)}\big) \\ &+ \tfrac{1}{\Delta\theta}\cos\big(\theta_{k-1/2}\big)\big(D_y c_{i(k+1/2),j(k+1/2)} - D_y c_{i(k-1/2),j(k-1/2)}\big) \\ =&: T_2^x + T_2^y. \end{aligned} \tag{B.10}$$

We note that, for the discrete derivatives in the x-component, after adding a term with mixed signs for the k-index,

$$\begin{aligned} &D_x c_{i(k+1/2),j(k+1/2)} - D_x c_{i(k-1/2),j(k-1/2)} \\ &\quad = \big(D_x c_{i(k+1/2),j(k+1/2)} - D_x c_{i(k-1/2),j(k+1/2)}\big) + \big(D_x c_{i(k-1/2),j(k+1/2)} - D_x c_{i(k-1/2),j(k-1/2)}\big) \\ &\quad =: T^x_{2,1} + T^x_{2,2}. \end{aligned} \tag{B.11}$$

Now, focusing on $T^x_{2,1}$, we can add a telescoping sum ranging in the i-index from $i(k-1/2)$ to $i(k+1/2)$, so that

$$\begin{aligned} T^x_{2,1} &= -D_x c_{i(k-1/2),j(k+1/2)} + \sum_{r=i(k-1/2)+1}^{i(k+1/2)-1} \big[D_x c_{r,j(k+1/2)} - D_x c_{r,j(k+1/2)}\big] + D_x c_{i(k+1/2),j(k+1/2)} \\ &= \sum_{r=i(k-1/2)}^{i(k+1/2)-1} \big[D_x c_{r+1,j(k+1/2)} - D_x c_{r,j(k+1/2)}\big]. \end{aligned} \tag{B.12}$$

By adding and subtracting at each summand the term $\mathrm{d}_x c_{r+1/2,j(k+1/2)}$ (as defined on the dual mesh) we get

$$\begin{aligned} T^x_{2,1} &= \sum_{r=i(k-1/2)}^{i(k+1/2)-1} \frac{1}{2\Delta x}\Big[\big(c_{r+2,j(k+1/2)} - c_{r,j(k+1/2)}\big) - 2\big(c_{r+1,j(k+1/2)} - c_{r,j(k+1/2)}\big) \\ &\qquad + 2\big(c_{r+1,j(k+1/2)} - c_{r,j(k+1/2)}\big) - \big(c_{r+1,j(k+1/2)} - c_{r-1,j(k+1/2)}\big)\Big] \\ &= \sum_{r=i(k-1/2)}^{i(k+1/2)-1} \frac{\Delta x}{2}\big(\mathrm{d}^2_x c_{r+1,j(k+1/2)} + \mathrm{d}^2_x c_{r,j(k+1/2)}\big). \end{aligned} \tag{B.13}$$

For $T^x_{2,2}$, we do a similar calculation. By adding a telescoping sum we obtain

$$T^x_{2,2} = \sum_{r=j(k-1/2)}^{j(k+1/2)-1} \big[D_x c_{i(k-1/2),r+1} - D_x c_{i(k-1/2),r}\big], \tag{B.14}$$

and by adding and subtracting $\frac{1}{2\Delta x}[c_{i(k-1/2),r+1} - c_{i(k-1/2),r}]$ for each summand we can write

$$T^x_{2,2} = \sum_{r=j(k-1/2)}^{j(k+1/2)-1} \frac{\Delta y}{2}\big[\mathrm{d}_x\,\mathrm{d}_y c_{i(k-1/2)-1/2,r+1/2} + \mathrm{d}_x\,\mathrm{d}_y c_{i(k-1/2)+1/2,r+1/2}\big], \tag{B.15}$$

recalling the notation

$$\mathrm{d}_x\,\mathrm{d}_y c_{i+1/2,j+1/2} = \frac{1}{\Delta \mathbf{x}}[c_{i,j} - c_{i+1,j} - c_{i,j+1} + c_{i+1,j+1}]. \tag{B.16}$$

We now recall the definitions $C_{\Delta x} = \Delta\theta/\Delta x, C_{\Delta y} = \Delta\theta/\Delta y$, so that by inserting $T^x_{2,1}$ and $T^x_{2,2}$ from above into T^x_2 we are lead to

$$\begin{aligned} T^x_2 &= \frac{-\sin\big(\theta_{k-1/2}\big)}{2C_{\Delta x}} \sum_{r=i(k-1/2)}^{i(k+1/2)-1} \big[\mathrm{d}^2_x c_{r+1,j(k+1/2)} + \mathrm{d}^2_x c_{r,j(k+1/2)}\big] \\ &\quad + \frac{-\sin\big(\theta_{k-1/2}\big)}{2C_{\Delta y}} \sum_{r=j(k-1/2)}^{j(k+1/2)-1} \big[\mathrm{d}_x\,\mathrm{d}_y c_{i(k-1/2)-1/2,r+1/2} + \mathrm{d}_x\,\mathrm{d}_y c_{i(k-1/2)+1/2,r+1/2}\big]. \end{aligned} \tag{B.17}$$

We can use a similar argument for T^y_2, so that

$$\begin{aligned} T^y_2 &= \frac{\cos\big(\theta_{k-1/2}\big)}{2C_{\Delta x}} \sum_{r=i(k-1/2)}^{i(k+1/2)-1} [\mathrm{d}_x\,\mathrm{d}_y c_{r+1/2,j(k-1/2)-1/2} + \mathrm{d}_x\,\mathrm{d}_y c_{r+1/2,j(k-1/2)+1/2}] \\ &\quad + \frac{\cos\big(\theta_{k-1/2}\big)}{2C_{\Delta y}} \sum_{r=j(k-1/2)}^{j(k+1/2)-1} \big[\mathrm{d}^2_y c_{i(k+1/2),r+1} + \mathrm{d}^2_y c_{i(k+1/2),r}\big], \end{aligned} \tag{B.18}$$

Putting everything together, we get, using Jensen's inequality,

$$\begin{aligned}\|\mathrm{d}_\theta B[c_h]\|_{L^2}^2 \leq &\sum_{i,j,k} \left|Dc_{i(k+1/2),j(k+1/2)}\right|^2 \Delta\xi \\ &+ \max\left\{2\pi\lambda, \tfrac{1}{2\varepsilon_C}\right\} \sum_{i,j,k} \\ &\times \left[\left|\mathrm{d}_x^2 c_{i(k+1/2),j(k+1/2)}\right|^2 + 2\left|\mathrm{d}_x\, \mathrm{d}_y c_{i(k+1/2)+1/2,j(k+1/2)+1/2}\right|^2 + \left|\mathrm{d}_y^2 c_{i(k+1/2),j(k+1/2)}\right|^2\right] \Delta\xi \\ \leq &(2\pi)\left[\|\mathrm{d}_{\mathbf{x}} c_h\|_{L^2}^2 + \max\left\{2\pi\lambda, \tfrac{1}{2\varepsilon_C}\right\} \left\|\mathrm{d}_{\mathbf{x}}^2 c_h\right\|_{L^2}^2\right].\end{aligned}$$

Here we used that $|i(k+1/2)-1-i(k-1/2)| \leq \max\{2\pi\lambda C_{\Delta x}, 1\}$, $|j(k+1/2)-1-j(k-1/2)| \leq \max\{2\pi\lambda C_{\Delta y}, 1\}$ and that $C_{\Delta x}$ and $C_{\Delta y}$ are uniformly bounded below strictly from zero by $\varepsilon_C > 0$.

We also note that it is clear that, for any grid function $(g_{i,j})_{(i,j)}$, $\sum_{i,j} g_{i(k),j(k)} = \sum_{i,j} g_{i,j}$, since the k-index shift induces only a translation of the grid points.

This concludes the proof for the B_λ-term. □

Appendix C. Algebra for the Alikakos lemma

Algebra for (5.19). First we note that the discrete Gagliardo–Nirenberg inequality in dimension two implies

$$\|u\|_{L^2}^2 \leq \frac{1}{\kappa} C_{GN}^2 \|u\|_{L^1}^2 + \kappa \|u\|_{1,2}^2, \tag{C.1}$$

for any $0 < \kappa < 1$. This can be rewritten as

$$-\frac{\kappa}{1-\kappa}|u|_{1,2}^2 \leq -\|u\|_{L^2}^2 + \frac{1}{\kappa(1-\kappa)} C_{GN}^2 \|u\|_{L^1}^2. \tag{C.2}$$

We now let κ be such that $2^{-qk} = \frac{\kappa}{1-\kappa}$. We let ε be such that $D_T - \frac{\varepsilon}{2}\mathrm{Pe}^2 = \frac{1}{2}$, and define $p = 2^k - 1$,$a_k = \frac{2}{\varepsilon}(2^k-1)2^k = \frac{2}{\varepsilon}p(p+1)$, $\epsilon_k = \frac{1}{2^{qk}}$ and $c_k = 2^{\beta k}$, so that $\beta \geq 0$ is such that $C_{GN}^2 \frac{1}{\kappa(1-\kappa)} \leq 2^{\beta k}$ for all integers $k \geq 1$. (The latter is true for β sufficiently larger than q, but we still have to choose q.) Then, (C.2) can be written as

$$-\epsilon_k |u|_{1,2}^2 \leq -\|u\|_{L^2}^2 + c_k \|u\|_{L^1}^2. \tag{C.3}$$

We now require $-\frac{2p}{p+1} \leq -(a_k + \epsilon_k)\epsilon_k$, which can be written as

$$-\frac{2\left(2^k-1\right)}{2^k} \leq -\left(\frac{2}{\varepsilon}\left(2^k-1\right)2^k + \frac{1}{2^{qk}}\right)\frac{1}{2^{qk}}. \tag{C.4}$$

This inequality holds for all integers $k \geq 1$ for $q \geq 0$ large enough.

Hence, multiplying (C.3) by $(a_k + \epsilon_k)$ we get

$$-\frac{2p}{p+1}\left|(\rho_h)^{\frac{p+1}{2}}\right|_{1,2}^2 \leq -(a_k+\epsilon_k)\epsilon_k \left|(\rho_h)^{\frac{p+1}{2}}\right|_{1,2}^2 \leq -(a_k+\epsilon_k)\left\|(\rho_h)^{\frac{p+1}{2}}\right\|_{L^2}^2 + (a_k+\epsilon_k)c_k\left\|(\rho_h)^{\frac{p+1}{2}}\right\|_{L^1}^2. \tag{C.5}$$

This gives the algebra to derive (5.19). □

Algebra for (5.54). The discrete Gagliardo–Nirenberg inequality in dimension three implies,

$$\|u\|_{L^2}^2 \leq C_{GN}^2 \|u\|_{1,2}^{\frac{6}{5}} \|u\|_{L^1}^{\frac{4}{5}}. \tag{C.6}$$

Then, using the Hölder conjugates $p = \frac{5}{3}$ and $q = \frac{5}{2}$ we get

$$\|u\|_{L^2}^2 \leq \frac{1}{\kappa^{\frac{5}{2}}} C_{GN}^5 \|u\|_{L^1}^2 + \kappa^{\frac{5}{3}} \|u\|_{1,2}^2. \tag{C.7}$$

This can be rewritten as

$$-\frac{\kappa^{\frac{5}{3}}}{1-\kappa^{\frac{5}{3}}}|u|_{1,2}^2 \leq -\|u\|_{L^2}^2 + \frac{C_{GN}^5}{\kappa^{\frac{5}{2}}\left(1-\kappa^{\frac{5}{3}}\right)}\|u\|_{L^1}^2. \tag{C.8}$$

We choose q and κ such that $2^{-qk} = \frac{\kappa^{\frac{5}{3}}}{1-\kappa^{\frac{5}{3}}}$. From here on, we can do the same algebra as for (C.5) to conclude that

$$-\frac{2p}{p+1}|(f_h)^{\frac{p+1}{2}}|_{1,2}^2 \leq -(a_k+\epsilon_k)\epsilon_k\left|(f_h)^{\frac{p+1}{2}}\right|_{1,2}^2 \leq -(a_k+\epsilon_k)\left\|(f_h)^{\frac{p+1}{2}}\right\|_{L^2}^2 + (a_k+\epsilon_k)c_k\left\|(f_h)^{\frac{p+1}{2}}\right\|_{L^1}^2. \tag{C.9}$$

□