THE SCHARFETTER–GUMMEL SCHEME FOR AGGREGATION-DIFFUSION EQUATIONS

ANDRÉ SCHLICHTING* AND CHRISTIAN SEIS†

Abstract. In this paper, we propose a finite-volume scheme for aggregation-diffusion equations that is based on a Scharfetter–Gummel approximation of the nonlinear, nonlocal flux term. This scheme is analyzed concerning well-posedness and convergence towards solutions to the continuous problem. Also, it is proven that the numerical scheme has several structure-preserving features. More specifically, it is shown that the discrete solutions satisfy a free-energy dissipation relation analogous to the continuous model, and, as a consequence, the numerical solutions converge in the large time limit to stationary solutions, for which we provide a thermodynamic characterization.

1. Introduction

1.1. **The aggregation-diffusion equation.** We consider the aggregation-diffusion equation in a smoothly bounded domain Ω in $\mathbf{R}^d$, describing the evolution of a probability density $\rho \in \mathcal{P}(\Omega)$ according to

$$\partial_t \rho = \nabla \cdot (\kappa \nabla \rho + \rho \nabla W * \rho) \,, \tag{1}$$

where $\kappa > 0$ is the diffusion constant and $W : \mathbf{R}^d \to \mathbf{R}$ an interaction potential. The convolution in (1) is to be understood in Ω, that is

$$(W * \rho)(x) = \int_\Omega W(x-y)\rho(y)\,\mathrm{d}y.$$

Moreover, in order to preserve the total mass, we have to impose the no-flux boundary conditions

$$\kappa \partial_\nu \rho + \rho \partial_\nu W * \rho = 0 \qquad \text{on } \partial\Omega, \tag{2}$$

where ν denotes the outer normal vector on $\partial\Omega$. The initial datum will be denoted by ρ_0, and we assume that it is a probability distribution, $\rho_0 \in \mathcal{P}(\Omega)$.

Equation (1) arises as the infinite particle limit from systems of weakly interacting diffusions, first analysed by McKean [46, 47]. It is rigorously shown in [48, 52, 51] that the law of the empirical distribution of the particle system converges to the solution ρ of (1) in the so-called mean field limit.

The derivation already shows, that (1), maybe with an additional external potential, arises in many applications in which interactions between particles or agents are present. The examples range from opinion dynamics [33], granular materials [3, 18, 8] and

(*) Institut für Angewandte Mathematik, Universität Bonn, 53115 Bonn, Germany.
(†) Institut für Analysis und Numerik, Univerität Münster, 48149 Münster, Germany.
E-mail addresses: `schlichting@iam.uni-bonn.de, seis@uni-muenster.de`.

Date: April 29, 2020.
Key words and phrases. finite volume scheme, aggregation-diffusion equation, structure-preserving, free-energy dissipation relation, convergence, large time limit.

mathematical biology [38, 51, 10] to statistical mechanics [45], galactic dynamics [6], liquid-vapor transitions [42, 24], plasma physics [7], and synchronization [40].

In these model, all the physical properties of the underlying system are encoded through the interaction potential $W : \mathbf{R}^d \to \mathbf{R}$, which describes attractive and repulsive forces. We analyze the common case, where the force coming from the interaction potential between particles is antisymmetric and hence W is assumed to be symmetric. In addition, we assume W to be globally Lipschitz continuous and continuous differentiable away from the origin. These assumptions imply existence and uniqueness of weak solutions to (1) by standard arguments (see for instance [17, Theorem 2.2]), but rule out singular potentials blowing up at the origin, like the Newtonian potential for the Keller–Segel model [38]. In summary, for later reference, our assumptions are

(A1) $$W(x) = W(-x) \text{ and } W(0) = 0\,;$$

(A2) $$W \in C^1(\mathbf{R}^d \setminus \{0\})\,;$$

(A3) $$W \text{ is Lipschitz-continuous.}$$

The aggregation-diffusion equation (1) has a free energy functional acting as Lyapunov function for the evolution. It is defined as the sum of entropy and interaction energy,

$$\mathcal{F}(\rho) = \kappa \int_\Omega \rho(x) \log \rho(x) \,\mathrm{d}x + \frac{1}{2} \iint_{\Omega\times\Omega} W(x-y)\rho(x)\rho(y) \,\mathrm{d}x \,\mathrm{d}y\,. \tag{3}$$

Indeed, a short computation reveals that the free energy dissipation is given by

$$\frac{\mathrm{d}}{\mathrm{d}t}\mathcal{F}(\rho) = -\int_\Omega \rho |\nabla(\kappa \log \rho + W * \rho)|^2 \,\mathrm{d}x = -\mathcal{D}(\rho). \tag{4}$$

It is natural to assume that the free energy is initially finite, thus $\mathcal{F}(\rho_0) \in \mathbf{R}$. Moreover, by writing

$$\partial_t \rho = \nabla \cdot (\rho(\kappa \nabla \log \rho + \nabla W * \rho)) = -\nabla \cdot (\rho \nabla(-D\mathcal{F}(\rho)))\,, \tag{5}$$

it becomes evident that (1) is a gradient flow of the free energy $\mathcal{F}$ with respect to the Wasserstein distance (cf. [35, 18, 56]). In the thermodynamic interpretation, the term $\nabla(-D\mathcal{F}(\rho))$ is the (generalized) force coming from the second law of thermodynamics driving the system towards states of lower free energy. Hence, in accordance to thermodynamic principles, the dissipation in (4) can be expressed as the product of the force $\nabla(-D\mathcal{F}(\rho))$ with the flux $j = -\kappa\nabla\rho - \rho\nabla W * \rho$ from (1),

$$\mathcal{D}(\rho) = \int_\Omega (\kappa \nabla \log \rho + \nabla W * \rho) \cdot (\kappa \nabla \rho + \rho \nabla W * \rho) \,\mathrm{d}x. \tag{6}$$

The thermodynamic structure of the equation also provides a physical formulation of the stationary solutions to (1). On the one hand, it was already observed in [39], that $\rho \in \mathcal{P}(\Omega)$ is a stationary state if and only if ρ is a fixed point of the Kirkwood–Monroe fixed point map $T : \mathcal{P}_{\mathrm{ac}}(\Omega) \to \mathcal{P}_{\mathrm{ac}}^+(\Omega)$ given by

$$T(\rho)(x) = \frac{e^{-\kappa^{-1} W*\rho(x)}}{\int e^{-\kappa^{-1} W*\rho(y)} \,\mathrm{d}y}, \tag{7}$$

where $\mathcal{P}_{\mathrm{ac}}^{(+)}(\Omega)$ is the set of absolutely continuous positive probability measures. The map T allows to rewrite (1) in symmetric state-dependent form

$$\partial_t \rho = \kappa \nabla \cdot \Big(T(\rho) \nabla \frac{\rho}{T(\rho)} \Big). \tag{8}$$

Hence, every stationary state has a Boltzmann statistical representation and it immediately follows that any equilibrium satisfies the detailed balance condition: $j = -\kappa T(\rho)\nabla \frac{\rho}{T(\rho)} = -\kappa \nabla \rho - \rho \nabla W * \rho = 0$. The gradient flow formulation of (1) on the other hand provides that stationary points are critical points of $\mathcal{F}(\rho)$ and diminish the dissipation $\mathcal{D}(\rho) = 0$.

Finally, the free energy dissipation principle (4) gives also a way to show the large time behavior of (1). After proving suitable lower-semicontinuity properties of $\mathcal{F}$ and $\mathcal{D}$, one can argue by the LaSalle invariance principle (see, e.g., [57, Theorem 4.2 in Chapter IV]) that for $t \to \infty$ any solution converges to the set of stationary points of (1). This is proven for a related models with specific interaction potentials in [53, 3, 55, 24].

1.2. **Goals.** The main goal of this work is to provide a numerical scheme for (1), which preserves the free energy structure of the equation and is stable for all ranges of κ. In particular, the scheme should satisfy a discrete analog to the identity (4) and preserve the Boltzmann statistical form of stationary states in (7). For this purpose, we propose a (semi-)implicit finite volume scheme based on a Scharfetter–Gummel discretization of the flux. The resulting flux has a physical derivation by solving a suitable cell problem between neighboring cells (see Section 2.2). Indeed, we recover the free energy dissipation principle for the Scharfetter–Gummel discretization in the product form of force times flux (cf. (6)) in Proposition 1.

The free energy dissipation principle also gives us a way to characterize the stationary states: First, they are fixed points of a suitable discrete Kirkwood–Monroe map; second, they are critical points of the numerical free energy; and third, they are states with vanishing numerical dissipation. This characterization of stationary states, in Theorem 2 below, shows that the proposed discretization is consistent with the thermodynamic structure of (1).

Finally, the numerical free energy dissipation identity provides also a control on the gradient for solutions of the scheme (cf. Lemma 3), which allows us to obtain compactness (cf. Proposition 2) in order to conclude that the discretized solution converges to the weak solution of (1) in Theorem 3. In addition, by exploiting the characterization of stationary states as critical points of $\mathcal{F}$, we can also conclude that every stationary solution of the Scharfetter–Gummel scheme converges to a stationary solution of the continuous problem.

The paper is organized as follows: In Section 2, we introduce the numerical Scharfetter–Gummel finite volume scheme approximating the aggregation-diffusion equation (34) and motivate the particular form of the flux by discussing the associated one-dimensional cell problem. We subsequently present and discuss or main results concerning well-posedness and convergence of the scheme, characterization of stationary states, and the large time behavior. The section concludes with a discussion of related work. The proofs are all contained in Section 3.

2. The numerical scheme

In this section, we will introduce the numerical scheme and derive some of its most elementary and substantial features.

2.1. **Definition.**

We first introduce the general notation that is required to define a finite volume method. We start with the tesselation of Ω, that we assumed to be smoothly bounded earlier. For technical reasons, that we will briefly discuss later, it is convenient to choose a tesselation consisting of Voronoi cells. Hence, we let $\mathcal{T}^h$ be a Voronoi tesselation covering Ω such that K is compact and $K \cap \Omega \neq \emptyset$ for any $K \in \mathcal{T}^h$ and with maximal size

$$\sup_{K\in\mathcal{T}^h} \operatorname{diam}(K) \leq h\,.$$

We set $\hat{\Omega} = \cup_K K$, which contains Ω by construction. The generator of each cell $K \in \mathcal{T}^h$ will be denoted by x_K, and we set $d_{KL} = d(x_K, x_L)$, where d is the Euclidean distance on $\mathbf{R}^d$. If K and L are two neighboring cells, we write $L \sim K$, and we denote by $K|L$ the common edge, $K|L = \bar{K} \cap \bar{L}$. We furthermore denote by τ_{KL} the transmission coefficients,

$$\tau_{KL} = \frac{|K|L|}{d_{KL}}\,.$$

Here and in the following, the symbol $|\cdot|$ is used to denote an area, but we will also utilize it for volumes. Hence, $|K|$ is the volume of a cell K and $|\partial K|$ is the area of its surface.

We finally discretize time. The time step size will be denoted by δt and we set $t^n = n\,\delta t$ for any $n \in \mathbf{N}_0$.

With these preparations at hand, we are now in the position to introduce the finite volume approximations. To start with, we discretize the initial datum. By extending ρ_0 by zero to $\hat{\Omega}$, we may consider the averages $\rho_K^0 = \fint_K \rho_0\,\mathrm{d}x$ on each cell $K \in \mathcal{T}^h$, and set $\rho_h^0 = \{\rho_K^0\}_{K\in\mathcal{T}^h}$. It readily checked that the finite volume approximation ρ_h^0 of the initial configuration is a probability distribution. The set of all probability distributions on $\mathcal{T}^h$ will be denoted by $\mathcal{P}(\mathcal{T}^h)$, i.e.,

$$\mathcal{P}(\mathcal{T}^h) = \left\{ \{\rho_K\}_{K\in\mathcal{T}^h} : \rho_K \geq 0 \ \forall K \in \mathcal{T}^h \text{ and } \sum_K |K|\rho_K = 1 \right\}.$$

The general iteration scheme that constitutes a discrete evolution equation reads

$$|K|\frac{\rho_K^{n+1} - \rho_K^n}{\delta t} + \sum_{L\sim K} F_{KL}^{n+1} = 0, \tag{9}$$

where F_{KL}^{n+1} is the numerical flux from cell K to its neighbor L. Notice that there is no flux across the outer boundary $\partial\hat{\Omega}$ in accordance with (2). The Scharfetter–Gummel scheme approximates the simultaneous flux due to diffusion *and* advection across a common edge in terms of the Bernoulli function $B_\kappa : \mathbf{R} \to \mathbf{R}$ given for $\kappa > 0$ by

$$B_\kappa(s) = \begin{cases} \frac{s}{e^{\frac{s}{\kappa}}-1}, & \text{for } s \neq 0, \\ \kappa, & \text{for } s = 0. \end{cases} \tag{10}$$

This function is convex, strictly decreasing and satisfies

$$B_\kappa(s) \geq (s)^- = \max\{-s, 0\} \qquad \text{and} \qquad \lim_{\kappa\to 0} B_\kappa(s) \to (s)^- \,,$$

for any $s \in \mathbf{R}$. We propose the Scharfetter–Gummel numerical flux an approximation for the aggregation-diffusion equation (1) in the form

$$F_{KL}^{n+1} = \tau_{KL}\Big(B_\kappa\Big(d_{LK} q_{LK}^{n+1}\Big)\rho_K^{n+1} - B_\kappa\Big(d_{KL} q_{KL}^{n+1}\Big)\rho_L^{n+1}\Big), \tag{11}$$

where q_{KL}^{n+1} is a discretization of the aggregation convolution term,

$$q_{KL}^{n+1} = \sum_{J\in\mathcal{T}} |J| \, \frac{\rho_J^{n+1} + \rho_J^n}{2} \, \frac{W(x_K - x_J) - W(x_L - x_J)}{d_{KL}} \,. \tag{12}$$

We will motivate this definition of the numerical flux briefly in the next subsection 2.2. The arithmetic mean occurring in time in (12) is needed to have a numerical analogue of the free energy dissipation relation (4) for general interaction potentials W, see Theorem 1 below. The same choice was made in [1] to ensure dissipation for the upwind scheme. For further reference, we remark that both flux and convolution term are antisymmetric,

$$q_{KL}^{n+1} = -q_{LK}^{n+1} \qquad \text{and} \qquad F_{KL}^{n+1} = -F_{LK}^{n+1}. \tag{13}$$

We conclude this subsection by stating three equivalent formulations of (11) and (12), which will conveniently simplify later discussions and computations. First, by introducing the unidirectional numerical fluxes from cell K to L denoted by j_{KL}^{n+1}, the flux can be written in divergence form,

$$j_{KL}^{n+1} = B_\kappa\Big(d_{LK} q_{LK}^{n+1}\Big)\rho_K^{n+1} \qquad \text{and hence} \qquad F_{KL}^{n+1} = \tau_{KL}\Big(j_{KL}^{n+1} - j_{LK}^{n+1}\Big) \,. \tag{14}$$

Using the antisymmetry of the convolution term (13), and the definition of the Bernoulli function (10), we can furthermore write

$$\frac{F_{KL}^{n+1}}{\tau_{KL}} = j_{KL}^{n+1} - j_{LK}^{n+1} = d_{KL} q_{KL}^{n+1} \frac{\rho_K^{n+1} e^{\frac{d_{KL} q_{KL}^{n+1}}{2\kappa}} - \rho_L^{n+1} e^{-\frac{d_{KL} q_{KL}^{n+1}}{2\kappa}}}{e^{\frac{d_{KL} q_{KL}^{n+1}}{2\kappa}} - e^{-\frac{d_{KL} q_{KL}^{n+1}}{2\kappa}}} \,. \tag{15}$$

In fact, in our motivation for the particular form of the Scharfetter–Gummel flux, we will derive this identity rather than (11). Finally, thanks to the elementary identity $B_\kappa(s) = \frac{s}{2}\left(\coth\left(\frac{s}{2\kappa}\right) - 1\right)$, we may rewrite (11) as

$$F_{KL}^{n+1} = |K_{|}L| q_{KL}^{n+1} \frac{\rho_K^{n+1} + \rho_L^{n+1}}{2} + \frac{1}{2}|K_{|}L| q_{KL}^{n+1} \coth\left(\frac{d_{KL} q_{KL}^{n+1}}{2\kappa}\right)\left(\rho_K^{n+1} - \rho_L^{n+1}\right) \,. \tag{16}$$

This formulation of the numerical flux is particularly helpful, as it (roughly) separates the aggregation term from the diffusion term.

2.2. **Cell problem.** We briefly give some background about the flux relation (15), which is obtained from the solution of the following one dimensional cell problem. For given ρ_K, ρ_L and q_{KL}, the resulting normalized (per unit interface area) net flux $f_{KL} = F_{KL}/\tau_{KL}$ is obtained as the solution of the boundary value problem

$$\begin{aligned} &f_{KL} = -\kappa \partial_x \rho(\cdot) + q_{KL}\rho(\cdot) \quad \text{on } (0, d_{KL}), \\ &\rho(0) = \rho_K \qquad \text{and} \qquad \rho(d_{KL}) = \rho_L. \end{aligned} \tag{17}$$

Hence besides $f_{KL} \in \mathbf{R}$, the function $\rho : [0,1] \to \mathbf{R}$ is part of the unknown in (17). It is readily checked that (15) is the solution to (17).

It might be insightful to study the solution f_{KL} of the cell problem in terms of its parameters ρ_K, ρ_L, $d_{KL}q_{KL}$, and κ. We thus introduce the function $\theta_\kappa : \mathbf{R}_+ \times \mathbf{R}_+ \times \mathbf{R} \to \mathbf{R}$ by

$$\theta_\kappa(a,b;v) = \begin{cases} v\,\dfrac{ae^{\frac{v}{2\kappa}} - be^{-\frac{v}{2\kappa}}}{e^{\frac{v}{2\kappa}} - e^{-\frac{v}{2\kappa}}}, & \text{for } v \neq 0, \\ \kappa(a-b), & \text{for } v = 0. \end{cases} \tag{18}$$

In this way, we can write

$$f_{KL} = \theta_\kappa(\rho_K, \rho_L; d_{KL}q_{KL}). \tag{19}$$

Hence, we observe that the Scharfetter–Gummel flux solves the two-point boundary problem (17), which is the main reason why it preserves many of the structural properties of the equation (1). It seems straightforward to generalize this idea to aggregation-diffusion equations with diffusion operators having nonlinear mobility, like in the case of the porous medium equation. Although the resulting flux cannot be expressed as a simple closed function of ρ_K, ρ_L, and v, it still has many physical properties of the continuum equation. These generalizations are studied in [29] for nonlinear diffusion with linear drift. Similar constructions for generalized Scharfetter–Gummel schemes are known in the literature on numerical methods for semiconductors [32, Section 4.2]. To avoid to work with nonexplicit functions for the flux, Bessemoulin-Chatard [4] introduced a class of *modified Scharfetter–Gummel* schemes, which she applied to nonlinear diffusion equations with linear drift. The generalization of these approaches to interaction equations with nonlinear diffusion remains to be investigated.

We conclude this subsection with a remark, in which we gather some properties of the solution function θ_κ that relate the Scharfetter–Gummel flux to physically known quantities and relations on the one hand, and the numerical upwind scheme on the other hand.

Remark 1. *For any $\kappa, a, b > 0$, the following holds:*

(i) *The function θ_κ is* one-homogeneous *in the first two variables, i.e., for any $\lambda > 0$ and $v \in \mathbf{R}$, it holds that $\theta_\kappa(\lambda a, \lambda b; v) = \lambda\theta_\kappa(a,b;v)$.*

(ii) *For small velocities, one recovers* Fick's law *and the next order, i.e.,*

$$\theta_\kappa(a,b;v) = \kappa(a-b) + \frac{a+b}{2}v + O(|v|^2) \qquad \text{as } |v| \to 0. \tag{20}$$

(iii) *The* no-flux velocity *v satisfying $\theta_\kappa(a,b;v) = 0$ is characterized by $v = -\kappa \log \frac{a}{b}$.*

(iv) *It holds the* Onsager relation*: $\mathbf{R} \ni v \mapsto \theta_\kappa(a,b;v)$ is one-to-one, since*

$$\min\{a,b\} \le \partial_v \theta_\kappa(a,b;v) \le \max\{a,b\}. \tag{21}$$

(v) *In the zero-diffusivity limit, θ_κ reduced to the* upwind flux*,*

$$\lim_{\kappa\to 0} \theta_\kappa(a,b;v) = \begin{cases} av & , \text{ for } v > 0, \\ bv & , \text{ for } v < 0, \\ 0 & , \text{ for } v = 0. \end{cases}$$

2.3. **Main results.** We introduce the minimal distance $d_{\min} = \min_{K,L} d_{KL}$. For our analysis, we have to ensure that the scheme is not degenerating in the sense that cells have the uniform isoperimetric property

$$\frac{|\partial K|}{|K|} \le \frac{C_{\mathrm{iso}}}{h}. \tag{22}$$

Moreover, the time step size has to be sufficiently small in the sense that

$$\delta t \le \frac{h\, d_{\min}}{8 C_{\mathrm{iso}} \kappa} \qquad \text{and} \qquad \delta t \le \frac{h}{12 C_{\mathrm{iso}} \operatorname{Lip}(W)}. \tag{23}$$

Note, that if the mesh is regular in the sense that $d_{\min} \gtrsim h$, the first condition in (23) becomes the usual parabolic scaling assumption $\delta t \lesssim h^2$.

Theorem 1 (Well-posedness). *Under the assumptions (22) and (23), the Scharfetter–Gummel scheme (9), (11), (12) with initial condition in $\mathcal{P}(\mathcal{T}^h)$ has a unique solution. This solution is mass preserving and after the first time step strictly positive. Moreover, the scheme is stable in $\ell^1(\mathcal{T}^h)$, that is for $\rho^0, \tilde\rho^0 \in \mathcal{P}(\mathcal{T}^h)$ exits some C such that for all n*

$$\|\rho^n - \tilde\rho^n\|_{\ell^1(\mathcal{T}^h)} \le C^n \|\rho^0 - \tilde\rho^0\|_{\ell^1(\mathcal{T}^h)} \,. \tag{24}$$

In analogy to (3), we define the numerical free energy, which we split into its entropic part and interaction energy given by

$$\mathcal{F}^h(\rho) = \kappa \mathcal{S}^h(\rho) + \mathcal{E}^h(\rho), \tag{25}$$

$$\text{with} \qquad \mathcal{S}^h(\rho) = \sum_K |K| \rho_K \log \rho_K$$

$$\text{and} \qquad \mathcal{E}^h(\rho) = \frac{1}{2}\sum_{K,L} |K||L| W(x_K - x_L)\rho_K \rho_L \,. \tag{26}$$

We establish a discrete version of the free energy dissipation relation (4). Besides the free energy functional, we need to introduce the relative entropy between $\rho^0, \rho^1 \in \mathcal{P}(\mathcal{T}^h)$ given by

$$\mathcal{H}(\rho \mid \tilde\rho) = \sum_K |K| \rho_K \log \frac{\rho_K}{\tilde\rho_K}, \tag{27}$$

which will occur as an additional dissipation in time due to the implicit time discretization.

Proposition 1 (Free energy dissipation and large time behavior). *Let $\{\rho_K^n\}_{K,n}$ be a solution to the Scharfetter–Gummel scheme (9), (11), (12). Then, for any $n \in \mathbf{N}_0$ it holds that*

$$\frac{\mathcal{F}^h(\rho^{n+1}) - \mathcal{F}^h(\rho^n)}{\delta t} + \kappa \frac{\mathcal{H}(\rho^n \mid \rho^{n+1})}{\delta t} = -\mathcal{D}^h(\rho^{n+1}) \,, \tag{28}$$

where $\mathcal{D}^h$ is the dissipation functional given by

$$\mathcal{D}^h(\rho^{n+1}) = \sum_K \sum_{L\sim K} \frac{\kappa |K{\scriptstyle|}L|}{d_{KL}} \, \alpha_\kappa\Big(\rho_K^{n+1}, \rho_L^{n+1}, d_{KL} q_{KL}^{n+1}\Big) \ge 0, \tag{29}$$

where $\alpha_\kappa : \mathbf{R}_+ \times \mathbf{R}_+ \times \mathbf{R} \to \mathbf{R}_+$ is given in terms of θ_κ from (18) by

$$\alpha_\kappa(a,b;v) = \Big(\log\Big(a e^{\frac{v}{2\kappa}}\Big) - \log\Big(b e^{-\frac{v}{2\kappa}}\Big)\Big)\, \theta_\kappa(a,b;v). \tag{30}$$

We remark that due to the nonnegativity of dissipation functional and relative entropy, the free energy functional is decreasing during the evolution,

$$\mathcal{F}^h(\rho^n) \le \mathcal{F}^h(\rho^0),$$

for any $n \in \mathbf{N}$.

Remark 2. *In comparison to (4), we have the additional term $\mathcal{H}(\rho^n \mid \rho^{n+1})$ providing a weak BV control on the discrete time gradient (cf. Lemma 3), which is typical for implicit schemes [50].*

To make the connection to (4) more apparent, we expand $\mathcal{D}^h$ using (20)

$$\begin{aligned}\mathcal{D}^h(\rho) =& \sum_K \sum_{L\sim K} \frac{|K|L|}{d_{KL}} \Big(\kappa(\log\rho_K - \log\rho_L) + d_{KL}q_{KL}\Big)\,\theta_k(a,b;d_{KL}q_{KL})\\ =& \sum_K \sum_{L\sim K} d_{KL}|K|L|\Big(\kappa\nabla_{KL}\log\rho + q_{LK}\Big)\Big(\kappa\nabla_{KL}\rho + \frac{\rho_K+\rho_L}{2}q_{LK} + O(d_{KL}|q_{KL}|^2)\Big),\end{aligned}$$

where we write $\nabla_{KL}\rho = \frac{\rho_L-\rho_K}{d_{KL}}$ for the discrete gradient. By recalling the form of q as convolution with W in (12), we expect in the continuum limit, that (28) becomes (4).

The numerical free energy dissipation principle raises the question, if the stability estimate (24) can be improved to one, which does not degenerate in the limit $h, \delta t \to 0$. The classical entropy method [58] entails that two solutions of the limit equation (1) are exponential stable in relative entropy and hence also in L^1. Proposition 1 shows that Scharfetter–Gummel discretization is structure preserving and we conjecture that the entropy method is also applicable in this case.

The stationary solutions to (9) are densities $\rho \in \mathcal{P}(\mathcal{T}^h)$ such that for all $K \in \mathcal{T}^h$ it holds

$$0 = \sum_{L\sim K} F_{KL} = \sum_{L\sim K} \tau_{KL}(j_{KL}[\rho] - j_{LK}[\rho]) = \sum_{L\sim K} \tau_{KL}\theta_\kappa(\rho_K,\rho_L;d_{KL}q_{KL}[\rho]), \tag{31}$$

where we used the identities (14) and (19) with $\rho^{n+1} = \rho^n = \rho$. The stationary states have the following characterization, which is completely analog to the situation for the aggregation-diffusion equation studied in [17, Proposition 2.4].

Theorem 2 (Characterization of stationary states)**.** *Let $h > 0$ and let $\rho \in \mathcal{P}(\mathcal{T}^h)$, then the following statements are equivalent:*

(i) *ρ is a stationary state satisfying (31);*
(ii) *ρ solves the Kirkwood–Monroe fixed point equation*

$$\rho_K = \frac{\exp(-\kappa^{-1}\sum_J |J|\rho_J W(x_K - x_J))}{Z^h(\rho)}, \tag{32}$$

for any $K \in \mathcal{T}^h$, where $Z^h(\rho) = \sum_{\tilde K} |\tilde K| \exp(-\kappa^{-1}\sum_J |J|\rho_J W(x_{\tilde K} - x_J))$;
(iii) *ρ is a critical point of the free energy functional $\mathcal{F}^h$ on $\mathcal{P}(\mathcal{T}^h)$;*
(iv) *ρ diminishes the dissipation, i.e. $\mathcal{D}^h(\rho) = 0$.*

Note, that the fixed point identity entails, in particular, that any stationary solution ρ is strictly positive.

With the characterization of stationary states and thanks to the free energy dissipation relation (28), we can establish the large time behavior of the scheme. More precisely, we show that for any initial datum, approximate solutions converge towards a stationary state. As a particular consequence, this result shows that the set of stationary states has to be non-empty.

Theorem 3 (Large time behavior of the scheme)**.** *Let $\{\rho_K^n\}_{K,n}$ be a solution to the Scharfetter–Gummel scheme (9), (11), (12). Let Π^h be the set of stationary solutions. Then ρ^n approaches Π^h in the large time limit,*

$$\lim_{n\to\infty} \operatorname{dist}_{\ell^1(\mathcal{T}^h)}(\rho^n, \Pi^h) = 0. \tag{33}$$

The statement (33) is equivalent to the existence of a stationary state $\bar{\rho}^h \in \Pi^h$ and of a subsequence n_k such that $\rho^{n_k} \to \bar{\rho}^h$ in $\ell^1(\mathcal{T}^h)$. A sufficient condition for the uniqueness of the limit would be that Π^h consists only of isolated stationary points. We leave a deeper understanding of Π^h in dependence of the interaction potential W for future work.

Given the discrete solution $\{\rho_K^n\}$ of a numerical finite volume scheme, we can introduce the *approximate solution* $\rho_{\delta t,h} \in L^1(\mathbf{R}_+ \times \hat{\Omega})$ of the PDE, given by

$$\rho_{\delta t,h} = \rho_K^n \quad \text{for a.e. } (t,x) \in \left[t^n, t^{n+1}\right) \times K.$$

The estimate (28) also carries a priori gradient information about the discrete gradient in the dissipation functional. Exploiting this, we are able to establish sufficient compactness of the discrete scheme in the limit $h \to 0$ to establish the convergence result.

Theorem 4 (Convergence)**.** *The approximate solution sequence $\{\rho_{\delta t,h}\}$ converges as $h \to 0$ in $L^2((0,T); L^1(\Omega))$ towards a distributional solution of the continuous problem. Moreover, any stationary solution of the Scharfetter–Gummel scheme converges in $L^1(\Omega)$ to a stationary solution of the continuous problem.*

2.4. **Related work.** A considerable amount of numerical schemes has been proposed and analyzed for advection-diffusion equations in general and aggregation-diffusion equations in particular. In the following, we give a (necessarily incomplete) list of references.

The class of schemes that we adapt in the present work goes back to the work of Scharfetter and Gummel in 1969 [49]. In this work, the authors' objective is to derive effective numerical methods for the simulation of semiconductors, where both advection and diffusion dominated regimes could occur. Further contributions for numerically solving semiconductor models include, for instance, mixed exponential methods, e.g., [9, 36, 37], or upwind finite volume schemes, e.g., [20, 21, 22]. For more references about the developments of schemes for drift-diffusion models, with particular background on semiconductor models, see, e.g., [32, 11].

In the context of aggregation-diffusion equations (1), there has been quite some research activity in recent years, with a special focus on structure-preserving properties. For instance, in [1, 2], a (semi-)implicit discretization of (1) and nonlinear variants based on upwind fluxes is proposed and shown to converge. In [44], the starting point for a scheme is formula (8), which is discretized by symmetric differences. This work shows positivity and free energy dissipation, but the scheme has the drawback that it is formulated for $\kappa = 1$ only, and we expect stability issues for such an approach for small diffusivity constants. We also mention Lagrangian particle approximations, e.g., [14], which is motivated by the Lagrangian particle interpretation of the aggregation equation [16]. For further references on this subject, we refer to the recent review [15].

In the present work, we consider linear diffusions, but also a generalization of the Scharfetter–Gummel scheme to aggregation equations with nonlinear diffusions seems possible. Related to this, we mention [29, 4], where advection-diffusion equations with nonlinear diffusions (and nonlinear fluxes) are studied. Notice that the work [4] is similar to ours in the sense that steady states and large time dynamics are investigated. (Here, the motivation is again a model for semiconductors.) Both works and the fact that many relevant nonlinear aggregation-diffusion equations own a free energy functional similar to (3) suggest that analysis of the finite free energy setting seems to be accessible also for nonlinear diffusions.

We want to mention that the theoretical investigation of the large time behavior of Scharfetter–Gummel-related schemes for advection-diffusion equation is contained in [12, 43] based on discrete function inequalities, which go back to [5]. Already before, the scheme in [19] uses a nonlinear approximation of the diffusion flux, which in the linear setting would essentially rely on the observation that the Laplacian can be interpreted as an advection operator with nonlinear flux according to the identity $\Delta\rho = \nabla\cdot(\rho\nabla\log\rho)$, and shows preservation of the large time behavior of the continuous model.

Concerning the rate of convergence for this and related schemes, we refer to the recent works [41, 25], which show for the upwind scheme applied to the pure aggregation equation, i.e., $\kappa = 0$, an explicit rate of order $\frac{1}{2}$ with respect to the Wasserstein distance. Since the proposed Scharfetter–Gummel scheme is a direct generalization to the case with diffusion of strength κ, we conjecture that a similar analysis is possible in this case. We leave this project for future research.

It is not apparent to us how the Scharfetter–Gummel scheme of the present work is connected to the gradient flow formulation (5). However, the free energy dissipation relation (28) suggests that there might exist such a formulation. As an example, the gradient flow formulation for Markov chains in [23] was motivated by the related upwind scheme and generalized to discrete McKean–Vlasov dynamics on graphs in [27]. Recently, the two works [13, 28] provide a gradient flow formulation for the aggregation equation on finite volumes. Besides, for linear Fokker–Planck equations, gradient flow formulations are obtained in [26, 34], but with no obvious way on how to generalize those to the nonlinear setting.

3. Proofs

3.1. **Well-posedness and first properties.**

In this subsection we provide the proof of Theorem 1.

The following two lemmas represent general results for the Scharfetter–Gummel scheme and do not rely on the specific choice of the driving vector field q_{KL} in (12). First, we show that the scheme is conservative.

Lemma 1. *Let $\rho_{\delta t,h}$ be an approximate solution to the Scharfetter–Gummel scheme (9) and (11) with antisymmetric driving vectorfield $\{q^n_{KL}\}_{K,L,n}$, that is $q^n_{KL} = -q^n_{KL}$ for all $K, L \in \mathcal{T}^h$ and $n \in \mathbf{N}_0$. Then $\rho_{\delta t,h}$ is mass preserving in the sense that*

$$\int_\Omega \rho_{\delta t,h}(t,x)\,\mathrm{d}x = \int_\Omega \rho^0_h(x)\,\mathrm{d}x, \tag{34}$$

for any $t > 0$.

Proof. The conservativity (34) of the scheme is a consequence of the antisymmetry (13), which implies by symmetrization

$$\sum_K \sum_{L\sim K} F_{KL}^{n+1} = \frac{1}{2}\sum_K \sum_{L\sim K}\Big(F_{KL}^{n+1} + F_{LK}^{n+1}\Big) = 0$$

and hence

$$\sum_K |K|\rho_K^{n+1} = \sum_K |K|\rho_K^n + \delta t \sum_K \sum_{L\sim K} F_{KL}^{n+1} = \sum_K |K|\rho_K^n.$$

The statement in (34) now follows by iteration and the fact that $\int_\Omega \rho_{\delta t,h}\,\mathrm{d}x = \sum_K |K|\rho_K^n$ for some $n \in \mathbf{N}_0$. ■

Our next goal is to show the positivity of the scheme.

Lemma 2. *Let $\rho_{\delta t,h}$ be an approximate solution to the Scharfetter–Gummel scheme (9) and (11) with antisymmetric driving vectorfield $\{q_{KL}^n\}_{K,L,n}$ of potential form, that is for some $\{V_K^n\}_{K,n}$ it holds*

$$d_{KL}q_{KL}^n = V_K^n - V_L^n \qquad \text{for all } K, L \in \mathcal{T} \text{ and } n \in \mathbf{N}_0.$$

If the initial datum ρ_h^0 is nonnegative, then ρ_K^n is positive for all $K \in \mathcal{T}^h$ and $n \in \mathbf{N}$.

Proof of Lemma 2. For the proof, it is convenient to consider instead of ρ_K^n, the following transformed quantity $h_K^n = \rho_K^n e^{\kappa^{-1}V_K^n}$. A short computation reveals that h_K^n solves the iteration scheme

$$\begin{aligned}|K|\frac{h_K^{n+1}e^{-\kappa^{-1}V_K^{n+1}} - h_K^n e^{-\kappa^{-1}V_K^n}}{\delta t} &= \sum_{L\sim K}\tau_{KL}\frac{V_K^{n+1} - V_L^{n+1}}{e^{\frac{1}{\kappa}V_K^{n+1}} - e^{\frac{1}{\kappa}V_{L^{n+1}}}}\Big(h_L^{n+1} - h_K^{n+1}\Big) \\ &= \kappa\sum_{L\sim K}\tau_{KL}\frac{h_L^{n+1} - h_K^{n+1}}{\Lambda\Big(e^{\frac{1}{\kappa}V_K^{n+1}}, e^{\frac{1}{\kappa}V_L^{n+1}}\Big)},\end{aligned} \tag{35}$$

where $\Lambda : \mathbf{R}_+ \times \mathbf{R}_+ \to \mathbf{R}_+$ is the logarithmic mean given by

$$\Lambda(a,b) = \begin{cases}\frac{a-b}{\log a - \log b}, & a, b > 0,\\ a, & a = b.\end{cases}$$

We observe that (35) is the implicit finite volume scheme for an elliptic linear operator and hence satisfying the maximum principle, that is, if $h_K^n \geq 0$ for all $K \in \mathcal{T}^h$, we immediately get $h_K^{n+1} \geq 0$ for all K. Indeed, suppose that for some n and $J \in \mathcal{T}^h$, it holds $h_J^{n+1} = \min_{K\in\mathcal{T}^h} h_K^{n+1} < 0$ and $h_K^n \geq 0$ for all $K \in \mathcal{T}^h$, then we have by (35),

$$0 > |J|\frac{h_J^{n+1}e^{-\kappa^{-1}V_J^{n+1}} - h_J^n e^{-\kappa^{-1}V_J^n}}{\delta t} = \kappa\sum_{L\sim J}\frac{\tau_{JL}}{\Lambda\Big(e^{\frac{1}{\kappa}V_J^{n+1}}, e^{\frac{1}{\kappa}V_L^{n+1}}\Big)}\Big(h_L^{n+1} - h_J^{n+1}\Big) \geq 0,$$

which is a contradiction.

To obtain the positivity, let $h_L^n \geq 0$ for all $L \in \mathcal{T}^h$ and $h_{L^*}^n > 0$ for some $L^* \in \mathcal{T}^h$. Moreover, towards a contradiction, we assume that $h_{K^*}^{n+1} = 0$ for some $K^* \in \mathcal{T}^h$. By using $K = K^*$ in (35), we immediately see that

$$\frac{|K|}{\delta t} h_{K^*}^n e^{-\frac{1}{\kappa} V_{K^*}^{n+1}} + \sum_{L\sim K^*} \tau_{KL} \frac{h_L^{n+1}}{\Lambda(e^{\frac{1}{\kappa}V_{K^*}^{n+1}}, e^{\frac{1}{\kappa}V_L^{n+1}})} = 0.$$

By the nonnegativity of $\{h_K^n\}_n$, this implies that $h_L^{n+1} = 0$ for any $L \sim K^*$ because the denominators above are finite. This argument can be reapplied to any $K \sim K^*$, and by iteration we arrive $h_K^{n+1} = 0$ for any $K \in \mathcal{T}^h$. Inserting this information back into (35), we deduce that $h_K^n = 0$ for any $K \in \mathcal{T}^h$, which contradicts the hypothesis that $h_{L^*}^n > 0$. ■

It remains to show that the scheme is well-defined. For this, we recall that in view of the formulation (15) of the scheme and definition (18), the scheme can be written in the form

$$\rho_K^{n+1} = \rho_K^n - \frac{\delta t}{|K|} \sum_{L\sim K} \tau_{KL} \theta_\kappa\Big(\rho_K^{n+1}, \rho_L^{n+1}; d_{KL} q_{KL}^{n+1}\Big). \tag{36}$$

Proof of Theorem 1. We show existence with the help of a fixed point argument. For this purpose, motivated by (36), we consider the function

$$\Xi_K[\varphi] = \rho_K^n - \frac{\delta t}{|K|} \sum_{L\sim K} \tau_{KL} \theta_\kappa \left(\varphi_K, \varphi_L; d_{KL} q_{KL}^{n+1}[\varphi]\right),$$

for any $\varphi = \{\varphi_K\}_{K\in\mathcal{T}}$ and a given probability distribution $\{\rho_K^n\}_{K\in\mathcal{T}}$. Observe that a fixed point of the function $\Xi = \{\Xi_K\}_K$ solves the Scharfetter–Gummel scheme. Our goal is to find a fixed point in the set $M = \Big\{\varphi = \{\varphi_K\}_K :\ \|\varphi\|_{L^1(\hat\Omega)} \le 2\Big\}$.

First, we show that for sufficiently small time steps, Ξ maps M into M. Using (21) and the triangle inequality, observe that

$$\begin{aligned} \|\Xi[\varphi]\|_{L^1(\hat\Omega)} &\le \|\rho^n\|_{L^1(\hat\Omega)} + \delta t \sum_K \sum_{L\sim K} \tau_{KL} |\theta_\kappa(\varphi_K, \varphi_L; 0)| \\ &\quad + \delta t \sum_K \sum_{L\sim K} \tau_{KL} \max\{|\varphi_K|, |\varphi_L|\}\, d_{KL} \Big|q_{KL}^{n+1}[\varphi]\Big|. \end{aligned} \tag{37}$$

By Remark 1, we have that $|\theta_\kappa(\varphi_K, \varphi_L; 0)| \le \kappa|\varphi_K - \varphi_L| \le \kappa\,(|\varphi_K| + |\varphi_L|)$, and thus, by relabeling

$$\delta t \sum_K \sum_{L\sim K} \tau_{KL} |\theta_\kappa(\varphi_K, \varphi_L; 0)| \le \frac{2\kappa\,\delta t}{d_{\min}} \sum_K \sum_{L\sim K} |K|L||\varphi_K| = \frac{2\kappa\,\delta t}{d_{\min}} \sum_K |\partial K|\,|\varphi_K|.$$

Thanks to the isoperimetric property (22), the latter is controlled by the L^1 norm of φ,

$$\delta t \sum_K \sum_{L\sim K} \tau_{KL} |\theta_\kappa(\varphi_K, \varphi_L; 0)| \le \frac{2C_{\mathrm{iso}}\kappa\,\delta t}{h d_{\min}} \|\varphi\|_{L^1(\hat\Omega)}.$$

For $\varphi \in M$ and δt small in the sense of (23), the right-hand side is bounded by $1/2$. The first two terms in (37) are thus controlled in total by $3/2$.

We turn to the estimate of the third expression in (37). Notice that, by the Lipschitz assumption on W in (A3), we have

$$\Big|q_{KL}^{n+1}[\varphi]\Big| \le \frac{1}{2} \sum_J |J|\,(|\varphi_J| + \rho_J^n)\,\mathrm{Lip}(W)$$

$$= \frac{1}{2}\left(\|\varphi\|_{L^1(\hat\Omega)} + \|\rho^n\|_{L^1(\hat\Omega)}\right)\operatorname{Lip}(W) \le \frac{3}{2}\operatorname{Lip}(W). \tag{38}$$

Therefore,

$$\delta t\sum_K\sum_{L\sim K}\tau_{KL}\max\Big\{|\varphi_K|,|\varphi_L|\Big\}d_{KL}\left|q_{KL}^{n+1}[\varphi]\right| \le \frac{3}{2}\operatorname{Lip}(W)\,\delta t\sum_K\sum_{L\sim K}|K|L|\Big(|\varphi_K|+|\varphi_L|\Big).$$

Similarly as above, relabeling and applying the isoperimetric property (22) yields

$$\delta t\sum_K\sum_{L\sim K}\tau_{KL}\max\Big\{|\varphi_K|,|\varphi_L|\Big\}d_{KL}\left|q_{KL}^{n+1}[\varphi]\right| \le \frac{3C_{\text{iso}}\operatorname{Lip}(W)\,\delta t}{h}\|\varphi\|_{L^1(\hat\Omega)}.$$

Hence, if δt is small as in (23) and $\varphi\in M$, the right-hand side is bounded by $1/2$. Plugging these estimates into (37), then yields that

$$\Big\|\Xi[\varphi]\Big\|_{L^1(\hat\Omega)} \le 1+\frac{1}{2}+\frac{1}{2}\le 2,$$

and thus $\Xi[\varphi]\in M$, as desired.

We will now establish that Ξ is a contraction on M. Let $\varphi,\psi\in M$ be given. Then, using the triangle inequality and (21), we find using $\lambda(s)=|s|\frac{e^s}{|e^s-e^{-s}|}$ that

$$\begin{aligned}\Big\|\Xi[\varphi]-\Xi[\psi]\Big\|_{L^1(\hat\Omega)} &\le \delta t\sum_K\sum_{L\sim K}\tau_{KL}\left|\theta_\kappa\Big(\varphi_K,\varphi_L;d_{KL}q_{KL}^{n+1}[\varphi]\Big)-\theta_\kappa\Big(\psi_K,\psi_L;d_{KL}q_{KL}^{n+1}[\psi]\Big)\right|\\
&\le 2\kappa\,\delta t\sum_K\sum_{L\sim K}\tau_{KL}\lambda\left(\frac{d_{KL}q_{KL}^{n+1}[\varphi]}{2\kappa}\right)|\varphi_K-\psi_K|\\
&\quad+2\kappa\,\delta t\sum_K\sum_{L\sim K}\tau_{KL}\lambda\left(-\frac{d_{KL}q_{KL}^{n+1}[\varphi]}{2\kappa}\right)|\varphi_L-\psi_L|\\
&\quad+\delta t\sum_K\sum_{L\sim K}\tau_{KL}\max\Big\{|\psi_K|,|\psi_L|\Big\}d_{KL}\Big|q_{KL}^{n+1}[\varphi]-q_{KL}^{n+1}[\psi]\Big|.\end{aligned}$$

Because $\lambda(s)\le|s|+\frac{1}{2}$ for any $s\in\mathbf{R}$, (38) and

$$\left|q_{KL}^{n+1}[\varphi]-q_{KL}^{n+1}[\psi]\right| \le \frac{1}{2}\operatorname{Lip}(W)\sum_J|J|\,|\varphi_J-\psi_J| = \frac{1}{2}\operatorname{Lip}(W)\|\varphi-\psi\|_{L^1(\hat\Omega)},$$

the latter can be rewritten as

$$\begin{aligned}\|\Xi[\varphi]-\Xi[\psi]\|_{L^1(\hat\Omega)} &\le \delta t\sum_K\sum_{L\sim K}\tau_{KL}\left(\frac{3}{2}d_{KL}\operatorname{Lip}(W)+\kappa\right)|\varphi_K-\psi_K|\\
&\quad+\delta t\sum_K\sum_{L\sim K}\tau_{KL}\left(\frac{3}{2}d_{KL}\operatorname{Lip}(W)+\kappa\right)|\varphi_L-\psi_L|\\
&\quad+\frac{1}{2}\operatorname{Lip}(W)\,\delta t\sum_K\sum_{L\sim K}|K|L|\Big(|\psi_K|+|\psi_L|\Big)\|\varphi-\psi\|_{L^1(\hat\Omega)}.\end{aligned}$$

Relabeling and using the isoperimetric property (22) and $\psi\in M$ gives

$$\begin{aligned}\Big\|\Xi[\varphi]-\Xi[\psi]\Big\|_{L^1(\hat\Omega)} &\le 3\operatorname{Lip}(W)\,\delta t\sum_K|\partial K|\,|\varphi_K-\psi_K|\\
&\quad+2\frac{\kappa\,\delta t}{d_{\min}}\sum_K|\partial K|\,|\varphi_K-\psi_K|\end{aligned}$$

$$+ \operatorname{Lip}(W)\,\delta t \sum_K |\partial K|\,|\psi_K|\,\|\varphi-\psi\|_{L^1(\hat\Omega)}$$
$$\le \left(\frac{5C_{\mathrm{iso}}\operatorname{Lip}(W)\,\delta t}{h} + \frac{2C_{\mathrm{iso}}\kappa\,\delta t}{h\,d_{\min}}\right)\|\varphi-\psi\|_{L^1(\hat\Omega)}.$$

Under the smallness assumption (23) on the time step size, the latter implies the contraction estimate

$$\|\Xi[\varphi]-\Xi[\psi]\|_{L^1(\hat\Omega)} \le \frac{2}{3}\|\varphi-\psi\|_{L^1(\hat\Omega)}. \tag{39}$$

By Banach's fixed point theorem, there exists a unique fixed point $\rho^{n+1}=\{\rho_K^{n+1}\}_K$ in M.

It remains to argue that this solution is indeed the unique solution. For this, we make use of the previous results in Lemmas 1 and 2. Thus, if $\tilde\rho_{\delta t,h}$ is a solution, we deduce from Lemmas 1 and 2 that $\tilde\rho_{\delta t,h}$ is a positive probability distribution for all $t>0$. In particular, by induction, for any n, $\tilde\rho^{n+1}$ is a fixed point of Ξ in M. By uniqueness $\rho^{n+1}=\tilde\rho^{n+1}$.

Finally, by considering the fixed point map $\tilde\Xi$ associated to the solution $\tilde\rho^n$ with initial value $\tilde\rho^0$, we get with (39) the bound

$$\|\rho^{n+1}-\tilde\rho^{n+1}\|_{L^1(\hat\Omega)} \le 3\|\rho^n-\tilde\rho^n\|_{L^1(\hat\Omega)}.$$

An iteration gives the stability estimate (24). ∎

3.2. **Free energy dissipation principle and consequences.**

Proof of Proposition 1. By the result of Lemma 2, we have that $\rho_K^{n+1}>0$ for all $K\in\mathcal{T}^h$, hence all occurrences of $\log\rho_K^{n+1}$ are well-defined. We start with the entropy and use the convention that $0\cdot\log 0 = 0 = 0\cdot\log\infty$ to arrive at

$$\frac{\mathcal{S}^h(\rho^{n+1})-\mathcal{S}^h(\rho^n)}{\delta t} = \sum_K \log\!\left(\rho_K^{n+1}\right)|K|\frac{\rho_K^{n+1}-\rho_K^n}{\delta t} + \frac{1}{\delta t}\sum_K |K|\rho_K^n\log\frac{\rho_K^{n+1}}{\rho_K^n}\,.$$

We proceed by recalling the definition of relative entropy from (27) and use the scheme (9) with identity (14) and a symmetrization

$$\frac{\mathcal{S}^h(\rho^{n+1})-\mathcal{S}^h(\rho^n)+\mathcal{H}(\rho^n\mid\rho^{n+1})}{\delta t} = -\sum_K\sum_{L\sim K}\log\rho_K^{n+1}F_{KL}^{n+1}$$
$$= -\frac12\sum_K\sum_{L\sim K}|K|L|\,\frac{\log\rho_K^{n+1}-\log\rho_L^{n+1}}{d_{KL}}\left(j_{KL}^{n+1}-j_{LK}^{n+1}\right).$$

Now, we turn to the discrete interaction energy (26). Here, we use the following discrete product rule for $K,J\in\mathcal{T}^h$

$$\rho_K^{n+1}\rho_J^{n+1}-\rho_K^n\rho_J^n = \left(\rho_K^{n+1}-\rho_K^n\right)\frac{\rho_J^{n+1}+\rho_J^n}{2} + \left(\rho_J^{n+1}-\rho_J^n\right)\frac{\rho_K^{n+1}+\rho_K^n}{2}\,,$$

and the symmetry of W in (A1) to write

$$\frac{\mathcal{E}^h(\rho^{n+1})-\mathcal{E}^h(\rho^n)}{\delta t} = \sum_{K,J} W(x_J-x_K)|J|\frac{\rho_J^{n+1}+\rho_J^n}{2}|K|\frac{\rho_K^{n+1}-\rho_K^n}{\delta t}.$$

Applying the scheme (9) with (14) and another symmetrization, we arrive at

$$\frac{\mathcal{E}^h(\rho^{n+1})-\mathcal{E}^h(\rho^n)}{\delta t}$$

$$
\begin{aligned}
&= -\sum_{J,K}\sum_{L\sim K} W(x_J - x_K)|J|\frac{\rho_J^{n+1}+\rho_J^n}{2}\frac{|K|L|}{d_{KL}}\Big(j_{KL}^{n+1} - j_{LK}^{n+1}\Big)\\
&= -\frac{1}{2}\sum_{K}\sum_{L\sim K}|K|L|\Big(j_{KL}^{n+1} - j_{LK}^{n+1}\Big)\sum_J |J|\frac{\rho_J^{n+1}+\rho_J^n}{2}\frac{W(x_J - x_K) - W(x_J - x_L)}{d_{KL}}\\
&= -\frac{1}{2}\sum_{K}\sum_{L\sim K}|K|L|\, q_{KL}^{n+1}\left(j_{KL}^{n+1} - j_{LK}^{n+1}\right),
\end{aligned}
$$

by definition (12).

Finally, we combine both calculations and invoke (15) to prove (28),

$$
\begin{aligned}
&\frac{\mathcal{F}^h(\rho^{n+1}) - \mathcal{F}^h(\rho^n) + \mathcal{H}(\rho^n \mid \rho^{n+1})}{\delta t}\\
&\qquad = -\frac{1}{2}\sum_K\sum_{L\sim K}|K|L|\,\frac{\kappa\log\rho_K^{n+1} - \kappa\log\rho_L^{n+1} + d_{KL}q_{KL}^{n+1}}{d_{KL}}\left(j_{KL}^{n+1} - j_{LK}^{n+1}\right)\\
&\qquad = -\frac{1}{2}\sum_K\sum_{L\sim K}\kappa|K|L|\,\frac{\log\left(\rho_K^{n+1}e^{\frac{d_{KL}q_{KL}^{n+1}}{2\kappa}}\right) - \log\left(\rho_L^{n+1}e^{-\frac{d_{KL}q_{KL}^{n+1}}{2\kappa}}\right)}{d_{KL}}\times\\
&\qquad\qquad\qquad\times d_{KL}q_{KL}^{n+1}\frac{\rho_K^{n+1}e^{\frac{d_{KL}q_{KL}^{n+1}}{2\kappa}} - \rho_L^{n+1}e^{-\frac{d_{KL}q_{KL}^{n+1}}{2\kappa}}}{e^{\frac{d_{KL}q_{KL}^{n+1}}{2\kappa}} - e^{-\frac{d_{KL}q_{KL}^{n+1}}{2\kappa}}}.
\end{aligned}
$$

This concludes the proof of Proposition 1. ■

Before turning to the proof of Theorem 2, we list some auxiliary elementary properties of the function α_κ occurring in the definition of the dissipation (29) and following from the ones of θ_κ in Remark 1.

Remark 3. *The function α_κ from (30) satisfies the following properties:*

- *(i) For any $v \in \mathbf{R}$ is $\alpha_\kappa(\cdot,\cdot;v)$ jointly convex on $\mathbf{R}_+\times\mathbf{R}_+$.*
- *(ii) It holds $\alpha_\kappa(a,b,r) = 0$ if and only if $v = -\kappa\log\frac{a}{b}$.*
- *(iii) It holds $\alpha_\kappa(a,b;0) = \kappa(a-b)\log\frac{a}{b}$.*
- *(iv) It holds $\lim_{\kappa\to 0}\kappa\alpha_\kappa(a,b;v) = a(v_+)^2 + b(v_-)^2$.*

Proof of Theorem 2. Let ρ be a solution to (32). Then it holds for any $K\sim L$

$$
\frac{\rho_K}{\rho_L} = \exp\left(-\kappa^{-1}\sum_J |J|\rho_J\Big(W(x_K - x_J) - W(x_K - x_J)\Big)\right) = \exp\Big(-\kappa^{-1}d_{KL}q_{KL}[\rho]\Big).
$$

Hence, we have that $\theta_\kappa(\rho_K,\rho_L,;d_{KL}q_{KL}[\rho]) = 0$ by Remark 1(iii), showing that (ii) implies (i).

Now, let ρ be a solution to (31). We introduce the notation

$$
\Phi_K[\rho] = \frac{\exp(-\kappa^{-1}\sum_J|J|\rho_J W(x_K - x_J))}{\sum_{\tilde K}|\tilde K|\exp(-\kappa^{-1}\sum_J|J|\rho_J W(x_{\tilde K} - x_J))}.
$$

Our goal is to show that ρ solves the fixed point equation (32), i.e., $\rho_K = \Phi_K[\rho]$. Notice first that by using the following auxiliary relations

$$d_{KL}q_{KL}[\rho] = -\kappa \log \frac{\Phi_K[\rho]}{\Phi_L[\rho]} \quad \text{and} \quad \exp\left(\frac{d_{KL}}{2\kappa} q_{KL}[\rho]\right) = \sqrt{\frac{\Phi_L[\rho]}{\Phi_K[\rho]}},$$

we may write for any $\{h_K\}_K$, similarly as in the proof of Lemma 2, that

$$\begin{aligned}\theta_\kappa\Big(h_K\Phi_K[\rho], h_L\Phi_L[\rho]; d_{KL}q_{KL}(\rho)\Big) &= d_{KL}q_{KL}[\rho]\frac{h_k\sqrt{\Phi_K[\rho]\Phi_L[\rho]} - h_L\sqrt{\Phi_K[\rho]\Phi_L[\rho]}}{\sqrt{\frac{\Phi_L[\rho]}{\Phi_K[\rho]}} - \sqrt{\frac{\Phi_K[\rho]}{\Phi_L[\rho]}}} \\ &= -\kappa(h_K - h_L)\frac{\log \Phi_K[\rho] - \log \Phi_L[\rho]}{\frac{1}{\Phi_K[\rho]} - \frac{1}{\Phi_L[\rho]}} \\ &= \kappa(h_K - h_L)\frac{1}{\Lambda\Big(\frac{1}{\Phi_K[\rho]}, \frac{1}{\Phi_L[\rho]}\Big)}.\end{aligned}$$

Therefore, supposing that $\rho_K = h_K\Phi_K[\rho]$ for some reals h_K, which is possible thanks to the positivity of $\Phi_K[\rho]$, equation (31) translates into

$$0 = \sum_{L\sim K} \tau_{KL}\frac{h_K - h_L}{\Lambda(\Phi_K[\rho]^{-1}, \Phi_L[\rho]^{-1})}.$$

Testing this equation by some φ_K and doing a summation by parts, we arrive at the weak formulation

$$0 = \sum_K \sum_{L\sim K} \tau_{KL}\frac{(\varphi_K - \varphi_L)(h_K - h_L)}{\Lambda(\Phi_K[\rho]^{-1}, \Phi_L[\rho]^{-1})}.$$

By choosing $\varphi_K = h_K$ and recalling that the logarithmic mean Λ is positive because the $\Phi_K[\rho]$'s are positive, we observe that $\{h_K\}_K$ has to be constant, that means, $h_K = C$ for some $C \in \mathbf{R}$. In particular, it holds that $\rho_K = C\Phi_K[\rho]$, for any cell K, and since both $\{\rho_K\}_K$ and $\{\Phi_K[\rho]\}_K$ are probability distributions, we must have $h_K = C = 1$, which is want we aimed to prove. We have thus showed that (i) and (ii) are equivalent.

Now, we consider the variation of $\mathcal{F}^h$ from (25) on $\mathcal{P}(\mathcal{T}^h)$. For this fix $s : \mathcal{T}^h \to \mathbf{R}$ with $\sum_K |K|s_K = 0$ and consider the variation

$$\lim_{\varepsilon\to 0} \frac{\mathcal{F}^h(\rho + \varepsilon\, s) - \mathcal{F}^h(\rho)}{\varepsilon} = \sum_K |K| D_K\mathcal{F}^h(\rho)s_K$$

Hence, we can only identify $D\mathcal{F}^h$ with a vector upto a constant $c \in \mathbf{R}$ as

$$D_K\mathcal{F}^h(\rho) = \kappa \log \rho_K + \sum_L |L|W(x_K - x_L)\rho_L + c.$$

Since, $\rho \in \mathcal{P}(\mathcal{T}^h)$, we immediately recover $c = \kappa \log Z^h(\rho)$ as in (ii). Hence, each probability distribution ρ on $\mathcal{P}(\mathcal{T}^h)$ satisfying (32) is a critical point of $\mathcal{F}^h$ and vise versa, showing the equivalence of (ii) and (iii).

Likewise, we have from Remark 1 and Remark 3 that $\alpha_\kappa(\rho_K, \rho_L, d_{KL}q_{KL}[\rho]) = 0$ if and only if $\theta_\kappa(\rho_K, \rho_L, d_{KL}q_{KL}[\rho]) = 0$, which implies by (31) that ρ is a stationary solution if and only if $\mathcal{D}^h(\rho) = 0$, showing the equivalence of (i) and (iv). ∎

Proof of Theorem 3. The proof of the large time behavior follows along a standard argument from the theory of dynamical systems (see for instance [54, Section 6]). We consider $\mathcal{P}(\mathcal{T}^h) \subset \ell^1(\mathcal{T}^h)$ as a convex compact subset endowed with the $\ell^1(\mathcal{T}^h)$ topology. By the global well-posedness for the scheme in $\mathcal{P}(\mathcal{T}^h)$ from Theorem 1, we can define the ω-limit set for any $\rho^0 \in \mathcal{P}(\mathcal{T}^h)$ given by

$$\omega(\rho^0) = \Big\{\mu \in \mathcal{P}(\mathcal{T}^h) : \rho^{n_k} \to \mu \text{ in } \ell^1(\mathcal{T}^h) \text{ for some subsequence } n_k \to \infty\Big\}.$$

Since the scheme leaves $\mathcal{P}(\mathcal{T}^h)$ invariant, each positive orbit $\mathcal{O}^+(\rho^0) = \bigcup_{n\geq 0} \rho^n$ is compact in $\mathcal{P}(\mathcal{T}^h)$ and we find a convergent subsequence in $\mathcal{P}(\mathcal{T}^h)$, showing that ω-limit is non-empty. Since the scheme is by the stability estimate (24) in particular continuous in $\ell^1(\mathcal{T}^h)$, we obtain by [54, Lemma 6.5] that any ω-limit set $\omega(\rho_0)$ is positive invariant, that is, for $\mu \in \omega(\rho^0)$ follows that $\mathcal{O}^+(\mu) \subseteq \omega(\rho^0)$. The compactness also implies by standard arguments that $\operatorname{dist}_{\ell^1(\mathcal{T}^h)}(\rho^n, \omega(\rho^0)) \to 0$ as $t \to \infty$ (see [54, Lemma 6.7]).

Hence, it is left to characterize the ω-limit. To do so, we note that the free energy $\mathcal{F}^h$ is bounded from below on $\mathcal{P}(\mathcal{T}^h)$, since for any $\rho \in \mathcal{P}(\mathcal{T}^h)$ we have $\mathcal{S}(\rho) \geq 0$ and since W is bounded from below on Ω, it holds

$$\mathcal{F}^h(\rho) \geq -\|W\|_\infty.$$

By the free energy dissipation relation (28) we have that $\mathcal{F}^h(\rho^n)$ is also monotone decreasing along the scheme, hence $\mathcal{F}^h(\rho^n) \to \mathcal{F}^\infty \in \mathbf{R}$. In particular, for any $\overline{\rho}^0 \in \omega(\rho^0)$ holds $\mathcal{F}^h(\overline{\rho}^0) = \mathcal{F}^\infty$. Starting the scheme from any $\overline{\rho}^0 \in \omega(\rho^0)$, we have from (28) for any $N \in \mathbf{N}$ the identity

$$\mathcal{F}^\infty + \kappa \sum_{n=0}^{N} \mathcal{H}(\overline{\rho}^{n+1}|\overline{\rho}^n) + \delta t \sum_{n=1}^{N} \mathcal{D}^h(\overline{\rho}^n) = \mathcal{F}^\infty,$$

which implies that $\mathcal{H}(\overline{\rho}^{n+1}|\overline{\rho}^n) = 0 = \mathcal{D}^h(\overline{\rho}^{n+1})$ for any $n \in \mathbf{N}_0$, since both, relative entropy and dissipation are nonnegative. Hence, $\omega(\rho^0)$ consists of elements $\overline{\rho}^0 \in \mathcal{P}(\mathcal{T}^h)$ satisfying $\mathcal{D}^h(\overline{\rho}^0) = 0$, implying that any $\overline{\rho}^0 \in \omega(\rho^0)$ is a stationary state by Theorem 2. ■

3.3. **Convergence of the scheme.** In this section, we finally turn to the proof of the convergence of the Scharfetter–Gummel scheme, Theorem 4. In particular, we may assume throughout this subsection that h is small in the sense that

$$h \leq \frac{\kappa}{\operatorname{Lip}(W)}. \tag{40}$$

In a first step, we have to establish estimates on the discrete temporal and spatial gradients.

Lemma 3 (Gradient estimates). *The following estimates hold true for any $N \in \mathbf{N}$,*

$$\kappa^2\, \delta t \sum_{n=0}^{N} \sum_{K} \sum_{L\sim K} \frac{|K|L|}{d_{KL}} \Big(\sqrt{\rho_K^{n+1}} - \sqrt{\rho_L^{n+1}}\Big)^2 \lesssim \mathcal{F}^h(\rho^0) + T \operatorname{Lip}(W)^2, \tag{41}$$

$$\kappa^2\, \delta t \sum_{n=0}^{N} \bigg(\sum_{K} \sum_{L\sim K} |K|L|\, |\rho_K^{n+1} - \rho_L^{n+1}|\bigg)^2 \lesssim \mathcal{F}^h(\rho^0) + T \operatorname{Lip}(W)^2, \tag{42}$$

$$\kappa \sum_{n=0}^{N} \bigg(\sum_{K} |K|\, |\rho_K^{n+1} - \rho_K^n|\bigg)^2 \leq \mathcal{F}^h(\rho^0). \tag{43}$$

The reader who is familiar with the entropy theory for the heat equation will identify our first estimate (41) as a discrete version of the entropy bound on the Fisher information, which is a consequence of the dissipation estimate (4). On the discrete level, these estimates rely on Proposition 1.

We furthermore remark that the gradient estimates (41) and (42) make use of the smallness condition in (40). We emphasize that this is not a condition for well-posedness or stability of the scheme. Moreover, for very small diffusivities for which (40) is violated, a variation of the analysis below would yield the weak BV estimate

$$\sum_K \sum_{L\sim K} |K_{|}L|\,|q_{KL}^{n+1}|\,|\rho_K^{n+1}-\rho_L^{n+1}| \le \sum_K \sum_{L\sim K} |K_{|}L|\Big(\rho_K^{n+1}(q_{KL}^{n+1})^+ + \rho_L^{n+1}(q_{KL}^{n+1})^-\Big) \lesssim 1,$$

which is known for the upwind scheme [50].

Proof. We start by noticing that, because $(\log(a)-\log(b))(a-b)(a+b)\ge 2(a-b)^2$ for any positive $a,b>0$ by the concavity of the logarithm, we have

$$\alpha_\kappa(a,b,v) \ge \frac{(ae^{\frac{v}{2\kappa}} - be^{-\frac{v}{2\kappa}})^2}{ae^{\frac{v}{2\kappa}} + be^{-\frac{v}{2\kappa}}}\frac{v}{e^{\frac{v}{2\kappa}} - e^{-\frac{v}{2\kappa}}}.$$

For notational simplicity, we write $z=\frac{v}{2\kappa}$ and suppose that $|z|\le 1$ in the following, which we will justify later. We furthermore set

$$\phi(z) = \frac{(ae^z - be^{-z})^2}{ae^z + be^z},$$

and notice that

$$\begin{aligned}
\phi(z) &\ge \phi(0) + \phi'(0)z - \frac{1}{2}\max_{z\in[0,1]}|\phi''(z)|z^2\\
&\ge \frac{(a-b)^2}{a+b} + 2(a-b)z - \frac{(a-b)^3}{(a+b)^2}z - C_1(a+b)z^2\\
&\ge \frac{1}{2}\frac{(a-b)^2}{a+b} - C_2(a+b)z^2,
\end{aligned}$$

for some constant $C_1, C_2>0$ by Taylor expansion and the elementary inequality $a\,b\le \varepsilon a^2 + \frac{1}{4\varepsilon}b^2$, for any $a,b,\varepsilon>0$.

For $v = d_{KL}q_{KL}^{n+1}$ and our smallness assumption (40), we indeed have that $|z|\le 1$. Truly, in view of the definition of q_{KL}^{n+1} in (12) and because the approximate solution is a probability measure, it holds that $|z| = d_{KL}|q_{KL}^{n+1}|/2\kappa \le \frac{h}{\kappa}\operatorname{Lip}(W)\le 1$. Since $|e^z - e^{-z}|\lesssim |z|$ for $|z|\le 1$, we thus infer from Proposition 1 and the nonnegativity of the relative entropy $\mathcal{H}(\rho^{n+1},\rho^n)$, see the discussion below, that

$$\begin{aligned}
\frac{\mathcal{F}^h(\rho^{n+1})-\mathcal{F}^h(\rho^n)}{\delta t} &+ \kappa^2\sum_K\sum_{L\sim K}\frac{|K_{|}L|}{d_{KL}}\frac{(\rho_K^{n+1}-\rho_L^{n+1})^2}{\rho_K^{n+1}+\rho_L^{n+1}}\\
&\lesssim \sum_K\sum_{L\sim K}\frac{|K_{|}L|}{d_{KL}}(\rho_K^{n+1}+\rho_L^{n+1})\left|d_{KL}q_{KL}^{n+1}\right|^2.
\end{aligned}$$

Notice that the term on the right-hand side is bounded by

$$\operatorname{Lip}(W)^2\sum_K\sum_{L\sim K} h|K_{|}L|(\rho_K^{n+1}+\rho_L^{n+1}) \lesssim (\operatorname{Lip}(W))^2\|\rho^{n+1}\|_{L^1(\Omega)} = \operatorname{Lip}(W)^2,$$

where we have used relabeling arguments, the isoperimetric property (22), and a similar reasoning as above in order to bound q_{KL}^{n+1}. Hence, summing over n and multiplying by δt gives

$$\kappa^2\,\delta t\sum_{n=0}^{N}\sum_K\sum_{L\sim K}\frac{|K{\scriptstyle|}L|}{d_{KL}}\frac{(\rho_K^{n+1}-\rho_L^{n+1})^2}{\rho_K^{n+1}+\rho_L^{n+1}}\lesssim \mathcal{F}^h(\rho_0)+T\,\mathrm{Lip}(W)^2.$$

It remains to apply the elementary inequality $(\sqrt{a}-\sqrt{b})^2\le\frac{(a-b)^2}{a+b}$ to deduce (41).

In order to prove (42), we have to use the Cauchy–Schwarz inequality,

$$\begin{aligned}&\kappa^2\,\delta t\sum_{n=0}^{N}\left(\sum_K\sum_{L\sim K}|K{\scriptstyle|}L|\,|\rho_K^{n+1}-\rho_L^{n+1}|\right)^2\\&\quad\le\kappa^2\,\delta t\sum_{n=0}^{N}\left(\sum_K\sum_{L\sim K}\frac{|K{\scriptstyle|}L|}{d_{KL}}\frac{(\rho_K^{n+1}-\rho_L^{n+1})^2}{\rho_K^{n+1}+\rho_L^{n+1}}\right)\left(\sum_K\sum_{L\sim K}|K{\scriptstyle|}L|d_{KL}(\rho_K^{n+1}+\rho_L^{n+1})\right),\end{aligned}$$

and notice that

$$\sum_K\sum_{L\sim K}|K{\scriptstyle|}L|d_{KL}(\rho_K^{n+1}+\rho_L^{n+1})\lesssim 1,$$

by a standard relabeling argument, using the isoperimetric property of the mesh (22) and the fact that the approximate solution is a probability distribution. This proves (42).

The estimate of the time derivative in (43) is an immediate consequence of the entropy dissipation (28) in Proposition 1. Indeed, dropping the dissipation term on the right-hand side of (28), summing over n and using the Pinsker inequality yields

$$\frac{\kappa}{2}\sum_{n=0}^{N}\left(\sum_K|K|\,|\rho_K^{n+1}-\rho_K^n|\right)^2\le\kappa\sum_{n=0}^{N}\mathcal{H}(\rho^{n+1},\rho^n)\le\mathcal{F}^h(\rho^0).$$

This concludes the proof of Lemma 3. ■

The previous lemma provides bounds on gradients in terms of the discrete free energy of the initial datum $\rho^0=\rho_h^0$. Our next lemma ensures that this quantity is indeed bounded for finite entropy initial data ρ_0.

Lemma 4 (Initial energy)**.** *It holds that*

$$\mathcal{F}^h(\rho_h^0)\le\mathcal{F}(\rho_0)+h\,\mathrm{Lip}(W). \tag{44}$$

Proof. From the convexity of the function $\varphi(z)=z\log z$, we obtain via Jensen's inequality that

$$\mathcal{S}^h(\rho_h^0)=\sum_K|K|\varphi\left(\frac{1}{|K|}\int_K\rho_0\,\mathrm{d}x\right)\le\sum_K\int_K\varphi(\rho_0)\,\mathrm{d}x=\int_\Omega\varphi(\rho_0)\,\mathrm{d}x=\mathcal{S}(\rho_0).$$

Moreover, using the definitions of the interaction energy and the Lipschitz property of the potential W,

$$\begin{aligned}\mathcal{E}^h(\rho_h^0)&=\frac12\sum_{K,L}\int_K\int_L W(x_K-x_L)\rho_0(x)\rho_0(y)\,\mathrm{d}x\,\mathrm{d}y\\&\le\frac12\int_\Omega\int_\Omega W(x-y)\rho_0(x)\rho_0(y)\,\mathrm{d}x\,\mathrm{d}y+h\,\mathrm{Lip}(W)\sum_{K,L}\int_K\rho_0\,\mathrm{d}x\int_L\rho_0\,\mathrm{d}y\\&=\mathcal{E}(\rho_0)+h\,\mathrm{Lip}(W).\end{aligned}$$

A combination of both estimates yields the statement of the lemma. ■

In the next step, we translate the gradient bounds into continuity estimates for the approximate solution.

Lemma 5 (Translations)**.** *For any $\tau > 0$, it holds that*

$$\int_0^T \left(\int_\Omega |\rho_{h,\delta t}(t+\tau, x) - \rho_{h,\delta t}(t,x)| \,\mathrm{d}x \right)^2 \mathrm{d}t \lesssim \tau, \tag{45}$$

uniformly in h. Moreover, for $\eta \in \mathbf{R}^d$, it holds that

$$\int_0^T \left(\int_{\mathbf{R}^d} |\rho_{h,\delta t}(t, x+\eta) - \rho_{h,\delta t}(t,x)| \,\mathrm{d}x \right)^2 \mathrm{d}t = o(1), \tag{46}$$

as $|\eta| \to 0$, uniformly in h.

In the second statement of the lemma, we think of $\rho_{h,\delta t}$ being extended trivially to all of $\mathbf{R}^d$, so that the spatial translations are all well-defined. Notice that we lose an order of convergence in (46) because of the presence of the boundary $\partial\Omega$.

Proof. We start with the argument for the spatial translations estimated in (46). Notice that since $\rho_{h,\delta t}$ is a probability measure on $\hat{\Omega}$, we may always assume that $|\eta| \le 1$. We first split the integral in x according

$$\begin{aligned} &\int_{\mathbf{R}^d} \Big|\rho_{h,\delta t}(t, x+\eta) - \rho_{h,\delta t}(t,x)\Big| \,\mathrm{d}x \\ &\quad = \int_{\{x\in\hat{\Omega},\, x+\eta\in\hat{\Omega}\}} \Big|\rho_{h,\delta t}(t, x+\eta) - \rho_{h,\delta t}(t,x)\Big| \,\mathrm{d}x + \int_{\{x\in\hat{\Omega},\, x\pm\eta\notin\hat{\Omega}\}} \rho_{h,\delta t}(t,x)\,\mathrm{d}x \end{aligned} \tag{47}$$

Let us start with the treatment of the first integral term. Similarly as in [31], we let $[x,z]$ be the line segment between two points x and z and define the characteristic function of neighboring cells $K \sim L$ by

$$\chi_{KL}(x,z) = \begin{cases} 1 & \text{, if } [x,z] \cap K|L \neq \emptyset, \\ 0 & \text{, else .} \end{cases}$$

By geometric arguments, we observe that

$$\int_{\hat{\Omega}} \chi_{KL}(x, x+\eta)\,\mathrm{d}x \le |K|L|\,|\eta|.$$

With the help of the triangle inequality, we thus estimate

$$\begin{aligned} \int_{\{x\in\hat{\Omega},\, x+\eta\in\hat{\Omega}\}} \Big|\rho_{h,\delta t}(t, x+\eta) - \rho_{h,\delta t}(t,x)\Big| \,\mathrm{d}x &\le \sum_K \sum_{L\sim K} |\rho_K^n - \rho_L^n| \int_{\hat{\Omega}} \chi_{KL}(x, x+\eta)\,\mathrm{d}x \\ &\le |\eta| \sum_K \sum_{L\sim K} |K|L|\,|\rho_K^n - \rho_L^n|. \end{aligned}$$

Squaring both sides, integrating in time and invoking the gradient bound (42) together with the control of the initial datum in (44) yields an error of order $O(|\eta|)$.

Considering the second term in (47), we notice that

$$\int_{\{x\in\hat{\Omega},\, x\pm\eta\notin\hat{\Omega}\}} \rho_{h,\delta t}(t,x)\,\mathrm{d}x \le 2 \int_{\{x\in\hat{\Omega}\colon \mathrm{dist}(x,\partial\hat{\Omega})\le|\eta|\}} \rho_{h,\delta t}\,\mathrm{d}x.$$

To estimate the right-hand side, we let $\varphi(z) = z \log z$ and apply Jensen's inequality,

$$\int_{\mathrm{dist}(\cdot,\partial\hat{\Omega})\le|\eta|} \rho_{h,\delta t}\,\mathrm{d}x \le \Big|\{\mathrm{dist}(\cdot,\partial\hat{\Omega}) \le |\eta|\}\Big|\varphi^{-1}\left(⨍_{\mathrm{dist}(\cdot,\partial\hat{\Omega})\le|\eta|} \varphi(\rho_{h,\delta t})\,\mathrm{d}x\right)$$

Thanks to the bound on the free energy (28) and the initial datum (44), this estimate yields that

$$\int_{\{x\in\hat{\Omega},\, x\pm\eta\notin\hat{\Omega}\}} \rho_{h,\delta t}(t,x)\,\mathrm{d}x \le 2\Big|\{\mathrm{dist}(\cdot,\partial\hat{\Omega}) \le |\eta|\}\Big|\varphi^{-1}\left(\frac{C}{\Big|\{\mathrm{dist}(\cdot,\partial\hat{\Omega}) \le |\eta|\}\Big|}\right),$$

for some $C > 0$, independent of h and η. It is clear that φ^{-1} is growing only sublinearly, and therefore, we have an $o(1)$ bound as $|\eta| \to 0$, uniformly in h.

The estimate on the temporal increments (45) is obtained simply by using (43) together with the triangle inequality. ■

Before proving the convergence of the scheme, we provide an auxiliary lemma on the convergence of discrete gradients, which we introduce in the following. For this, we define the *diamonds*

$$D_{KL} = \{tx_K + (1-t)y : t \in (0,1], y \in K_{|}L\} \subset K,$$

for any $K, L \in \mathcal{T}$ with $L \sim K$. Of course, $K = \dot{\cup}_{L\sim K} D_{KL} \dot{\cup} \{x_K\}$ by construction. For further reference, we note that the volume of these cells can be computed according to the law

$$|D_{KL}| = \frac{d_{KL}}{2d}|K_{|}L|. \tag{48}$$

For a sequence $\zeta_{h,\delta t} = \{\zeta_K^n\}$ on $(0,T)\times\hat{\Omega}$, we may now introduce the discrete gradients

$$\nabla^h \zeta_{h,\delta t}(t,x) = \frac{d}{d_{KL}}(\zeta_L^n - \zeta_K^n)\nu_{KL} \quad \text{for any } (t,x) \in (t^n, t^{n+1}) \times D_{KL}. \tag{49}$$

Here and in the following, ν_{KL} denotes the outer (with respect to K) unit normal vector on the edge $K_{|}L$.

It is a well-known fact that these gradients are convergent.

Lemma 6 (Convergence of discrete gradients [30])**.** *Let $\zeta \in C^\infty([0,T]\times\mathbf{R}^d)$ be given and set $\zeta_{h,\delta t} = \zeta_K^n = \zeta(t^n, x_K)$ in $(t^n, t^{n+1}) \times K$ for any n and any K. Then*

$$\nabla^h \zeta_{h,\delta t} \to \nabla\zeta \quad \textit{weakly-}* \textit{ in } L^\infty((0,T)\times\Omega).$$

In fact, in [30, Lemma 2], an L^2 variant of this result was established and the present version follows from minor modifications of the original proof. For completeness and the convenience of the reader, we recall the details.

Proof. We start by noticing that $\zeta_{h,\delta t} \to \zeta$ uniformly in t and x by the smoothness of ζ. Moreover, by the definition of the discrete gradient, it holds that $\|\nabla^h \zeta_{h,\delta t}\|_{L^\infty} \le d\|\nabla\zeta\|_{C^0}$, and thus, there exists an L^∞ function g and a subsequence (not relabeled) such that

$$\nabla^h \zeta_{h,\delta t} \to g \quad \text{weakly-}* \text{ in } L^\infty((0,T)\times\Omega).$$

It remains thus to prove that $g = \nabla\zeta$ almost everywhere in $(0,T)\times\Omega$.

For this purpose, we let $\varphi \in C_c^\infty([0,T]\times\Omega;\mathbf{R}^d)$ be a given vector field. By the discreteness of the approximation $\zeta_{h,\delta t}$ and the divergence theorem, we have that

$$\begin{aligned}\int_0^T\int_\Omega \zeta_{h,\delta t}\nabla\cdot\varphi\,\mathrm{d}x\,\mathrm{d}t &= \sum_n\sum_K \zeta_K^n \int_{t^n}^{t^{n+1}}\int_K \nabla\cdot\varphi\,\mathrm{d}x\,\mathrm{d}t\\ &= \sum_n\sum_K\sum_{K\sim L}\zeta_K^n\,\nu_{KL}\cdot\int_{t^n}^{t^{n+1}}\int_{K|L}\varphi\,\mathrm{d}S\,\mathrm{d}t,\end{aligned}$$

where $\mathrm{d}S$ denotes the surface measure on edges. By (anti-)symmetrization arguments, we further obtain

$$\int_0^T\int_\Omega \zeta_{h,\delta t}\nabla\cdot\varphi\,\mathrm{d}x\,\mathrm{d}t = \frac{1}{2}\sum_n\sum_K\sum_{K\sim L}(\zeta_K^n-\zeta_L^n)\,\nu_{KL}\cdot\int_{t^n}^{t^{n+1}}\int_{K|L}\varphi\,\mathrm{d}S\,\mathrm{d}t.$$

Denoting by φ_{KL}^n the average of φ on $(t^n,t^{n+1})\times K|L$, and using (48), we may furthermore rewrite this identity as

$$\int_0^T\int_\Omega \zeta_{h,\delta t}\nabla\cdot\varphi\,\mathrm{d}x\,\mathrm{d}t = \sum_n\sum_K\sum_{K\sim L}\frac{d}{d_{KL}}(\zeta_K^n-\zeta_L^n)\,\nu_{KL}\cdot\int_{t^n}^{t^{n+1}}\int_{D_{KL}}\varphi_{KL}^n\,\mathrm{d}x\,\mathrm{d}t\,. \tag{50}$$

On the other hand, by the definition of the discrete gradient, we have that

$$\int_0^T\int_\Omega \nabla^h\zeta_{h,\delta t}\cdot\varphi\,\mathrm{d}x\,\mathrm{d}t = \sum_n\sum_K\sum_{L\sim K}\frac{d}{d_{KL}}(\zeta_L^n-\zeta_K^n)\,\nu_{KL}\cdot\int_{t^n}^{t^{n+1}}\int_{D_{KL}}\varphi\,\mathrm{d}x\,\mathrm{d}t\,. \tag{51}$$

Due to the fact that ζ and φ are smooth, we have the estimates $|\zeta_K^n-\zeta_L^n|\le d_{KL}\|\nabla\zeta\|_{C^0}$ and $|\varphi-\varphi_{KL}^n|\le \delta t\|\partial_t\varphi\|_{C^0}+h\|\nabla\varphi\|_{C^0}\lesssim h\|\varphi\|_{C^1}$, and thus by comparing (50) and (51), we have the discrete partial integration rule

$$\int_0^T\int_\Omega \nabla^h\zeta_{h,\delta t}\cdot\varphi\,\mathrm{d}x\,\mathrm{d}t = -\int_0^T\int_\Omega\zeta_{h,\delta t}\nabla\cdot\varphi\,\mathrm{d}x\,\mathrm{d}t + o(1)\qquad \text{as } h\to 0\,.$$

Passing in this identity to the limit, we then find

$$\int_0^T\int_\Omega g\cdot\nabla\varphi\,\mathrm{d}x\,\mathrm{d}t = -\int_0^T\int_\Omega \zeta\nabla\cdot\varphi\,\mathrm{d}x\,\mathrm{d}t.$$

Because φ was chosen arbitrarily, this proves $g=\nabla\zeta$ as desired. ■

With these preparations at hand, we are now in the position to derive the compactness of the sequence of approximate solutions. Here, the discrete gradient $\nabla^h\rho_{h,\delta t}$ is defined by (49).

Proposition 2 (Compactness)**.** *There exists $\rho\in L^2((0,T);L^1(\Omega))$ with $\nabla\rho\in L^1((0,T)\times\Omega)$, such that $\rho_{h,\delta t}\to\rho$ strongly in $L^2((0,T);L^1(\Omega))$ and $\nabla^h\rho_{h,\delta t}\to\nabla\rho$ weakly in $L^2((0,T);L^1(\Omega))$.*

Proof. The compactness of the sequence in $L^2((0,T);L^1(\Omega))$ is a consequence of Lemma 5 and the Riesz–Fréchet–Kolmogorov theorem. We now turn to the weak convergence of the gradients. The algebraic arguments are essentially identical to the convergence of the gradients in the smooth setting. We will thus skip those and focus on the parts, which are different.

We first show the convergence of the gradients of the roots $\nabla^h\sqrt{\rho_{h,\delta t}}$, for which we need the control on the Fisher information. In view of (48), it is a short computation to obtain the relation

$$\int_{\delta t}^T \int_{\hat\Omega} |\nabla^h \sqrt{\rho_{h,\delta t}}|^2 \,\mathrm{d}x\,\mathrm{d}t = \frac{d}{2}\,\delta t \sum_{n=1}^N \sum_K \sum_{L\sim K} \frac{|K|L|}{d_{KL}} \left(\sqrt{\rho_K^n} - \sqrt{\rho_L^n}\right)^2 .$$

Thanks to (41), the right-hand side is bounded uniformly in h, and thus, for every $\varepsilon > 0$, we find a function $f \in L^2((\varepsilon, T)\times\Omega; \mathbf{R}^d)$ and a subsequence, not relabeled, both dependent on ε, such that

$$\nabla^h \sqrt{\rho_{h,\delta t}} \to f \quad \text{weakly in } L^2((\varepsilon,T)\times\Omega).$$

We have to show that $f = \nabla\sqrt{\rho}$, and thus, the statement becomes independent of ε, and holds true for the full sequence.

We choose a smooth test function $\varphi \in C_c^\infty((0,T)\times\Omega;\mathbf{R}^d)$ and suppose that ε is small in the sense that $\varphi(t,\cdot) = 0$ for any $t \le \varepsilon$. We also assume that $\delta t \le \varepsilon$. Notice then that by algebraic reformulations similar to those in Lemma 6, we have the two identities

$$\int_0^T \int_\Omega \sqrt{\rho_{h,\delta t}}\, \nabla\cdot\varphi \,\mathrm{d}x\,\mathrm{d}t = \sum_n \sum_K \sum_{L\sim K} \frac{d}{d_{KL}} \Big(\sqrt{\rho_K^n} - \sqrt{\rho_L^n}\Big) \nu_{KL} \cdot \int_{t^n}^{t^{n+1}} \int_{D_{KL}} \varphi_{KL}^n \,\mathrm{d}x\,\mathrm{d}t,$$

where φ_{KL}^n is the average of φ over $(t^n, t^{n+1}) \times D_{KL}$, and

$$-\int_0^T \int_\Omega \nabla^h \sqrt{\rho_{h,\delta t}} \cdot \varphi \,\mathrm{d}x\,\mathrm{d}t = \sum_n \sum_K \sum_{L\sim K} \frac{d}{d_{KL}} \Big(\sqrt{\rho_K^n} - \sqrt{\rho_L^n}\Big) \nu_{KL} \cdot \int_{t^n}^{t^{n+1}} \int_{D_{KL}} \varphi \,\mathrm{d}x\,\mathrm{d}t.$$

Now, using $|\varphi - \varphi_{KL}^n| \lesssim h\|\varphi\|_{C^1}$ uniformly in $(t^n, t^{n+1}) \times D_{KL}$, which holds true thanks to the smoothness of φ, and (48), we find that

$$\begin{aligned}
&\left|\sum_n \sum_K \sum_{L\sim K} \frac{d}{d_{KL}} \Big(\sqrt{\rho_K^n} - \sqrt{\rho_L^n}\Big) \nu_{KL} \cdot \int_{t^n}^{t^{n+1}} \int_{D_{KL}} (\varphi - \varphi_{KL}^n) \,\mathrm{d}x\,\mathrm{d}t\right| \\
&\qquad \lesssim h\|\varphi\|_{C^1}\, \delta t \sum_n \sum_K \sum_{L\sim K} |K|L| \left|\sqrt{\rho_L^n} - \sqrt{\rho_K^n}\right| \\
&\qquad \lesssim h\|\varphi\|_{C^1} \sqrt{T} \left(\delta t \sum_n \sum_K \sum_{L\sim K} \frac{|K|L|}{d_{KL}} \Big(\sqrt{\rho_K^n} - \sqrt{\rho_L^n}\Big)^2\right)^{\frac12} \left(\sum_K \sum_{L\sim K} d_{KL}|K|L|\right)^{\frac12}.
\end{aligned}$$

By (41) and the isoperimetric property (22), the right-hand side vanishes as $h \to 0$. Therefore, we have

$$\int_0^T \int_\Omega \sqrt{\rho_{h,\delta t}}\, \nabla\cdot\varphi \,\mathrm{d}x\,\mathrm{d}t = -\int_0^T \int_\Omega \nabla^h \sqrt{\rho_{h,\delta t}} \cdot \varphi \,\mathrm{d}x\,\mathrm{d}t + o(1) \qquad \text{as } h \to 0\,.$$

We notice that $\sqrt{\rho_{h,\delta t}} \to \sqrt{\rho}$ in $L^2((0,T)\times\Omega)$ thanks to the elementary inequality $(\sqrt{a} - \sqrt{b})^2 \le |a-b|$. Therefore, sending h to zero, we thus obtain

$$\int_0^T \int_\Omega \sqrt{\rho}\, \nabla\cdot\varphi \,\mathrm{d}x\,\mathrm{d}t = -\int_0^T \int_\Omega f\cdot\varphi \,\mathrm{d}x\,\mathrm{d}t.$$

This statement implies that $f = \nabla\sqrt{\rho}$ since φ was arbitrary.

Furthermore, by the Cauchy–Schwarz inequality and the fact that ρ is a probability distribution, we observe that

$$\int_0^T \left(\int_\Omega |\nabla\rho|\,\mathrm{d}x\right)^2 \mathrm{d}t \le \int_0^T \int_\Omega \frac{|\nabla\rho|^2}{\rho}\,\mathrm{d}x\,\mathrm{d}t = 4\int_0^T \int_\Omega \left|\nabla\sqrt{\rho}\right|^2 \mathrm{d}x\,\mathrm{d}t,$$

and thus $\nabla\rho \in L^2((0,T);L^1(\Omega))$.

In order to show the convergence of the gradients, notice that thanks to the bound in (42), we have an estimate on the spatial gradient. Indeed, using formula (48), we observe that

$$\int_{\delta t}^T \left(\int_\Omega |\nabla^h \rho_{h,\delta t}|\,\mathrm{d}x\right)^2 \mathrm{d}t = \frac{\delta t}{4}\sum_{n=1}^N \left(\sum_K \sum_{L\sim K} |K_{|}L|\,|\rho_L^n - \rho_K^n|\right)^2.$$

It thus follows that, for every positive ε, there exists a vector-valued measure $g \in L^2((\varepsilon,T);\mathcal{M}(\Omega;\mathbf{R}^d))$ and a subsequence (not relabeled), such that

$$\nabla^h \rho_{h,\delta t} \to g \qquad \text{weakly-}* \text{ in } L^2((\varepsilon,T);\mathcal{M}(\Omega)).$$

If we know that $g = \nabla\rho$, the desired convergence result immediately follows because we have just seen that $\nabla\rho \in L^2((0,T);L^1(\Omega))$. We skip the argument for this missing ingredient, as it closely resembles the previous part of the proof. Notice only that the gradient bound in (42) has to be used instead of the estimate on the Fisher information in (41). ∎

We are now in the position to prove the convergence of the scheme, which is the first statement of Theorem 4. The convergence of the stationary solutions will be addressed afterwards.

Proof of Theorem 4. Convergence of the scheme. The compactness of the sequence is established in Proposition 2. Notice also that the sequence of the initial data ρ_h^0 converges in $L^1(\Omega)$ towards ρ_0 by construction. It remains thus to prove that the limit function ρ solves the continuous equation (1). For this purpose, we choose a subsequence of time steps δt such that $N\,\delta t = T$ for some integer N, consider a test function $\zeta \in C_c^\infty([0,T)\times\mathbf{R}^d)$, and suppose that δt is sufficiently small such that $\zeta(t) = 0$ for $t \ge T - 2\,\delta t$. We set $\zeta_K^n = \zeta(t^n, x_K)$, and write $\zeta_{h,\delta t} = \zeta_K^n$ in $(t^n, t^{n+1})\times K$.

Testing the discrete equation (9) with $\delta t\,\zeta_K^n$ and using the flux identity (16) gives

$$\begin{aligned}
0 &= \sum_{n=0}^{N-1}\sum_K |K|\,\zeta_K^n\left(\rho_K^{n+1} - \rho_K^n\right) + \delta t\sum_{n=0}^{N-1}\sum_K\sum_{L\sim K} |K_{|}L|\,\zeta_K^n\, q_{KL}^{n+1}\,\frac{\rho_K^{n+1}+\rho_L^{n+1}}{2}\\
&\quad + \frac{\delta t}{2}\sum_{n=0}^{N-1}\sum_K\sum_{L\sim K} |K_{|}L|\,\zeta_K^n\, q_{KL}^{n+1}\coth\!\left(\frac{d_{KL}q_{KL}^{n+1}}{2\kappa}\right)\left(\rho_K^{n+1}-\rho_L^{n+1}\right)\\
&= I_1^h + I_2^h + I_3^h.
\end{aligned}$$

Convergence of I_1^h: We first turn to the convergence of the term that involves the time derivative. Performing a discrete integration by parts in the time variable, we have that

$$I_1^h = -\sum_{n=1}^{N-1}\sum_K |K|(\zeta_K^n - \zeta_K^{n-1})\rho_K^n - \sum_K |K|\zeta_K^0\rho_K^0 + \sum_K |K|\zeta_K^{N-1}\rho_K^N.$$

Notice that by our choice of ζ, it holds that $\zeta_K^{N-1} = \zeta(t^{N-1}, x_K) = \zeta((N-1)\,\delta t, x_K) = 0$, and thus, the last term on the right-hand side vanishes. For the middle term, we observe that $\zeta_K^0 = \zeta(0, x_K) = \zeta(0, x) + O(|x - x_K|)$ for any $x \in K$ by the smoothness of the test function, and thus, because ρ_0 is a probability distribution,

$$\sum_K |K| \zeta_K^0 \rho_K^0 = \sum_K \int_K \zeta(0, x)\rho_0(x)\,\mathrm{d}x + O(h) = \int_\Omega \zeta(0, x)\rho_0(x)\,\mathrm{d}x + O(h).$$

Finally, using the smoothness of ζ again, we expand $\zeta_K^n - \zeta_K^{n-1} = \partial_t\zeta(t^{n-1}, x_K)\,\delta t + O(\delta t)^2 = \partial_t\zeta(t, x)\,\delta t + O(\delta t\, h)$ for any $(t, x) \in (t^n, t^{n+1}) \times K$, and thus, similarly as before,

$$\begin{aligned} -\sum_{n=1}^{N-1}\sum_K |K|(\zeta_K^n - \zeta_K^{n-1})\rho_K^n &= -\sum_{n=1}^{N-1}\int_{t^n}^{t^{n-1}}\sum_K \int_K \frac{\zeta_K^n - \zeta_K^{n-1}}{\delta t}\rho_{h,\delta t}(t, x)\,\mathrm{d}x\,\mathrm{d}t \\ &= -\int_{\delta t}^T \int_{\hat\Omega} \partial_t\zeta(t, x)\rho_{h,\delta t}(t, x)\,\mathrm{d}x\,\mathrm{d}t + O(h). \end{aligned}$$

Because $\rho_{h,\delta t}$ is converging stongly in $L^1((0, T) \times \Omega)$ and because

$$\int_{\hat\Omega\setminus\Omega} \rho_{h,\delta t}\,\mathrm{d}x = \int_\Omega (\rho - \rho_{h\,dt})\,\mathrm{d}x, \tag{52}$$

since both ρ and $\rho_{h,\delta t}$ are probability distributions, we find that

$$I_1^h \to -\int_0^T \int_\Omega \partial_t\zeta\,\rho\,\mathrm{d}x\,\mathrm{d}t - \int_\Omega \zeta(0, x)\rho_0(x)\,\mathrm{d}x, \qquad \text{as } h \to 0\,.$$

Convergence of I_2^h: By a discrete integration by parts, we have

$$\begin{aligned} I_2^h &= \frac{\delta t}{2}\sum_n\sum_K\sum_{L\sim K} |K{\scriptstyle|}L|\Big(\zeta_K^n - \zeta_L^n\Big)q_{KL}^{n+1}\frac{\rho_K^{n+1} + \rho_L^{n+1}}{2} \\ &= \frac{\delta t}{2}\sum_n\sum_K\sum_{L\sim K} |K{\scriptstyle|}L|\Big(\zeta_K^n - \zeta_L^n\Big)q_{KL}^{n+1}\rho_K^{n+1} \\ &\qquad + \frac{\delta t}{2}\sum_n\sum_K\sum_{L\sim K} |K{\scriptstyle|}L|\Big(\zeta_K^n - \zeta_L^n\Big)q_{KL}^{n+1}\frac{\rho_L^{n+1} - \rho_K^{n+1}}{2}. \end{aligned}$$

As a consequence of the Lipschitz property of the interaction potential (A3) and the fact that the numerical solution is a probability distribution, it holds that $|q_{KL}^{n+1}| \le \mathrm{Lip}(W)$. Moreover, by the smoothness of ζ, we have that $|\zeta_K^n - \zeta_L^n| \le 2h\|\nabla\zeta\|_{C^0}$, and thus

$$\begin{aligned} &\left|\delta t\sum_n\sum_K\sum_{L\sim K} |K{\scriptstyle|}L|\Big(\zeta_K^n - \zeta_L^n\Big)q_{KL^{n+1}}\Big(\rho_L^{n+1} - \rho_K^{n+1}\Big)\right| \\ &\qquad \lesssim h\|\nabla\zeta\|_{C^0}\,\mathrm{Lip}(W)\sum_n \delta t\sum_K |K{\scriptstyle|}L|\left|\rho_L^{n+1} - \rho_K^{n+1}\right|, \end{aligned}$$

and the right-hand side is of order $O(h)$ by the virtue of the gradient estimate in (42). Therefore,

$$I_2^h = \frac{\delta t}{2}\sum_n\sum_K\sum_{L\sim K} |K{\scriptstyle|}L|\Big(\zeta_K^n - \zeta_L^n\Big)q_{KL}^{n+1}\,\rho_K^{n+1} + o(1) \qquad \text{as } h \to 0.$$

Similarly for the time variable, we estimate slightly more carefully $|\zeta_K^n - \zeta_L^n| \le d_{KL}\|\nabla\zeta\|_{C^0}$ and use identity (48) to arrive at

$$\left|\frac{1}{2}\sum_n\sum_K\sum_{L\sim K}\int_{t^n}^{t^{n+1}}\int_{D_{KL}\cap\Omega}\Big(\rho_{h,\delta t}(t+\delta t,x)-\rho(t,x)\Big)\,\mathrm{d}x\,\mathrm{d}t\,\frac{|K|L|}{|D_{KL}|}(\zeta_K^n-\zeta_L^n)\,q_{KL}^{n+1}\right|$$
$$\le d\|\nabla\zeta\|_{C^0}\,\mathrm{Lip}(W)\int_0^T\int_\Omega\Big|\rho_{h,\delta t}(t+\delta t,x)-\rho(t,x)\Big|\,\mathrm{d}x\,\mathrm{d}t.$$

In a similar manner, by the virtue of (52), it holds that

$$\left|\frac{1}{2}\sum_n\sum_K\sum_{L\sim K}\int_{t^n}^{t^{n+1}}\int_{D_{KL}\setminus\Omega}\rho_{h,\delta t}(t+\delta t,x)\,\mathrm{d}x\,\mathrm{d}t\,\frac{|K|L|}{|D_{KL}|}(\zeta_K^n-\zeta_L^n)q_{KL}^{n+1}\right|$$
$$\le d\|\nabla\zeta\|_{C^0}\,\mathrm{Lip}(W)\int_0^T\int_\Omega\Big|\rho_{h,\delta t}(t+\delta t,x)-\rho(t+\delta t,x)\Big|\,\mathrm{d}x\,\mathrm{d}t.$$

As a consequence of the stong convergence established above and the continuity in time of approximate solutions (45), we see that in both estimates the right-hand side vanishes as h (and then, by (23), also δt) converges to 0. Using (48) again, we thus conclude that

$$I_2^h=\sum_n\sum_K\sum_{L\sim K}\int_{t^n}^{t^{n+1}}\int_{D_{KL}\cap\Omega}\rho(t,x)\,\mathrm{d}x\,\mathrm{d}t\frac{d}{d_{KL}}(\zeta_K^n-\zeta_L^n)q_{KL}^{n+1}+o(1),\qquad \text{as } h\to 0\,.$$

To estimate the flux term, we let $\varepsilon > 0$ be an arbitrarily fixed number and suppose that h is small such that $h < \varepsilon$. We start by noticing that, because W is differentiable away from the origin by (A2), we have the expansion

$$W(x_K-x_J)=W(x_L-x_J)+\nabla W(x_L-x_J)\cdot(x_K-x_L)+o(d_{KL}),$$

for any cell J that satisfies $\mathrm{dist}(J, x_K) \ge 2\varepsilon$ and then also $\mathrm{dist}(J, x_L) \ge \varepsilon$. Thus, using the continuity of ∇W away from the origin, cf. (A2), as $d_{KL} \le 2h$, we have for any $x \in K$ and $y \in J$ that

$$\frac{W(x_K-x_J)-W(x_L-x_J)}{d_{KL}}=\nabla W(x-y)\cdot\frac{x_K-x_L}{|x_K-x_L|}+o(1),\qquad \text{as } h\to 0\,.$$

Therefore, we have to leading order, using in addition the Lipschitz property of W in (A3) and the fact that $\rho_{h,\delta t}$ is a probability distribution,

$$\left|q_{KL}^{n+1}+\sum_J\frac{\rho_J^{n+1}+\rho_J^n}{2}\int_J\nabla W(x-y)\,\mathrm{d}y\cdot\nu_{KL}\right|$$
$$\le 2\,\mathrm{Lip}(W)\sum_{J\text{ s.t. }\mathrm{dist}(J,x_K)\le 2\varepsilon}|J|\frac{\rho_J^{n+1}+\rho_J^n}{2}+o(1)$$
$$\le \mathrm{Lip}(W)\int_{B_{4\varepsilon}(x_K)}\Big(\rho_{h,\delta t}(t+\delta t,y)+\rho_{h,\delta t}(t,y)\Big)\,\mathrm{d}y+o(1),$$

for any $t \in [t^n, t^{n+1}]$ and any fixed ε, as $h \to 0$. It remains to replace the approximate solution in the convolution integral. Doing so, we obtain for any $t \in [t^n, t^{n+1}]$ and $x \in K$,

$$\left|q_{KL}^{n+1}+\int_\Omega\rho(t,y)\nabla W(x-y)\,\mathrm{d}y\cdot\nu_{KL}\right|$$
$$\lesssim \mathrm{Lip}(W)\int_\Omega\Big|\rho_{h,\delta t}(t+\delta t,y)+\rho_{h,\delta t}(t,y)-2\rho(t,y)\Big|\,\mathrm{d}y$$

$$+ \operatorname{Lip}(W) \int_{\hat{\Omega}\setminus\Omega} \Big(\rho_{h,\delta t}(t+\delta t, y) + \rho_{h,\delta t}(t,y)\Big)\,\mathrm{d}y$$
$$+ \operatorname{Lip}(W) \int_{B_{4\varepsilon}(x_K)\cap\Omega} \Big(\rho_{h,\delta t}(t+\delta t, y) + \rho_{h,\delta t}(t,y)\Big)\,\mathrm{d}y + o(1),$$

as $h \to 0$. Thanks to the continuity in time (45), the previously established convergence of the approximating sequence, and identity (52), and using $|\zeta_K^n - \zeta_L^n| \le d_{KL}\|\nabla\zeta\|_{C^0}$ again, we arrive at

$$\left| I_2^h - \sum_n \sum_K \sum_{L\sim K} \int_{t^n}^{t^{n+1}} \int_{D_{KL}\cap\Omega} \rho(t,x)(\nabla W * \rho)(t,x)\,\mathrm{d}x\,\mathrm{d}t \cdot \frac{d}{d_{KL}}(\zeta_L^n - \zeta_K^n)\nu_{KL} \right|$$
$$\lesssim \operatorname{Lip}(W) \int_0^T \int_{B_{4\varepsilon}(x_K)\cap\Omega} \rho\,\mathrm{d}y\,\mathrm{d}t + o(1), \qquad \text{as } h \to 0\,.$$

Alternatively, using the definition of the discrete gradient in (49), this estimate can be written as

$$\left| I_2^h - \int_0^T \int_\Omega \rho(\nabla W * \rho)\cdot\nabla^h\zeta_{h,\delta t}\,\mathrm{d}x\,\mathrm{d}t \right| \lesssim \operatorname{Lip}(W) \int_0^T \int_{B_{4\varepsilon}(x_K)\cap\Omega} \rho\,\mathrm{d}y\,\mathrm{d}t + o(1),$$

as $h \to 0$. Apparently, the (h-indenpendent) first term on the right-hand side vanishes as $\varepsilon \to 0$. Therefore, because $\rho\nabla W * \rho \in L^1((0,T)\times\Omega)$ as a consequence of (A3), we may invoke Lemma 6 and find that

$$I_2^h \to \int_0^T \int_\Omega \rho(\nabla W * \rho)\cdot\nabla\zeta\,\mathrm{d}x\,\mathrm{d}t, \qquad \text{as } h \to 0\,.$$

Convergence of I_3^h: Again, we start with a discrete integration by parts to rewrite I_3^h,

$$I_3^h = \frac{\delta t}{4} \sum_n \sum_K \sum_{L\sim K} |K|L| \Big(\zeta_K^n - \zeta_L^n\Big)\, q_{KL}^{n+1} \coth\left(\frac{d_{KL} q_{KL}^{n+1}}{2\kappa}\right) \Big(\rho_K^{n+1} - \rho_L^{n+1}\Big).$$

In order to get rid of the nonlinearity, we notice that the function $s \mapsto s\coth(s)$ is regular at the origin and $s\coth(s) = 1 + O(s^2)$ as $s \to 0$. Then, by using the bounds $|q_{KL}^{n+1}| \le \operatorname{Lip}(W)$ and $|\zeta_K^n - \zeta_L^n| \le 2h\|\nabla\zeta\|_{C^0}$ as before, we have that the second order contribution is bounded by

$$\frac{\delta t}{\kappa} \sum_n \sum_K \sum_{L\sim K} |K|L|\, |\zeta_K^n - \zeta_L^n|\, d_{KL}\, |q_{KL}^{n+1}|^2\, |\rho_K^{n+1} - \rho_L^{n+1}|$$
$$\lesssim \frac{h^2 \operatorname{Lip}(W)^2\,\delta t}{\kappa} \|\nabla\zeta\|_{C^0} \sum_n \sum_K \sum_{L\sim K} |K|L|\, |\rho_K^{n+1} - \rho_L^{n+1}|$$
$$\le \frac{h^2 \operatorname{Lip}(W)^2}{\kappa^2} \|\nabla\zeta\|_{C^0} \left(\kappa^2\,\delta t\, T \sum_n \Big(\sum_K \sum_{L\sim K} |K|L|\, |\rho_K^{n+1} - \rho_L^{n+1}| \Big)^2 \right)^{\frac{1}{2}}$$

and the right-hand side converges to 0 as $h \to 0$ by the virtue of (42). Therefore, we have to leading order

$$I_3^h = \frac{\kappa\,\delta t}{2} \sum_n \sum_K \sum_{L\sim K} \frac{|K|L|}{d_{KL}} (\zeta_K^n - \zeta_L^n)(\rho_K^{n+1} - \rho_L^{n+1}) + o(1), \qquad \text{as } h \to 0.$$

Using the regularity of ζ again, we find that $\zeta_K^n - \zeta_L^n = \nabla\zeta(t,x)\cdot(x_k - x_L) + O(h d_{KL})$ for any $(t,x) \in (t^{n+1}, t^{n+2}) \times D_{KL}$, where we have used that $d_{KL}, \delta t \lesssim h$. Here, the time interval (t^{n+1}, t^{n+2}) is chosen such that the gradients $\nabla\zeta$ and $\nabla^h \rho_{h,\delta t}$ are evaluated at the same time points. Indeed, we observe from (48), that it holds

$$\begin{aligned}\Bigg| \frac{\kappa\,\delta t}{2} \sum_n \sum_K \sum_{L\sim K} \frac{|K|L|}{d_{KL}} (\zeta_K^n - \zeta_L^n)(\rho_K^{n+1} - \rho_L^{n+1}) \\ + \kappa d \sum_n \sum_K \sum_{L\sim K} \int_{t^{n+1}}^{t^{n+2}} \int_{D_{KL}} \nabla\zeta(t,x)\,\mathrm{d}x\,\mathrm{d}t \cdot \nu_{KL} \frac{\rho_K^{n+1} - \rho_L^{n+1}}{d_{KL}} \Bigg| \\ \lesssim h\kappa\,\delta t \sum_n \sum_K \sum_{L\sim K} |K|L|\,|\rho_K^{n+1} - \rho_L^{n+1}|\,,\end{aligned}$$

and the right-hand side vanishes as $h \to 0$ thanks to the gradient estimate in (42). Therefore, using the definition of the discrete gradient in (49),

$$\begin{aligned} I_3^h &= -\kappa d \sum_n \sum_K \sum_{L\sim K} \int_{t^{n+1}}^{t^{n+2}} \int_{D_{KL}} \nabla\zeta(t,x)\,\mathrm{d}x\,\mathrm{d}t \cdot \nu_{KL} \frac{\rho_K^{n+1} - \rho_L^{n+1}}{d_{KL}} + o(1) \\ &= \kappa \int_{\delta t}^{T} \int_{\hat\Omega} \nabla\zeta(t,x) \cdot \nabla^h \rho_{h,\delta t}(t,x)\,\mathrm{d}x\,\mathrm{d}t + o(1), \qquad \text{as } h \to 0. \end{aligned}$$

As the final step, we implement the gradient convergence established in Proposition 2, and conclude that

$$I_3^h \to \kappa \int_0^T \int_\Omega \nabla\zeta \cdot \nabla\rho\,\mathrm{d}x\,\mathrm{d}t, \qquad \text{as } h \to 0.$$

Summary. Putting together the convergence results for I_1^h, I_2^h and I_3^h, passing to the limit in the identity $I_1^h + I_2^h + I_3^h = 0$ yields

$$\begin{aligned} -\int_0^T \int_\Omega \partial_t\zeta\rho\,\mathrm{d}x\,\mathrm{d}t + \int_0^T \int_\Omega \rho(\nabla W * \rho) \cdot \nabla\zeta\,\mathrm{d}x\,\mathrm{d}t + \kappa \int_0^T \int_\Omega \nabla\zeta \cdot \nabla\rho\,\mathrm{d}x\,\mathrm{d}t \\ = \int_\Omega \zeta(0,x)\rho^0(x)\,\mathrm{d}x, \end{aligned}$$

for any test function ζ. This is the distributional formulation of the aggregation-diffusion equation. ■

Similar to the convergence of the scheme, the convergence of stationary solutions is based on the Riesz–Fréchet–Kolmogorov compactness theorem. As a preparation, we derive estimates on discrete gradients.

Lemma 7. *Let $\{\rho_K\}_K$ be a stationary solution of the Scharfetter–Gummel scheme. Then it holds*

$$\sum_K \sum_{L\sim K} |K|L|\,|\rho_K - \rho_L| \lesssim \frac{\mathrm{Lip}(W)}{\kappa}.$$

Proof. We recall that by Theorem 2, the stationary solutions obeys the equation

$$\rho_K = \frac{1}{Z(\rho)} \exp\!\Bigg(-\frac{1}{\kappa} \sum_J |J|\,W(x_K - x_J)\rho_J \Bigg).$$

Making use of the elementary estimate $|\exp(a) - \exp(b)| \le \Big(\exp(a) + \exp(b)\Big)\,|a-b|$, we thus have that

$$|\rho_K - \rho_L| \le \frac{1}{\kappa}\,(\rho_K + \rho_L) \sum_J |J|\,\Big|W(x_K - x_J) - W(x_L - x_J)\Big|\rho_J.$$

Because the aggregation potential is Lipschitz (A3), and $\{\rho_K\}_K$ a probability distribution, we find

$$|\rho_K - \rho_L| \le \frac{d_{KL}\,\mathrm{Lip}(W)}{\kappa}(\rho_K + \rho_L).$$

It remains to apply a relabeling argument, the isoperimetric property (22) of the scheme, and, again, the fact that $\{\rho_K\}_K$ is a probability distribution to deduce that

$$\sum_K \sum_{L\sim K} |K{\scriptstyle|}L|\,|\rho_K - \rho_L| \lesssim \frac{\mathrm{Lip}(W)}{\kappa}. \qquad \blacksquare$$

We now define the finite volume approximation ρ_h of the stationary state $\{\rho_K\}_K$ by setting

$$\rho_h = \rho_K \quad \text{in } K.$$

We provide a continuity result. For this, we extend ρ_h trivially to all of $\mathbf{R}^d$.

Lemma 8. *For $\eta \in \mathbf{R}^d$, it holds that*

$$\int_{\mathbf{R}^d} |\rho_h(x+\eta) - \rho_h(x)|\,\mathrm{d}x = o(1), \tag{53}$$

as $|\eta| \to 0$, uniformly in h.

Proof. We first notice that stationary states have finite entropy, more precisely,

$$\sum_K |K|\rho_K \log \rho_K \lesssim 1, \tag{54}$$

uniformly in h. Indeed, in view of the properties (A1) and (A3) of the potential and the fact that $\{\rho_K\}_K$ is a probability distribution, it holds that

$$\sum_J |J|\,|W(x_K - x_J)|\rho_J \le \mathrm{Lip}(W)\,\mathrm{diam}(\hat\Omega) \sum_J |J|\rho_J = \mathrm{Lip}(W)\,\mathrm{diam}(\hat\Omega).$$

Therefore, invoking the characterization of stationary states in Theorem 2, we deduce that

$$\kappa|\log \rho_K| \le 2\,\mathrm{Lip}(W)\,\mathrm{diam}(\hat\Omega) + \log|\hat\Omega|. \tag{55}$$

Using again that $\{\rho_K\}_K$ is a probability distribution, (54) follows immediately.

The proof of (53) now follows almost identical to the one in the time-dependent setting, (46). We thus omit the details. ■

We finally provide the argument for the second statement of Theorem 4.

Proof of Theorem 4. Convergence of stationary states. We first remark that as a consequence of Lemma 8 and the Riesz–Fréchet–Kolmogorov theorem, the sequence $\{\rho_h\}_h$ is compact in $L^1(\Omega)$, and thus, there exists $\rho \in L^1(\Omega)$ such that $\rho_h \to \rho$ in $L^1(\Omega)$, as $h \to 0$. Our goal is to show that ρ is a stationary solution of the aggregation-diffusion equation (1).

As in the previous two lemmas, our starting point is the characterization of stationary states in Theorem 2,

$$\kappa \log \rho_K = -\sum_L |L|\, W(x_K - x_L)\, \rho_L + \log Z^h(\rho_h). \tag{56}$$

We have seen in the proof of Lemma 8 that the right-hand side of this identity is bounded uniformly in h, see (55), therefore, by the just established L^1 convergence and the dominated convergence theorem, it follows that

$$\|\kappa \log \rho_h - \kappa \log \rho\|_{L^1(\Omega)} \to 0 \qquad \text{as } h \to 0.$$

Regarding the convergence of the right-hand side of (56), we notice that, for any $x \in K$, it holds

$$\begin{aligned}\left|\sum_L |L| W(x_K - x_L)\rho_L - (W * \rho)(x)\right| &\le \sum_L \int_{L\cap\Omega} |W(x_K - x_L)|\, |\rho_L - \rho(y)|\, \mathrm{d}y \\ &\quad + \sum_L \int_{L\cap\Omega} |W(x_K - x_L) - W(x-y)|\rho(y)\, \mathrm{d}y \\ &\quad + \sum_L \int_{L\setminus\Omega} |W(x_K - x_L)|\rho_L\, \mathrm{d}y,\end{aligned}$$

and thus, using the properties (A1) and (A3) of the aggregation potential and identity (52), which also holds true in the stationary case, we obtain that

$$\begin{aligned}&\left|\sum_L |L|\, W(x_K - x_L)\, \rho_L - (W * \rho)(x)\right| \\ &\qquad \le 2\,\mathrm{Lip}(W)\,\mathrm{diam}(\Omega)\|\rho_h - \rho\|_{L^1(\Omega)} + h\,\mathrm{Lip}(W)\|\rho\|_{L^1(\Omega)} = o(1),\end{aligned}$$

as $h \to 0$. We easily deduce that

$$\sum_K \int_K \left|\sum_L |L| W(x_K - x_L)\rho_L - (W * \rho)(x)\right| \mathrm{d}x \to 0 \qquad \text{as } h \to 0,$$

or, in words, the first term on the right-hand side of (56) converges to $-W * \rho$ in $L^1(\Omega)$.

We consider now a mean-zero test function $\varphi \in C_c^\infty(\Omega)$ and define its (mean-zero) finite-volume approximation, as usual, by $\varphi_h = \varphi_K = ⨍_K \varphi\, \mathrm{d}x$ on K. Then $\varphi_h \to \varphi$ uniformly on Ω. Testing (56) by φ_h then yields

$$\sum_K |K| \left(\kappa \log \rho_K + \sum_L |L| W(x_K - x_L)\rho_L\right) \varphi_K = 0,$$

and passing to the limit $h \to 0$ gives

$$\int_\Omega (\kappa \log \rho + W * \rho)\, \varphi\, \mathrm{d}x = 0,$$

which is, since φ was arbitrary and mean-zero, equivalent to

$$\kappa \log \rho + W * \rho + \kappa \log Z(\rho) = 0,$$

where $Z(\rho) = ⨍_\Omega \exp(-\kappa^{-1} W * \rho)\, \mathrm{d}x$. This is a characterization of stationary solutions for the continuous problem as can be verified by inspection of the free energy dissipation (4). ∎

Acknowledgement

This work is funded by the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) under Germany's Excellence Strategy EXC 2044 – 390685587, *Mathematics Münster: Dynamics–Geometry–Structure,* and EXC 2047 – 390685813, *Hausdorff Center for Mathematics*, as well as the Collaborative Research Center 1060 – 211504053, *The Mathematics of Emergent Effects* at the Universität Bonn.

References

[1] R. Bailo, J. A. Carrillo, and J. Hu. Fully Discrete Positivity-Preserving and Energy-Dissipating Schemes for Aggregation-Diffusion Equations with a Gradient Flow Structure. *arXiv:1811.11502.*

[2] 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. *arXiv:2002.10821.*

[3] D. Benedetto, E. Caglioti, J. A. Carrillo, and M. Pulvirenti. A non-Maxwellian steady distribution for one-dimensional granular media. *J. Statist. Phys.*, 91(5/6):979–990, 1998.

[4] M. Bessemoulin-Chatard. A finite volume scheme for convection-diffusion equations with nonlinear diffusion derived from the Scharfetter-Gummel scheme. *Numer. Math.*, 121(4):637–670, 2012.

[5] M. Bessemoulin-Chatard, C. Chainais-Hillairet, and F. Filbet. On discrete functional inequalities for some finite volume schemes. *IMA J. Numer. Anal.*, 35(3):1125–1149, jul 2014.

[6] J. Binney and S. Tremaine. *Galactic dynamics.* Princeton University Press, 2008.

[7] J. A. Bittencourt. *Fundamentals of plasma physics.* Pergamon Press, Oxford, 1986.

[8] F. Bolley, I. Gentil, and A. Guillin. Uniform convergence to equilibrium for granular media. *Arch. Ration. Mech. Anal.*, 208(2):429–445, 2013.

[9] F. Brezzi, L. D. Marini, and P. Pietra. Two-dimensional exponential fitting and applications to drift-diffusion models. *SIAM J. Numer. Anal.*, 26(6):1342–1355, 1989.

[10] M. Burger, V. Capasso, and D. Morale. On an aggregation model with long and short range interactions. *Nonlinear Anal. Real World Appl.*, 8(3):939–958, jul 2007.

[11] C. Cancès, C. Chainais-Hillairet, J. Fuhrmann, and B. Gaudeul. A numerical analysis focused comparison of several finite volume schemes for an unipolar degenerated drift-diffusion model. *arXiv:1907.11126.*

[12] C. Cancès, C. Chainais-Hillairet, M. Herda, and S. Krell. Large time behavior of nonlinear finite volume schemes for convection-diffusion equations. *arXiv:1911.04962.*

[13] C. Cancès, T. O. Gallouët, and G. Todeschi. A variational finite volume scheme for Wasserstein gradient flows. *arXiv:1907.08305.*

[14] J. A. Carrillo, K. Craig, and F. S. Patacchini. A blob method for diffusion. *Calc. Var. Partial Differential Equations*, 58(2), feb 2019.

[15] J. A. Carrillo, K. Craig, and Y. Yao. Aggregation-Diffusion Equations: Dynamics, Asymptotics, and Singular Limits. In *Active Particles, Volume 2*, pages 65–108. Birkhäuser, Cham, 2019.

[16] J. A. Carrillo, M. Di Francesco, A. Figalli, T. Laurent, and D. Slepčev. Global-in-time weak measure solutions and finite-time aggregation for nonlocal interaction equations. *Duke Math. J.*, 156(2):229–271, 2011.

[17] J. A. Carrillo, R. S. Gvalani, G. A. Pavliotis, and A. Schlichting. Long-time behaviour and phase transitions for the mckean–vlasov equation on the torus. *Arch. Ration. Mech. Anal.*, 235(1):635–690, jul 2019.

[18] J. A. Carrillo, R. J. McCann, and C. Villani. Kinetic equilibration rates for granular media and related equations: entropy dissipation and mass transportation estimates. *Rev. Mat. Iberoamericana*, 19(3):971–1018, 12 2003.

[19] C. Chainais-Hillairet and F. Filbet. Asymptotic behaviour of a finite-volume scheme for the transient drift-diffusion model. *IMA J. Numer. Anal.*, 27(4):689–716, 2007.

[20] C. Chainais-Hillairet, J.-G. Liu, and Y.-J. Peng. Finite volume scheme for multi-dimensional drift-diffusion equations and convergence analysis. *M2AN Math. Model. Numer. Anal.*, 37(2):319–338, 2003.

[21] C. Chainais-Hillairet and Y.-J. Peng. Convergence of a finite-volume scheme for the drift-diffusion equations in 1D. *IMA J. Numer. Anal.*, 23(1):81–108, 2003.
[22] C. Chainais-Hillairet and Y.-J. Peng. Finite volume approximation for degenerate drift-diffusion system in several space dimensions. *Math. Models Methods Appl. Sci.*, 14(3):461–481, 2004.
[23] S.-N. Chow, W. Huang, Y. Li, and H. Zhou. Fokker–Planck equations for a free energy functional or Markov process on a graph. *Arch. Ration. Mech. Anal.*, 203(3):969–1008, 2012.
[24] J. G. Conlon and A. Schlichting. A non-local problem for the Fokker-Planck equation related to the Becker-Döring model. *Discret. Contin. Dyn. Syst.*, 39(4):1–57, 2019.
[25] F. Delarue, F. Lagoutière, and N. Vauchelet. Convergence analysis of upwind type schemes for the aggregation equation with pointy potential. *Annales Henri Lebesgue*, 3:217–260, 2020.
[26] K. Disser and M. Liero. On gradient structures for Markov chains and the passage to Wasserstein gradient flows. *Netw. Heterog. Media*, 10(2):233–253, 2015.
[27] M. Erbar, M. Fathi, V. Laschos, and A. Schlichting. Gradient flow structure for McKean–Vlasov equations on discrete spaces. *Discrete Contin. Dyn. Syst.*, 36(12):6799–6833, 2016.
[28] A. Esposito, F. S. Patacchini, A. Schlichting, and D. Slepčev. Nonlocal-interaction equation on graphs: gradient flow structure and continuum limit. *arXiv:1912.09834*.
[29] R. Eymard, J. Fuhrmann, and K. Gärtner. A finite volume scheme for nonlinear parabolic equations derived from one-dimensional local Dirichlet problems. *Numer. Math.*, 102(3):463–495, 2006.
[30] R. Eymard and T. Gallouët. *H*-convergence and numerical schemes for elliptic problems. *SIAM J. Numer. Anal.*, 41(2):539–562, 2003.
[31] R. Eymard, T. Gallouët, and R. Herbin. Finite volume methods. In *Handbook of numerical analysis, Vol. VII*, Handb. Numer. Anal., VII, pages 713–1020. North-Holland, Amsterdam, 2000.
[32] P. Farrell, N. Rotundo, D. H. Doan, M. Kantner, J. Fuhrmann, and T. Koprucki. *Handbook of Optoelectronic Device Modeling and Simulation*, volume 2, chapter Chapter 50 “Drift-Diffusion Models”, pages 733–771. CRC Press Taylor & Francis Group, 2017.
[33] R. Hegselmann and U. Krause. Opinion dynamics and bounded confidence: models, analysis and simulation. *J. Artif. Soc. Soc. Simul.*, 5(3), 2002.
[34] M. Heida, M. Kantner, and A. Stephan. Consistency and convergence for a family of finite volume discretizations of the Fokker–Planck operator. *arXiv:2002.09385*.
[35] R. Jordan, D. Kinderlehrer, and F. Otto. The Variational Formulation of the Fokker–Planck Equation. *SIAM J. Math. Anal.*, 29(1):1–17, jan 1998.
[36] A. Jüngel. Numerical approximation of a drift-diffusion model for semiconductors with nonlinear diffusion. *Z. Angew. Math. Mech.*, 75(10):783–799, 1995.
[37] A. Jüngel and P. Pietra. A Discretization Scheme for a Quasi-Hydrodynamic Semiconductor Model. *Math. Models Methods Appl. Sci.*, 07(07):935–955, nov 1997.
[38] E. F. Keller and L. A. Segel. Model for chemotaxis. *J. Theor. Biol.*, 30(2):225–234, 1971.
[39] J. G. Kirkwood and E. Monroe. Statistical Mechanics of Fusion. *J. Chem. Phys.*, 9(7):514–526, jul 1941.
[40] Y. Kuramoto. Rhythms and turbulence in populations of chemical oscillators. *Phys. A*, 106(1-2):128–143, 1981. Statphys 14 (Proc. Fourteenth Internat. Conf. Thermodynamics and Statist. Mech., Univ. Alberta, Edmonton, Alta., 1980).
[41] F. Lagoutière and N. Vauchelet. Analysis and Simulation of Nonlinear and Nonlocal Transport Equations. In *Innovative Algorithms and Analysis*, pages 265–288. Springer, Cham, 2017.
[42] J. L. Lebowitz and O. Penrose. Rigorous Treatment of the Van Der Waals-Maxwell Theory of the Liquid-Vapor Transition. *J. Math. Phys.*, 7(1):98–113, jan 1966.
[43] L. Li and J.-G. Liu. Large time behaviors of upwind schemes and *B*-schemes for Fokker-Planck equations on $\mathbb{R}$ by jump processes. *Math. Comput.*, Feb. 2020.
[44] H. Liu and W. Maimaitiyiming. Positive and free energy satisfying schemes for diffusion with interaction potentials. *arXiv:1910.00151*.
[45] N. Martzel and C. Aslangul. Mean-field treatment of the many-body Fokker-Planck equation. *J. Phys. A*, 34(50):11225–11240, 2001.
[46] H. P. McKean, Jr. A class of Markov processes associated with nonlinear parabolic equations. *Proc. Nat. Acad. Sci. U.S.A.*, 56:1907–1911, 1966.

[47] H. P. McKean, Jr. Propagation of chaos for a class of non-linear parabolic equations. In *Stochastic Differential Equations (Lecture Series in Differential Equations, Session 7, Catholic Univ., 1967)*, pages 41–57. Air Force Office Sci. Res., Arlington, Va., 1967.

[48] K. Oelschläger. A martingale approach to the law of large numbers for weakly interacting stochastic processes. *Ann. Probab.*, 12(2):458–479, 1984.

[49] D. L. Scharfetter and H. K. Gummel. Large-signal analysis of a silicon Read diode oscillator. *IEEE Trans. Electron Devices*, 16(1):64–77, jan jan 1969.

[50] A. Schlichting and C. Seis. Analysis of the implicit upwind finite volume scheme with rough coefficients. *Numer. Math.*, 139(1):155–186, 2018.

[51] A. Stevens. The derivation of chemotaxis equations as limit dynamics of moderately interacting stochastic many-particle systems. *SIAM J. Appl. Math.*, 61(1):183–212, jan 2000.

[52] A.-S. Sznitman. Topics in propagation of chaos. In *École d'Été de Probabilités de Saint-Flour XIX—1989*, volume 1464 of *Lecture Notes in Math.*, pages 165–251. Springer, Berlin, 1991.

[53] Y. Tamura. Free energy and the convergence of distributions of diffusion processes of McKean type. *J. Fac. Sci., Univ. Tokyo, Sect. I A*, 34:443–484, 1987.

[54] G. Teschl. *Ordinary Differential Equations and Dynamical Systems*. Graduate Studies in Mathematics. American Mathematical Soc., 2012.

[55] J. Tugaut. Convergence to the equilibria for self-stabilizing processes in double-well landscape. *Ann. Probab.*, 41(3A):1427–1460, may 2013.

[56] C. Villani. *Topics in optimal transportation*, volume 58 of *Graduate Studies in Mathematics*. American Mathematical Society, Providence, RI, 2003.

[57] J. A. Walker. *Dynamical Systems and Evolution Equations*. Springer US, 1980.

[58] H. T. Yau. Relative entropy and hydrodynamics of Ginzburg-Landau models. *Lett. Math. Phys.*, 22(1):63–80, 1991.