# ENERGY-TRANSPORT SYSTEMS FOR OPTICAL LATTICES: DERIVATION, ANALYSIS, SIMULATION

MARCEL BRAUKHOFF AND ANSGAR JÜNGEL

Abstract. Energy-transport equations for the transport of fermions in optical lattices are formally derived from a Boltzmann transport equation with a periodic lattice potential in the diffusive limit. The limit model possesses a formal gradient-flow structure like in the case of the energy-transport equations for semiconductors. At the zeroth-order high temperature limit, the energy-transport equations reduce to the whole-space logarithmic diffusion equation which has some unphysical properties. Therefore, the first-order expansion is derived and analyzed. The existence of weak solutions to the time-discretized system for the particle and energy densities with periodic boundary conditions is proved. The difficulties are the nonstandard degeneracy and the quadratic gradient term. The main tool of the proof is a result on the strong convergence of the gradients of the approximate solutions. Numerical simulations in one space dimension show that the particle density converges to a constant steady state if the initial energy density is sufficiently large, otherwise the particle density converges to a nonconstant steady state.

## 1. Introduction

An optical lattice is a spatially periodic structure that is formed by interfering optical laser beams. The interference produces an optical standing wave that may trap neutral atoms [4]. The lattice potential mimics the crystal lattice in a solid, while the trapped atoms mimic the valance electrons in a solid state crystal. In contrast to solid materials, it is easily possible to adjust the geometry and depth of the potential of an optical lattice. Another advantage is that the dynamics of the atoms, if cooled down to a few nanokelvin, can be followed on the time scale of milliseconds. Therefore, optical lattices are ideal systems to study physical phenomena that are difficult to observe in solid crystals. Moreover, they are promising candidates to realize quantum information processors [14] and extremely precise atomic clocks [2].

The dynamics of ultracold fermionic clouds in an optical lattice can be modeled by the Fermi-Hubbard model with a Hamiltonian that is a result of the lattice potential created by interfering laser beams and short-ranged collisions [11]. In the semi-classical picture, the interplay between diffusive and ballistic regimes can be described by a Boltzmann

*Date*: April 6, 2017.

2000 *Mathematics Subject Classification.* 35K59, 35K65, 35Q20, 82B40.

*Key words and phrases.* Energy-transport models, optical lattice, degenerate equations, quadratic gradient, existence of weak solutions, finite differences.

The authors acknowledge partial support from the Austrian Science Fund (FWF), grants P27352, P30000, and W1245.

transport equation [13], which is able to model qualitatively the observed cloud shapes [23].

In this paper, we investigate moment equations which are formally derived from a Boltzmann equation in the diffusive regime. The motivation of our work is the observation that in the (relative) high-temperature limit, the lowest-order diffusion approximation of the Boltzmann equation leads to a logarithmic diffusion equation [20] which has nonphysical properties in the whole space (for instance, it loses mass). Our aim is to derive the next-order approximation, leading to energy-transport equations for the particle and energy densities, and to analyze and simulate the resulting system of degenerate parabolic equations under periodic boundary conditions [20].

The starting point is the (scaled) Boltzmann equation for the distribution function $f(x,p,t)$,

$$\partial_t f + u\cdot\nabla_x f + \nabla_x V\cdot\nabla_p f = Q(f), \tag{1}$$

where $x\in\mathbb{R}^d$ is the spatial variable, $p$ is the crystal momentum defined on the $d$-dimensional torus $\mathbb{T}^d$ with unit measure, and $t>0$ is the time. The velocity $u$ is defined by $u(p)=\nabla_p\varepsilon(p)$ with the energy $\varepsilon(p)$, $V(x,t)$ is the potential, and $Q(f)$ is the collision operator. Compared to the semiconductor Boltzmann equation, there are two major differences.

First, the band energy $\varepsilon(p)$ is given by the periodic dispersion relation

$$\varepsilon(p) = -2\varepsilon_0\sum_{i=1}^d \cos(2\pi p_i), \quad p\in\mathbb{T}^d. \tag{2}$$

The constant $\varepsilon_0$ is a measure for the tunneling rate of a particle from one lattice site to a neighboring one. In semiconductor physics, usually a parabolic band structure is assumed, $\varepsilon(p)=\frac12|p|^2$ [15]. This formula also appears in kinetic gas theory as the (microscopic) kinetic energy. The band energy in optical lattice is bounded, while it is unbounded when $\varepsilon(p)=\frac12|p|^2$. This has important consequences regarding the integrability of the equilibrium distribution (see below).

Second, the potential $V$ is given by $V=-U_0n$, where $n=\int_{\mathbb{T}^d}fdp$ is the particle density and $U_0>0$ models the strength of the on-site interaction between spin-up and spin-down components [23]. In semiconductor physics, $V$ is the electric potential which is a given function or determined self-consistently from the (scaled) Poisson equation [15]. The definition $V=-U_0n$ leads to unexpected "degeneracies" in the moment equations, see the discussion below.

The collision operator is given as in [23] by the relaxation-time approximation

$$Q(f) = \frac{1}{\tau}(\mathcal{F}_f - f),$$

where $\tau>0$ is the relaxation time and $\mathcal{F}_f$ is determined by minimizing the free energy for fermions associated to (1) under the constraints of mass and energy conservation (see Section 2 for details), leading to

$$\mathcal{F}_f(x,p,t) = \big(\eta+\exp(-\lambda_0(x,t)-\lambda_1(x,t)\varepsilon(p))\big)^{-1}, \quad x\in\mathbb{R}^d,\ p\in\mathbb{T}^d,\ t>0,$$

where $(\lambda_0, \lambda_1)$ are the Lagrange multipliers resulting from the mass and energy constraints. For $\eta = 1$, we obtain the Fermi-Dirac distribution, while for $\eta = 0$, $\mathcal{F}_f$ equals the Maxwell-Boltzmann distribution. We may consider $\mathcal{F}_f$ as a function of $(\lambda_0, \lambda_1)$ and write $\mathcal{F}(\lambda_0, \lambda_1; p) = [\eta + \exp((-\lambda_0 - \lambda_1\varepsilon(p))]^{-1}$.

The variable $\lambda_1$ can be interpreted as the negative inverse (absolute) temperature, while $\lambda_0$ is related to the so-called chemical potential [15]. Since the band energy is bounded, the equilibrium $\mathcal{F}_f$ is integrable even when $\lambda_1 > 0$, which means that the absolute temperature may be negative. In fact, negative absolute temperatures can be realized in experiments with cold atoms [22]. Negative temperatures occur in equilibrated (quantum) systems that are characterized by an inverted population of energy states. The thermodynamical implications of negative temperatures are discussed in [21].

In the following, we detail the main results of the paper.

**Formal derivation and entropy structure.** Starting from the Boltzmann equation (1), we derive formally moment equations in the limit of large times and dominant collisions. More precisely, the particle density $n = \int_{\mathbb{T}^d} \mathcal{F} dp$ and energy density $E = \int_{\mathbb{T}^d} \mathcal{F}\varepsilon dp$ solve the energy-transport equations

$$\partial_t n + \operatorname{div} J_n = 0, \quad \partial_t E + \operatorname{div} J_E - J_n \cdot \nabla V = 0 \tag{3}$$

for $x \in \mathbb{R}^d$, $t > 0$, where the particle and energy current densities are given by

$$J_n = -\sum_{i=0}^{1} D_{0i}\nabla\lambda_i - \lambda_1 D_{00}\nabla V, \quad J_E = -\sum_{i=0}^{1} D_{1i}\nabla\lambda_i - \lambda_1 D_{10}\nabla V, \tag{4}$$

and the diffusion coefficients depend nonlocally on $\mathcal{F}$ and hence on $(\lambda_0, \lambda_1)$; see Proposition 1. The structure of system (3)-(4) is similar to the semiconductor case [17] but the diffusion coefficients $D_j$ are different. For $V = -U_0 n$, the Joule heating term $J_n \cdot \nabla V$ contains the squared gradient $|\nabla n|^2$, while in the semiconductor case it contains $|\nabla V|^2$ which is generally smoother than $|\nabla n|^2$.

System (3)-(4) possesses a formal gradient-flow or entropy structure. Indeed, the entropy $H$, defined in Section 2.2, is nonincreasing in time,

$$\frac{dH}{dt} = -\int_{\mathbb{R}^d} \sum_{i,j=0}^{1} \nabla\mu_i^\top L_{ij} \nabla\mu_j dx \le 0;$$

see Proposition 3. Here, the functions $\mu_0 = \lambda_0 + \lambda_1 V$ and $\mu_1 = \lambda_1$ are called the dual entropy variables, and the coefficients $L_{ij}$ are defined in (20). In the dual entropy variables, the potential terms are eliminated, leading to the "symmetric" problem

$$\partial_t \begin{pmatrix} n \\ E \end{pmatrix} = \operatorname{div}\left( \begin{pmatrix} L_{00} & L_{01} \\ L_{10} & L_{11} \end{pmatrix} \nabla \begin{pmatrix} \mu_0 \\ \mu_1 \end{pmatrix} \right),$$

where the matrix $(L_{ij})$ is symmetric and positive definite. This formal gradient-flow structure allows for the development of an existence theory but only for uniformly positive definite diffusion matrices [10]. A general existence result (including electric potentials) is still missing.

A further major difficulty comes from the fact that the system possesses certain "degeneracies" in the mapping $(n,E)\mapsto\mu=(\mu_0,\mu_1)$ and the entropy production $-dH/dt$. For instance, the determinant of the Jacobi matrix $\partial(n,E)/\partial\mu$ may vanish at certain points. Such a situation also occurs for the semiconductor energy-transport equations but only at the *boundary* of the domain of definition (namely at $E=0$). In the present situation, the degeneracy may occur at points in the *interior* of the domain of definition. In view of these difficulties, an analysis of the general energy-transport model (3)-(4) is currently out of reach. This motivates our approach to introduce a simplified model.

**Analysis of high-temperature energy-transport models.** We show the existence of weak solutions to a simplified energy-transport model. It is argued in [20] that the temperature is large (relative to the nanokelvin scale) in the center of the atomic cloud for long times. Therefore, we simplify (3)-(4) by performing the high-temperature limit. For high temperatures, the relaxation time may be approximated by $\tau(n)=\tau_0/(n(1-n))$ [23, Suppl.]. As $\theta=-1/\lambda_1$ can be interpreted as the temperature, the high-temperature limit corresponds to the limit $\lambda_1\to 0$. Expanding $\mathcal{F}(\lambda_0,\lambda_1)$ around $(\lambda_0,0)$ up to zeroth order leads to the diffusion equation (see Section 3)

$$\partial_t n = \operatorname{div}\left(\frac{\tau_0\nabla n}{n(1-\eta n)}\right)\quad\text{in }\mathbb{R}^d. \tag{5}$$

In the case $\eta=0$, we obtain the logarithmic diffusion equation $\partial_t n=\tau_0\Delta\log n$ which predicts a nonphysical behavior. Indeed, in two space dimensions, it can be shown that the particle number is not conserved and the unique smooth solution exists for finite time only; see, e.g., [9, 24]. We expect a similar behavior when $\eta>0$. This motivates us to compute the next-order expansion. It turns out that at first order and with $V=-U_0n$, $(n,E)$ is solving the (rescaled) energy-transport equations

$$\partial_t n = \operatorname{div}\left(\frac{W\nabla n}{n(1-\eta n)}\right), \tag{6}$$

$$\partial_t W = \frac{2d-1}{2d}\operatorname{div}\left(\frac{\nabla W}{n(1-\eta n)}\right) - U\frac{W|\nabla n|^2}{n(1-\eta n)}. \tag{7}$$

where $U=U_0/(2d\varepsilon_0^2)$ and $W=1-UE$ is the "reverted" energy. The case $W=0$ corresponds to the maximal energy $E=1/U_0$. Taking into account the periodic lattice structure, we solve (6)-(7) on the torus $\mathbb{T}^d$, together with the initial conditions $n(0)=n^0$, $W(0)=W^0$ in $\mathbb{T}^d$.

The structure of the diffusion equation (6) is similar to (5), but the diffusion coefficient contains $W$ as a factor, adding a degeneracy to the singular logarithmic diffusion equation. It is an open problem whether this factor removes the unphysical behavior of the solution to (5) in $\mathbb{R}^d$. We avoid this problem by solving (6)-(7) in a bounded domain and by looking for strictly positive particle densities. Is is another open problem to prove the existence of solutions to (6)-(7) in the whole space.

Because of the squared gradient term in (7), the energy $E$ (or $W$) is not conserved but the total energy $W_{\rm tot}=W-(U/2)n^2$. In fact, in terms of $W_{\rm tot}$, the squared gradient term is eliminated,

$$\partial_t W_{\rm tot} = \operatorname{div}\left(\frac{\nabla W}{2n(1-\eta n)} + \frac{UW}{1-\eta n}\nabla n\right). \tag{8}$$

Unfortunately, this formulation does not help for the analysis since the treatment of $\partial_t(n^2) = 2n\partial_t n$ is delicate as $\partial_t n$ lies in the dual space $H^1(\mathbb{T}^d)'$ but $n$ is generally not an element of $H^1(\mathbb{T}^d)$ because of the degeneracy (we have only $W^{1/2}\nabla n \in L^2(\mathbb{T}^d)$).

The analysis of system (6)-(7) is very challenging since the first equation is degenerate in $W$, and the second equation contains a quadratic gradient term. In the literature, there exist existence results for degenerate equations with quadratic gradient terms [8, 12], but the degeneracy is of porous-medium type. A more complex degeneracy was investigated in [7]. In our case, the degeneracy comes from another variable, which is much more delicate to analyze.

Related problems appear in semiconductor energy-transport theory, but only partial results have been obtained so far. Let us review these results. The existence of stationary solutions to (3) with the current densities

$$J_n = -\nabla(n\theta) + n\nabla V, \quad J_E = -\kappa_0\nabla\theta + \frac{5}{2}\theta J_n, \quad E = \frac{3}{2}n\theta,$$

close to the constant equilibrium has been shown in [1]. The idea is that in such a situation, the temperature $\theta$ is strictly positive which removes the degeneracy in the term $\nabla(n\theta)$. The parabolic system was investigated in [18, 19], and the global existence of weak solutions was shown without any smallness condition but for a simplifed energy equation. Again, the idea was to prove a uniform positivity bound for the temperature, which removes the degeneracy. A more general result (but without electric potential) was achieved in [25] for the system

$$\partial_t n = \Delta(\theta^\alpha n), \quad \partial_t(\theta n) = \Delta(\theta^{\alpha+1}n) + \frac{n}{\tau}(1-\theta)$$

in a bounded domain, where $0 < \alpha < 1$. The global existence of weak solutions to the corresponding initial-boundary-value problem was proved. Again, the idea is a positivity bound for $\theta$ but this bound required a nontrivial cut-off procedure and several entropy estimates.

In this paper, we make a step forward in the analysis of nonlinear parabolic systems with nonstandard degeneracies by solving (6)-(7) without any positive lower bound for $W$. Since $W$ may vanish, we can expect a gradient estimate for $n$ only on $\{W > 0\}$. Although the quadratic gradient term also possesses $W$ as a factor, the treatment of this term is highly delicate, because of low time regularity. Therefore, we present a result only for a time-discrete version of (6)-(7), namely for its implicit Euler approximation

$$\frac{1}{\triangle t}(n^k - n^{k-1}) = \operatorname{div}\left(\frac{W^k\nabla n^k}{n(1-\eta n^k)}\right), \tag{9}$$

$$\frac{1}{\triangle t}(W^k - W^{k-1}) = \frac{2d-1}{2d}\operatorname{div}\left(\frac{\nabla W^k}{n^k(1-\eta n^k)}\right) - U\frac{W^k|\nabla n^k|^2}{n^k(1-\eta n^k)} \tag{10}$$

for $x \in \mathbb{T}^d$, where $\triangle t > 0$ and $(n^{k-1}, W^{k-1})$ are given functions. We show the existence of a weak solution $(n^k, W^k)$ satisfying $n^k \geq 0$, $W^k \geq 0$ and $W^k n^k$, $W^k \in H^1(\mathbb{T}^d)$; see Theorem 9. In one space dimension and under a smallness assumption on the variance of $W^{k-1}$ and $n^{k-1}$, the strict positivity of $W^k$ can be proved; see Theorem 11.

The existence proof is based on the solution of a regularized and truncated problem by means of the Leray-Schauder fixed-point theorem. Standard elliptic estimates provide bounds uniform in the approximation parameters. The key step is the proof of the strong convergence of the gradient of the particle density. For this, we show a general result for degenerate elliptic problems; see Proposition 8. This result seems to be new. Standard results in the literature need the ellipticity of the differential operator [5]. Unfortunately, we are not able to perform the limit $\triangle t \to 0$ since some estimates in the proof of Proposition 8 are not uniform in $\triangle t$; also see Remark 10 for a discussion.

**Numerical simulations.** The time-discrete system (9)-(10) is discretized by finite differences in one space dimension and solved in an semi-implicit way. The large-time behavior exhibits an interesting phenomenon. If the initial energy $W^0$ is constant and sufficiently large, the solution $(n(t), W(t))$ converges to a constant steady state. However, if the constant $W^0$ is too small, the stationary particle density is nonconstant. In both cases, the time decay is exponential fast, but the decay rate becomes smaller for smaller constants $W^0$ since the diffusion coefficient in (6) is smaller too.

The paper is organized as follows. Section 2 is devoted to the formal derivation of the general energy-transport model and its entropy structure, similar to the semiconductor case [3]. The high-temperature expansion is performed in Section 3, leading to the energy-transport system (6)-(7). The strong convergence of the gradients is shown in Section 4. In Section 5 the existence result is stated and proved. The numerical simulations are presented in Section 6, and the Appendix is concerned with the calculation of some integrals involving the velocity $u(p)$ and energy $\varepsilon(p)$.

## 2. Formal derivation and entropy structure

2.1. **Derivation from a Boltzmann equation.** We consider the following semiclassical Boltzmann transport equation for the distribution function $f(x,p,t)$ in the diffusive scaling:

$$\alpha\partial_t f_\alpha + u\cdot\nabla_x f_\alpha + \nabla V_\alpha\cdot\nabla_p f_\alpha = \frac{1}{\alpha}Q_\alpha(f_\alpha), \quad (x,p)\in\mathbb{R}^d\times\mathbb{T}^d,\ t>0, \tag{11}$$

where $\alpha > 0$ is the Knudsen number [3], $(x,p)$ are the phase-space variables (space and crystal momentum), and $t > 0$ is the time. We recall that the velocity equals $u(p) = \nabla_p\varepsilon(p)$, where the energy $\varepsilon(p)$ is given by (2). The potential $V_\alpha$ is defined by $V_\alpha = -U_0 n_\alpha$. In the physical literature [23], the collision operator $Q_\alpha$ is given by the relaxation-time approximation

$$Q_\alpha(f) = \frac{1}{\tau_\alpha}(\mathcal{F}_f - f),$$

where the function $\mathcal{F}_f$ is determined by maximizing the free energy (18) associated to (11) under the constraints

$$\int_{\mathbb{T}^d}(\mathcal{F}_f - f)dp = 0, \quad \int_{\mathbb{T}^d}(\mathcal{F}_f - f)\varepsilon(p)dp = 0, \tag{12}$$

which express mass and energy conservation during scattering events. The solution of this problem is given by

$$\mathcal{F}_f(x,p,t) = \frac{1}{\eta + \exp(-\lambda_0(x,t) - \lambda_1(x,t)\varepsilon(p))},$$

where $\lambda_0$ and $\lambda_1$ are the Lagrange multipliers and $\eta \geq 0$ is a parameter which may take the values $\eta = 0$ (Maxwell-Boltzmann statistics) or $\eta = 1$ (Fermi-Dirac statistics). The relaxation time $\tau_\alpha \geq 0$ generally depends on the particle density but at this point we do not need to specifiy the dependence.

We show the following result.

**Proposition 1** (Derivation)**.** *Let $f_\alpha$ be a (smooth) solution to the Boltzmann equation* (11)*. We assume that the formal limits $f = \lim_{\alpha\to 0} f_\alpha$, $g = \lim_{\alpha\to 0}(f_\alpha - \mathcal{F}_{f_\alpha})/\alpha$, and $\tau = \lim_{\alpha\to 0}\tau_\alpha$ exist. Then the particle and energy densities*

$$n = n[\mathcal{F}_f] = \int_{\mathbb{T}^d}\mathcal{F}_f dp, \quad E = E[\mathcal{F}_f] = \int_{\mathbb{T}^d}\mathcal{F}_f\varepsilon(p)dp$$

*are solutions to* (3)-(4)*, and the diffusion coefficients $D_{ij} = (D_{ij}^{k\ell}) \in \mathbb{R}^{d\times d}$ are defined by*

$$D_{ij}^{k\ell} = \tau\int_{\mathbb{T}^d} u_k u_\ell \mathcal{F}_f(1 - \eta\mathcal{F}_f)\varepsilon(p)^{i+j}dp, \quad i,j = 0,1, \; k,\ell = 1,\ldots,d.$$

The proof of the proposition is similar to those of Propositions 1 and 2 in [17]. For the convenience of the reader, we present the (short) proof.

*Proof.* To derive the balance equations, we multiply the Boltzmann equation (11) by 1 and $\varepsilon$, respectively, and integrate over $\mathbb{T}^d$:

$$\partial_t n[f_\alpha] + \frac{1}{\alpha}\mathrm{div}_x\int_{\mathbb{T}^d} u f_\alpha dp = 0, \tag{13}$$

$$\partial_t E[f_\alpha] + \frac{1}{\alpha}\mathrm{div}_x\int_{\mathbb{T}^d}\varepsilon u f_\alpha dp - \frac{1}{\alpha}\nabla_x V_\alpha\cdot\int_{\mathbb{T}^d} u f_\alpha dp = 0. \tag{14}$$

The integrals involving the collision operator vanish in view of mass and energy conservation; see (12). We have integrated by parts in the last integral on the left-hand side of (14). Next, we insert the Chapman-Enskog expansion $f_\alpha = \mathcal{F}_{f_\alpha} + \alpha g_\alpha$ (which in fact defines $g_\alpha$) in (13)-(14) and observe that the function $p \mapsto u(p)\varepsilon(p)^j\mathcal{F}_{f_\alpha}(p)$ is odd for any $j \in \mathbb{N}_0$ such that its integral over $\mathbb{T}^d$ vanishes. This leads to

$$\partial_t n[\mathcal{F}_{f_\alpha}] + \alpha\partial_t n[g_\alpha] + \mathrm{div}_x\int_{\mathbb{T}^d} u g_\alpha dp = 0,$$

$$\partial_t E[\mathcal{F}_{f_\alpha}] + \alpha\partial_t E[g_\alpha] + \mathrm{div}_x\int_{\mathbb{T}^d} u\varepsilon g_\alpha dp - \nabla_x V\cdot\int_{\mathbb{T}^d} u g_\alpha dp = 0.$$

Passing to the formal limit $\alpha \to 0$ gives the balance equations (3) with

$$J_n = \int_{\mathbb{T}^d} ugdp, \quad J_E = \int_{\mathbb{T}^d} u\varepsilon gdp. \tag{15}$$

To specify the current densities, we insert the Chapman-Enskog expansion in (11),

$$\alpha\partial_t(\mathcal{F}_{f_\alpha} + \alpha g_\alpha) + u \cdot \nabla_x(\mathcal{F}_{f_\alpha} + \alpha g_\alpha) + \nabla_x V \cdot \nabla_p(\mathcal{F}_{f_\alpha} + \alpha g_\alpha) = -\frac{g_\alpha}{\tau_\alpha},$$

and perform the formal limit $\alpha \to 0$,

$$u \cdot \nabla_x \mathcal{F}_f + \nabla_x V \cdot \nabla_p \mathcal{F}_f = -\frac{g}{\tau}. \tag{16}$$

A straightforward computation shows that

$$\nabla_x \mathcal{F}_f = \mathcal{F}_f(1-\eta\mathcal{F}_f)(\nabla_x\lambda_0 + \varepsilon\nabla_x\lambda_1), \quad \nabla_p\mathcal{F}_f = \mathcal{F}_f(1-\eta\mathcal{F}_f)u\lambda_1,$$

and inserting this into (16) gives an explicit expression for $g$:

$$g = -\tau\mathcal{F}_f(1-\eta\mathcal{F}_f)\big(u\cdot\nabla_x\lambda_0 + \varepsilon u\cdot\nabla_x\lambda_1 + \lambda_1\nabla_x V\cdot u\big).$$

Therefore, the current densities (15) lead to (4). This finishes the proof. □

In the following we write $\mathcal{F}_f = \mathcal{F}(\lambda)$, where

$$\mathcal{F}(\lambda) = \frac{1}{\eta + \exp(-\lambda_0 - \lambda_1\varepsilon(p))}, \quad \lambda = (\lambda_0, \lambda_1) \in \mathbb{R}^2, \ p \in \mathbb{T}^d. \tag{17}$$

**Proposition 2** (Diffusion matrix)**.** *The diffusion matrix $\mathcal{D} = (D_{ij}) \in \mathbb{R}^{2d\times 2d}$ is symmetric and positive definite.*

*Proof.* The proof is similar to Proposition 3 in [17]. Let $z = (w, y) \in \mathbb{R}^{2d}$ with $w$, $y \in \mathbb{R}^d$. Then

$$\begin{aligned} z^\top\mathcal{D}z &= w^\top D_{00}w + 2w^\top D_{01}y + y^\top D_{11}y \\ &= \int_{\mathbb{T}^d} \mathcal{F}(1-\eta\mathcal{F})\sum_{i=1}^d (u_iw_i + \varepsilon u_iy_i)\sum_{j=1}^d (u_jw_j + \varepsilon u_jy_j)dp \\ &= \int_{\mathbb{T}^d} \mathcal{F}(1-\eta\mathcal{F})\sum_{i=1}^d \big|u_i(w_i + \varepsilon y_i)\big|^2 dp \ge 0. \end{aligned}$$

Since $D_{ij}^{k\ell}$ is symmetric in $(i,j)$ and $(k,\ell)$, the symmetry of $\mathcal{D}$ is clear. □

## 2.2. **Entropy structure.**

The entropy structure of (3)-(4) follows from the abstract framework presented in [17]. In the following, we make this framework explicit. First, we introduce the entropy

$$\begin{aligned} H(t) &= \int_{\mathbb{R}^d} h(\lambda)dx, \quad \text{where} \\ h(\lambda) &= \int_{\mathbb{T}^d} \big(\mathcal{F}\log\mathcal{F} + \eta^{-1}(1-\eta\mathcal{F})\log(1-\eta\mathcal{F})\big)dp. \end{aligned}$$

The entropy density $h$ can be reformulated as

$$\begin{aligned} h(\lambda) &= \int_{\mathbb{T}^d}\bigg(\mathcal{F}\log\frac{\mathcal{F}}{1-\eta\mathcal{F}} - \frac{1}{\eta}\log\frac{1}{1-\eta\mathcal{F}}\bigg)dp \\ &= \int_{\mathbb{T}^d}\big(\mathcal{F}(\lambda_0+\lambda_1\varepsilon) - \frac{1}{\eta}\log(1+\eta e^{\lambda_0+\lambda_1\varepsilon})\big)dp \\ &= n\lambda_0 + E\lambda_1 - \frac{1}{\eta}\int_{\mathbb{T}^d}\log(1+\eta e^{\lambda_0+\lambda_1\varepsilon})dp. \end{aligned} \tag{18}$$

The following result shows that the entropy is nonincreasing in time.

**Proposition 3** (Entropy structure). *It holds that*

$$\frac{dH}{dt} = -\int_{\mathbb{R}^d}\sum_{i,j=0}^{1}\nabla\mu_i^\top L_{ij}\nabla\mu_j dx \le 0, \tag{19}$$

*where* $\mu_0 = \lambda_0 + \lambda_1 V$ *and* $\mu_1 = \lambda_1$ *are the so-called dual entropy variables and*

$$L_{00} = D_{00}, \quad L_{01} = L_{10} = D_{01} - D_{00}V, \quad L_{11} = D_{11} - 2D_{01}V + D_{00}V^2. \tag{20}$$

*Proof.* Identity (18) implies that

$$\frac{\partial h}{\partial\lambda_i} = \frac{\partial n}{\partial\lambda_i}\lambda_0 + \frac{\partial E}{\partial\lambda_i}\lambda_1, \quad i = 0, 1,$$

and consequently,

$$\begin{aligned} \partial_t h(\lambda) &= \frac{\partial h}{\partial\lambda_0}\partial_t\lambda_0 + \frac{\partial h}{\partial\lambda_1}\partial_t\lambda_1 \\ &= \lambda_0\bigg(\frac{\partial n}{\partial\lambda_0}\partial_t\lambda_0 + \frac{\partial n}{\partial\lambda_1}\partial_t\lambda_1\bigg) + \lambda_1\bigg(\frac{\partial E}{\partial\lambda_0}\partial_t\lambda_0 + \frac{\partial E}{\partial\lambda_1}\partial_t\lambda_1\bigg) \\ &= \lambda_0\partial_t n + \lambda_1\partial_t E. \end{aligned}$$

Therefore, using (3)-(4) and integration by parts,

$$\begin{aligned} \frac{dH}{dt} &= \int_{\mathbb{R}^d}\partial_t h(\lambda)dx = \int_{\mathbb{R}^d}\big(J_n\cdot\nabla\lambda_0 + J_E\cdot\nabla\lambda_1 + \nabla V\cdot J_n\lambda_1\big)dx \\ &= -\int_{\mathbb{R}^d}\big(D_{00}|\nabla\mu_0|^2 + 2(D_{01}-D_{00}V)\nabla\mu_0\cdot\nabla\mu_1 \\ &\qquad + (D_{11} - 2D_{01}V + D_{00}V^2)|\nabla\mu_1|^2\big)dx, \end{aligned}$$

which proves the identity in (19). Using the positive definiteness of $\mathcal{D}$, a computation shows that $(L_{ij})$ is positive definite too, and the inequality in (19) follows. □

2.3. **Singularities and degeneracies in the energy-transport system.** We denote by $n$ and $E$ the particle and energy densities depending on the dual entropy variable $\mu = (\mu_0, \mu_1) = (\lambda_0 + V\lambda_1, \lambda_1)$. We have the (implicit) formulation

$$n(\mu) = \int_{\mathbb{T}^d} \frac{dp}{\eta + \exp(-\mu_0 + U_0 n(\mu) - \mu_1 \varepsilon(p))},$$
$$E(\mu) = \int_{\mathbb{T}^d} \frac{\varepsilon(p)dp}{\eta + \exp(-\mu_0 + U_0 n(\mu) - \mu_1 \varepsilon(p))}.$$

**Lemma 4.** *Let $\omega_i(\mu) := \int_{\mathbb{T}^d} \mathcal{F}(1-\eta\mathcal{F})\varepsilon(p)^i dp$, $i \in \mathbb{N}_0$. Then*

$$\det \frac{\partial(n,E)}{\partial\mu} = \frac{\omega_0\omega_2 - \omega_1^2}{1 - U_0\mu_1\omega_0}.$$

*Proof.* We differentiate

$$\frac{\partial n}{\partial\mu_0} = \left(1 + U_0\mu_0\frac{\partial n}{\partial\mu_0}\right)\omega_0, \quad \frac{\partial n}{\partial\mu_1} = Un + U_0\mu_1\omega_0\frac{\partial n}{\partial\mu_0} + \omega_1.$$

This gives after a rearrangement

$$\frac{\partial n}{\partial\mu_0} = \frac{\omega_0}{1 - U_0\mu_1\omega_0}, \quad \frac{\partial n}{\partial\mu_1} = \frac{Un\omega_0 + \omega_1}{1 - U_0\mu_1\omega_0}.$$

In a similar way, we obtain

$$\frac{\partial E}{\partial\mu_0} = \frac{\omega_1}{1 - U_0\mu_1\omega_0}, \quad \frac{\partial E}{\partial\mu_1} = \frac{Un\omega_1 + U_0\mu_1(\omega_1 - \omega_0\omega_2) + \omega_2}{1 - U_0\mu_1\omega_0},$$

and with

$$\det \frac{\partial(n,E)}{\partial\mu} = \frac{\partial n}{\partial\mu_0}\frac{\partial E}{\partial\mu_1} - \frac{\partial n}{\partial\mu_1}\frac{\partial E}{\partial\mu_0} = \frac{(\omega_0\omega_2 - \omega_1^2)(1 - U_0\mu_1\omega_0)}{(1 - U_0\mu_1\omega_0)^2},$$

the conclusion follows. ☐

Since $\mu_1$ can be positive and $\omega_0 > 0$, the expression $1 - U_0\mu_1\omega_0$ may vanish, so the determinant of $\partial(n,E)/\partial\mu$ may be not finite. Moreover, the numerator of the determinant may vanish, and the function $\mu \mapsto (n,E)$ may be not invertible. This is made more explicit in the following remark.

**Remark 5** (Case $\mu = 0$)**.** In the Maxwell-Boltzmann case, we can make the numerator of the determinant in Lemma 4 explicit. Indeed, it is clear that $\omega_0 = n$ and $\omega_1 = E$. For the computation of $\omega_2$, we observe first that

$$\begin{aligned} n &= \int_{\mathbb{T}^d} \mathcal{F}dp = \exp(\mu_0 - U_0 n\mu_1)\prod_{k=1}^d \int_{\mathbb{T}} \exp\big(-2\varepsilon_0\cos(2\pi p_k)\big)dp_k \\ &= \exp(\mu_0 - U_0 n\mu_1)I_0^d, \end{aligned} \tag{21}$$

$$E = \int_{\mathbb{T}^d} \mathcal{F}\varepsilon dp = -2\varepsilon_0\exp(\mu_0 - U_0 n\mu_1)\sum_{i=1}^d \prod_{k\neq i}^d \int_{\mathbb{T}} \exp\big(-2\varepsilon_0\cos(2\pi p_k)v\big)dp_k$$

$$\times \int_{\mathbb{T}} \exp\big(-2\varepsilon_0\cos(2\pi p_i)\big)\cos(2\pi p_i)dp_i$$

$$= -2d\varepsilon_0\exp(\mu_0 - U_0 n\mu_1)I_0^{d-1}I_1, \tag{22}$$

where, by symmetry,

$$I_0 := \int_{\mathbb{T}} \exp\big(-2\varepsilon_0\cos(2\pi p_1)\big)dp_1, \quad I_1 := \int_{\mathbb{T}} \exp\big(-2\varepsilon_0\cos(2\pi p_1)\big)\cos(2\pi p_1)dp_1.$$

Now, we have

$$\omega_2 = 4\varepsilon_0^2\sum_{i=1}^d\int_{\mathbb{T}^d}\cos^2(2\pi p_i)\mathcal{F}dp + 8\varepsilon_0^2\sum_{i=1}^d\sum_{j=1,\, j\neq i}^d\int_{\mathbb{T}^d}\cos(2\pi p_i)\cos(2\pi p_j)\mathcal{F}dp.$$

Expanding the exponentials in $\mathcal{F}$ and using (21)-(22), we find that

$$\int_{\mathbb{T}^d}\cos(2\pi p_i)\cos(2\pi p_j)\mathcal{F}dp = \exp(\mu_0 - U_0 n\mu_1)I_0^{d-2}I_1^2 = \frac{E^2}{(2d\varepsilon_0)^2 n},$$

and this expression is independent of $i \neq j$. For $i = j$, we use the identity $\sin(2\pi p_i)\mathcal{F} = (\partial\mathcal{F}/\partial p_i)/(4\pi\varepsilon_0\mu_1)$ and integration by parts:

$$\int_{\mathbb{T}^d}\cos^2(2\pi p_i)\mathcal{F}dp = \int_{\mathbb{T}^d}(1-\sin^2(2\pi p_i))\mathcal{F}dp = n - \int_{\mathbb{T}^d}\frac{\sin(2\pi p_i)}{4\pi\varepsilon_0\mu_1}\frac{\partial\mathcal{F}}{\partial p_i}dp$$
$$= n + \int_{\mathbb{T}^d}\frac{\cos(2\pi p_i)}{2\varepsilon_0\mu_1}\mathcal{F}dp.$$

Summing over $i = 1, \ldots, n$ gives

$$4\varepsilon_0^2\sum_{i=1}^d\int_{\mathbb{T}^d}\cos^2(2\pi p_i)\mathcal{F}dp = 4\varepsilon_0^2 dn - \frac{E}{\mu_1}.$$

We conclude that $\omega_2 = 4\varepsilon_0^2 dn - E/\mu_1 + (d-1)E^2/(dn)$ and consequently,

$$\omega_0\omega_2 - \omega_1^2 = 4\varepsilon_0^2 dn - \frac{En}{\mu_1} - \frac{E^2}{d}.$$

For certain values of $\mu_1$ or $(n, E)$, this expression may vanish such that $\det\partial(n,E)/\partial\mu = 0$ at these values. This shows that the relation between $(n, E)$ and $\mu$ needs to treated with care. □

**Remark 6** (Degeneracy in the entropy production)**.** A tedious computation, detailed in [6, Chapter 8.4], shows that the entropy production can be written as

$$\sum_{i,j=0}^1\int_{\mathbb{R}^d}\nabla\mu_i^\top L_{ij}\nabla\mu_j dx = g_1(\lambda)(1 - U_0\mu_1\omega_0)|\nabla n|^2 + g_2(\lambda)|g_3(\lambda)\nabla n - \nabla E|^2,$$

where $g_i(\lambda)$, $i = 1, 2, 3$, are functions depending on $\omega_i$, defined in Lemma 4, and on

$$\Gamma_i = \int_{\mathbb{T}^d}\varepsilon(p)^i|\nabla\varepsilon|^2\mathcal{F}(1-\eta\mathcal{F})dp, \quad i = 0, 1, 2.$$

The above formula shows that we lose the gradient estimate if $1 - U_0\mu_1\omega_0 = 0$. □

**Remark 7** (Comparison with the semiconductor case)**.** For the semiconductor energy-transport equations in the parabolic band approximation, we do not face the singularities and degeneracies occuring in the model for optical lattices. Indeed, let the potential $V$ be given (to simplify). According to Example 6.8 in [15], we have

$$n = \mu_1^{-3/2}\exp(\mu_0 + \mu_1 V), \quad E = \frac{3}{2}\mu_1^{-5/2}\exp(\mu_0 + \mu_1 V).$$

Then

$$\det\frac{\partial(n,E)}{\partial\mu} = \det\begin{pmatrix} n & nV - E \\ E & -5E/(2\mu_1) + EV \end{pmatrix} = -\frac{2}{3}E^2,$$

which is nonzero as long as $E > 0$. Furthermore, by Remark 8.12 in [15], it holds that $\omega_0 = n$, $\omega_1 = E$, and $\omega_2 = 5E^2/(3n)$, and so

$$\omega_0\omega_2 - \omega_1^2 = \frac{2}{3}E^2.$$

This expression is degenerate only at the boundary of the domain of definition (i.e. at $E = 0$). Often, such kind of degeneracies may be handled; an important example is the porous-medium equation. In the case of optical lattices, the degeneracy may occur in the interior of the domain of definition, which is much more delicate. □

## 3. High-temperature expansion

The Lagrange multiplier $\lambda_1$ is interpreted as the negative inverse temperature, so high temperatures correspond to small values of $|\lambda_1|$. In this section, we perform a high-temperature expansion of (3)-(4), i.e., we expand $\mathcal{F}(\lambda)$ around $(\lambda_0, 0)$ for small $|\lambda_1|$ up to first order. Our ansatz is

$$\begin{aligned} \mathcal{F}(\lambda) &= \mathcal{F}(\lambda_0,0) + \frac{\partial\mathcal{F}}{\partial\lambda_1}(\lambda_0,0)\lambda_1 + O(\lambda_1^2) \\ &= \mathcal{F}(\lambda_0,0) + \varepsilon\mathcal{F}(\lambda_0,0)\big(1 - \eta\mathcal{F}(\lambda_0,0)\big)\lambda_1 + O(\lambda_1^2). \end{aligned} \tag{23}$$

### 3.1. Zeroth-order expansion.

At zeroth-order, we have by (15), (16), and using formula (67) from the appendix,

$$\begin{aligned} n &= \int_{\mathbb{T}^d}\mathcal{F}(\lambda_0,0)dp + O(\lambda_1) = \mathcal{F}(\lambda_0,0) + O(\lambda_1), \\ J_n &= -\tau\int_{\mathbb{T}^d}\big(u\otimes u\nabla_x\mathcal{F}(\lambda_0,0) + u\nabla_x V\cdot\nabla_p\mathcal{F}(\lambda_0,0)\big)dp \\ &= -\tau\int_{\mathbb{T}^d}(u\otimes u)ndp + O(\lambda_1) = -\frac{\tau}{2}(4\pi\varepsilon_0)^2\nabla n + O(\lambda_1). \end{aligned}$$

Therefore, up to order $O(\lambda_1)$, we infer that

$$\partial_t n = \frac{1}{2}(4\pi\varepsilon_0)^2\operatorname{div}(\tau\nabla n), \quad x\in\mathbb{R}^d,\ t > 0,$$

At high temperature, the relaxation time depends on the particle density in a nonlinear way, $\tau = \tau_0/(n(1-\eta n))$ [20]. At low densities, i.e. $\eta = 0$, we obtain the logarithmic diffusion equation

$$\partial_t n = \varepsilon_1 \Delta \log n, \quad t > 0, \quad n(0,\cdot) = n_0 \quad \text{in } \mathbb{R}^d,$$

where $\varepsilon_1 = \frac{1}{2}\tau_0(4\pi\varepsilon_0)^2$. We already mentioned in the introduction that the (smooth) solution to this equation in two space dimensions loses mass, which is unphysical. Therefore, we compute the next-order expansion.

3.2. **First-order expansion.** We calculate, using (66),

$$\begin{aligned} n &= \int_{\mathbb{T}^d} \big(\mathcal{F}(\lambda_0,0) + \varepsilon(\mathcal{F}(1-\eta\mathcal{F}))(\lambda_0,0)\lambda_1\big)dp + O(\lambda_1^2) \\ &= \mathcal{F}(\lambda_0,0) + (\mathcal{F}(1-\eta\mathcal{F}))(\lambda_0,0)\lambda_1 \int_{\mathbb{T}^d} \varepsilon(p)dp + O(\lambda_1^2) = \mathcal{F}(\lambda_0,0) + O(\lambda_1^2), \\ E &= \mathcal{F}(\lambda_0,0)\int_{\mathbb{T}^d} \varepsilon(p)dp + (\mathcal{F}(1-\eta\mathcal{F}))(\lambda_0,0)\lambda_1 \int_{\mathbb{T}^d} \varepsilon(p)^2 dp + O(\lambda_1^2) \\ &= 2d\varepsilon_0^2(\mathcal{F}(1-\eta\mathcal{F}))(\lambda_0,0)\lambda_1 + O(\lambda_1^2) = 2d\varepsilon_0^2(\mathcal{F}(1-\eta\mathcal{F}))(\lambda_0,0)\lambda_1 + O(\lambda_1^2). \end{aligned}$$

Therefore, by (23),

$$\mathcal{F}(\lambda) = n + \varepsilon\frac{E}{2d\varepsilon_0^2} + O(\lambda_1^2),$$

and (16) and $\nabla_p \varepsilon = u$ give, up to order $O(\lambda_1^2)$,

$$\begin{aligned} g &= -\tau\big(u\cdot\nabla_x\mathcal{F}(\lambda) + \nabla_x V\cdot\nabla_p\mathcal{F}(\lambda)\big) \\ &= -\tau u\cdot\left(\nabla_x n + \frac{E}{2d\varepsilon_0^2}\nabla_x V\right) - \tau\varepsilon u\cdot\frac{\nabla_x E}{2d\varepsilon_0^2}. \end{aligned}$$

Then, by (15)-(16) and taking into account (67)–(69), we infer that, again up to first order,

$$\begin{aligned} J_n &= -\tau\int_{\mathbb{T}^d} u\otimes u dp \nabla_x n - \tau\int_{\mathbb{T}^d} u\otimes u dp\nabla_x V\frac{E}{d\varepsilon_1} = -8\pi^2\tau\varepsilon_0^2\nabla_x n - \frac{4\pi^2}{d}\tau E\nabla_x V, \\ J_E &= -\frac{\tau}{2d\varepsilon_0^2}\int_{\mathbb{T}^d}\varepsilon^2(u\otimes u)dp\nabla_x E = -4\pi^2\frac{2d-1}{d}\varepsilon_0^2\nabla_x E. \end{aligned}$$

Therefore, the first-order expansion leads to

$$\partial_t n = 8\pi^2\varepsilon_0^2 \operatorname{div}\left(\tau\nabla n + \frac{\tau}{2d\varepsilon_0^2}E\nabla V\right), \tag{24}$$

$$\partial_t E = 8\pi^2\varepsilon_0^2\frac{2d-1}{2d}\operatorname{div}(\tau\nabla E) - 8\pi^2\varepsilon_0^2\tau\nabla V\cdot\left(\nabla n + \frac{E}{2d\varepsilon_0^2}\nabla V\right). \tag{25}$$

We rescale the time by $t_s = (8\pi^2\varepsilon_0^2\tau_0)t$ and introduce $U = U_0/(2d\varepsilon_0^2)$ and $W = 1 - UE$. Then, writing again $t$ instead of $t_s$, system (24)-(25) becomes

$$\partial_t n = \operatorname{div}\left(\frac{W\nabla n}{n(1-\eta n)}\right), \quad \partial_t W = \frac{2d-1}{2d}\operatorname{div}\left(\frac{\nabla W}{n(1-\eta n)}\right) - U\frac{W|\nabla n|^2}{n(1-\eta n)}. \tag{26}$$

The existence of weak solutions to a time-discrete version of (26), together with periodic boundary conditions, is shown in Section 5.

## 4. A strong convergence result for the gradient

The key tool of the existence analysis of Section 5 is the following result on the strong convergence of the gradients of certain approximate solutions for the following equation. Let $\triangle t > 0$, $\overline{n} \in L^\infty(\Omega)$, $y \in L^\infty(\Omega) \cap H^1(\Omega)$, and $\Psi \in C^1(\mathbb{R})$. We consider the equation

$$\frac{1}{\triangle t}(n - \overline{n}) = \operatorname{div}\big(\nabla(y\Psi(n)) - \Psi(n)\nabla y\big) \tag{27}$$

for $n \in L^\infty(\Omega)$ such that $y\Psi(n) \in H^1(\Omega)$.

**Proposition 8.** *Let $\Omega \subset \mathbb{R}^d$ be a bounded domain, $\varepsilon > 0$, $\triangle t > 0$, let $\overline{n} \in L^\infty(\Omega)$ be such that $\overline{n} \geq 0$ in $\Omega$, and let $\Psi \in C^1(\mathbb{R})$ satisfy $\Psi' > 0$. Let $(y_\varepsilon)$ be a bounded sequence in $H^1(\Omega)$ satisfying $y_\varepsilon \geq C(\varepsilon) > 0$ for some $C(\varepsilon) > 0$, $y_\varepsilon \to y$ strongly in $L^2(\Omega)$ and weakly in $H^1(\Omega)$ as $\varepsilon \to 0$. Furthermore, let $n_\varepsilon \in L^2(\Omega)$ with $\Psi(n_\varepsilon) \in H^1(\Omega)$ be a weak solution to*

$$\frac{1}{\triangle t}\int_\Omega (n_\varepsilon - \overline{n})\phi dx + \int_\Omega y_\varepsilon \nabla\Psi(n_\varepsilon)\cdot\nabla\phi dx = 0 \tag{28}$$

*for all $\phi \in H^1(\Omega)$. Then there exist a function $n \in L^\infty(\Omega)$ such that $y\Psi(n) \in H^1(\Omega)$, being a weak solution of (27), and a subsequence of $(n_\varepsilon)$, which is not relabeled, such that, as $\varepsilon \to 0$,*

$$\begin{aligned} y^{1/2}y_\varepsilon^{1/2}\nabla\Psi(n_\varepsilon) &\to \nabla(y\Psi(n)) - \Psi(n)\nabla y \quad \textit{strongly in } L^2(\Omega), \\ 1_{\{y>0\}}(n_\varepsilon - n) &\to 0 \quad \textit{strongly in } L^2(\Omega). \end{aligned}$$

*Proof. Step 1.* First, we derive some uniform bounds. Set $M = \|\overline{n}\|_{L^\infty(\Omega)}$. Taking $(\Psi(n_\varepsilon) - \Psi(M))_+ = \max\{0, \Psi(n_\varepsilon) - \Psi(M)\}$ as a test function in (28), we find that

$$\begin{aligned} \frac{1}{\triangle t}\int_\Omega \big((n_\varepsilon - M) - (\overline{n} - M)\big)(\Psi(n_\varepsilon) - \Psi(M))_+ dx& \\ + \int_\Omega y_\varepsilon \nabla\Psi(n_\varepsilon)\cdot\nabla(\Psi(n_\varepsilon) - \Psi(M))_+ &= 0. \end{aligned}$$

Since $-(\overline{n} - M)(\Psi(n_\varepsilon) - \Psi(M))_+ \geq 0$, it follows that

$$\frac{1}{\triangle t}\int_\Omega (n_\varepsilon - M)_+(\Psi(n_\varepsilon) - \Psi(M))_+ dx + \int_\Omega y_\varepsilon|\nabla(\Psi(n_\varepsilon) - \Psi(M))_+|^2 dx \leq 0$$

and hence, $n_\varepsilon \leq M$ in $\Omega$. In a similar way, using $\Psi(n_\varepsilon)_- = \min\{0, \Psi(n_\varepsilon)\}$ as a test function and using $\overline{n} \geq 0$, we infer that $n_\varepsilon \geq 0$. This shows that $(n_\varepsilon)$ is bounded in $L^\infty(\Omega)$. Hence, there exists a subsequence which is not relabeled such that, as $\varepsilon \to 0$,

$$n_\varepsilon \rightharpoonup^* n \quad \text{weakly* in } L^\infty(\Omega) \tag{29}$$

for some function $n \in L^\infty(\Omega)$. Since $\Phi'$ is positive and continuous on $\mathbb{R}$, the boundedness of $(n_\varepsilon)$ in $L^\infty(\Omega)$ implies that there exists a constant $C > 0$ such that $1/C \leq \Psi'(n_\varepsilon) \leq C$

for all $\varepsilon > 0$. Next, we choose the test function $\phi = \Psi(n_\varepsilon)$ in (28) and use $\Psi(n_\varepsilon) \le \Psi(M)$ to find that

$$\frac{1}{\triangle t}\int_\Omega n_\varepsilon \Psi(n_\varepsilon)dx + \int_\Omega y_\varepsilon |\nabla\Psi(n_\varepsilon)|^2 dx \le \frac{\Psi(M)}{\triangle t}\int_\Omega \overline{n}dx.$$

We deduce that $(y_\varepsilon^{1/2}\nabla\Psi(n_\varepsilon))$ is bounded in $L^2(\Omega)$ and, for a subsequence,

$$y_\varepsilon^{1/2}\nabla\Psi(n_\varepsilon) \rightharpoonup \xi \quad \text{weakly in } L^2(\Omega) \tag{30}$$

for some function $\xi \in L^2(\Omega)$. Now, let $\phi$ be a smooth test function. Then we can take the limit, for a subsequence, in (28) and obtain

$$\frac{1}{\triangle t}\int_\Omega (n - \overline{n})\phi dx + \int_\Omega y^{1/2}\xi \cdot \nabla\phi dx = 0. \tag{31}$$

Note that this equation holds also for all $\phi \in H^1(\Omega)$.

*Step 2.* As $(y_\varepsilon)$ is strongly converging in $L^2(\Omega)$, we deduce from (29) that $y_\varepsilon^{1/2}(n_\varepsilon - n) \rightharpoonup 0$ weakly in $L^2(\Omega)$. We claim that this convergence is even strong. Indeed, the sequence

$$\nabla(y_\varepsilon n_\varepsilon^2) = 2\frac{y_\varepsilon^{1/2} n_\varepsilon}{\Psi'(n_\varepsilon)} y_\varepsilon^{1/2}\nabla\Psi(n_\varepsilon) + n_\varepsilon^2 \nabla y_\varepsilon$$

is uniformly bounded in $L^2(\Omega)$. Thus, $(y_\varepsilon n_\varepsilon^2)$ is bounded in $H^1(\Omega)$ and by compactness, for a subsequence, $y_\varepsilon n_\varepsilon^2 \to \zeta \ge 0$ strongly in $L^2(\Omega)$ or $y_\varepsilon^{1/2} n_\varepsilon \to \zeta^{1/2}$ strongly in $L^4(\Omega)$. The strong convergence of $(y_\varepsilon^{1/2})$ in $L^2(\Omega)$ and the weak* convergence of $(n_\varepsilon)$ in $L^\infty(\Omega)$ imply that $y_\varepsilon^{1/2} n_\varepsilon \rightharpoonup y^{1/2} n$ weakly in $L^2(\Omega)$. Therefore, $\zeta^{1/2} = y^{1/2} n$ and

$$y_\varepsilon^{1/2}(n_\varepsilon - n) \to 0 \quad \text{strongly in } L^2(\Omega). \tag{32}$$

This proves the claim.

*Step 3.* The next goal is to show that

$$1_{\{y>0\}}(n_\varepsilon - n) \to 0 \quad \text{strongly in } L^2(\Omega). \tag{33}$$

Taking into account (32), the strong convergence of $(y_\varepsilon^{1/2})$ in $L^2(\Omega)$, and the $L^\infty$ bound for $(n_\varepsilon)$, it follows that

$$\|y^{1/2}(n_\varepsilon - n)\|_{L^2(\Omega)} \le \|y_\varepsilon^{1/2}(n_\varepsilon - n)\|_{L^2(\Omega)} + \|y^{1/2} - y_\varepsilon^{1/2}\|_{L^2(\Omega)}\|n_\varepsilon - n\|_{L^\infty(\Omega)}$$

converges to zero. Thus, for a subsequence, $y^{1/2}(n_\varepsilon - n) \to 0$ a.e. in $\Omega$ and consequently, $n_\varepsilon - n \to 0$ a.e. in $\{y > 0\}$. Then the a.e. pointwise convergence $1_{\{y>0\}}(n_\varepsilon - n) \to 0$ and the dominated convergence theorem show (33).

*Step 4.* We wish to identify $\xi$ in (30) and (31). The bounds for $(\nabla y_\varepsilon)$ and $(y_\varepsilon\nabla\Psi(n_\varepsilon))$ in $L^2(\Omega)$ show that $(\nabla(y_\varepsilon\Psi(n_\varepsilon)))$ is bounded in $L^2(\Omega)$ and so, $(y_\varepsilon\Psi(n_\varepsilon))$ is bounded in $H^1(\Omega)$. By compactness, for a subsequence, $\nabla(y_\varepsilon\Psi(n_\varepsilon)) \rightharpoonup \nabla(y\Psi(n))$ weakly in $L^2(\Omega)$ and $y_\varepsilon\Psi(n_\varepsilon) \to y\Psi(n)$ strongly in $L^2(\Omega)$. We can identify the limit since $y_\varepsilon \to y$ strongly in $L^2(\Omega)$ and $\Psi(n_\varepsilon) \rightharpoonup^* \theta$ weakly* in $L^\infty(\Omega)$ with $\theta = \Psi(n)$ in $\{y > 0\}$ lead to $y_\varepsilon\Psi(n_\varepsilon) \rightharpoonup$

$y\Psi(n)$ weakly in $L^2(\Omega)$. Moreover, since $\Psi(n_\varepsilon) \to \Psi(n)$ strongly in $L^2(\Omega)$ and $\nabla y_\varepsilon \to \nabla y$ weakly in $L^2(\Omega)$, we infer that

$$y_\varepsilon \nabla\Psi(n_\varepsilon) = \nabla(y_\varepsilon\Psi(n_\varepsilon)) - \Psi(n_\varepsilon)\nabla y_\varepsilon \rightharpoonup \nabla(y\Psi(n)) - \Psi(n)\nabla y \quad \text{weakly in } L^2(\Omega). \tag{34}$$

Here, we have used additionally that $(y_\varepsilon\nabla\Psi(n_\varepsilon))$ is bounded in $L^2(\Omega)$ and that $L^1(\Omega)$ is dense in $L^2(\Omega)$. Similarly as above, we deduce that

$$\|(y^{1/2} - y_\varepsilon^{1/2})y_\varepsilon^{1/2}\nabla\Psi(n_\varepsilon)\|_{L^1(\Omega)} \le \|(y^{1/2} - y_\varepsilon^{1/2})\|_{L^2(\Omega)}\|y_\varepsilon^{1/2}\nabla\Psi(n_\varepsilon)\|_{L^2(\Omega)} \tag{35}$$

converges to zero. Therefore, $y^{1/2}\xi = \nabla(y\Psi(n)) - \Psi(n)\nabla y$ in $\Omega$.

*Step 5.* We obtain from (34) that

$$y^{1/2}\big(y_\varepsilon\nabla\Psi(n_\varepsilon) - y^{1/2}\xi\big) \rightharpoonup 0 \quad \text{weakly in } L^2(\Omega)$$

and consequently,

$$\begin{aligned}\int_\Omega y_\varepsilon\xi\cdot\big(y^{1/2}\nabla\Psi(n_\varepsilon) - \xi\big)dx &= \int_\Omega \xi\cdot y^{1/2}\big(y_\varepsilon\nabla\Psi(n_\varepsilon) - y^{1/2}\xi\big)dx \\ &\quad + \int_\Omega |\xi|^2\big(y - y_\varepsilon\big)dx \to 0,\end{aligned} \tag{36}$$

applying the dominated convergence theorem to the last integral. Furthermore, using the test function $\phi = y(\Psi(n_\varepsilon) - \Psi(n))$ in (28),

$$\begin{aligned}&\bigg|\int_\Omega y_\varepsilon y^{1/2}\nabla\Psi(n_\varepsilon)\cdot\big(y^{1/2}\nabla\Psi(n_\varepsilon) - \xi\big)dx\bigg| \\ &= \bigg|\int_\Omega \Big(y_\varepsilon y|\nabla\Psi(n_\varepsilon)|^2 - y_\varepsilon\nabla\Psi(n_\varepsilon)\cdot\big(\nabla(y\Psi(n)) - \Psi(n)\nabla y\big)\Big)dx\bigg| \\ &= \bigg|\int_\Omega y_\varepsilon\nabla\Psi(n_\varepsilon)\cdot\nabla(y(\Psi(n_\varepsilon) - \Psi(n))) - \int_\Omega y_\varepsilon^{1/2}\nabla\Psi(n_\varepsilon)\cdot\nabla y\big(y_\varepsilon^{1/2}(\Psi(n_\varepsilon) - \Psi(n))\big)dx\bigg| \\ &= \bigg|\frac{1}{\triangle t}\int_\Omega(\overline{n} - n_\varepsilon)y(\Psi(n_\varepsilon) - \Psi(n))dx - \int_\Omega y_\varepsilon^{1/2}\nabla\Psi(n_\varepsilon)\cdot\nabla y\big(y_\varepsilon^{1/2}(\Psi(n_\varepsilon) - \Psi(n))\big)dx\bigg| \\ &\le \frac{1}{\triangle t}\|n_\varepsilon - \overline{n}\|_{L^2(\Omega)}\|y(\Psi(n_\varepsilon) - \Psi(n))\|_{L^2(\Omega)} \\ &\quad + \|y_\varepsilon^{1/2}\nabla\Psi(n_\varepsilon)\|_{L^2(\Omega)}\bigg(\int_\Omega y_\varepsilon(\Psi(n_\varepsilon) - \Psi(n))^2|\nabla y|^2dx\bigg).\end{aligned}$$

By (33), we have $\|y(\Psi(n_\varepsilon) - \Psi(n))\|_{L^2(\Omega)} \to 0$ and by (32), $y_\varepsilon(\Psi(n_\varepsilon) - \Psi(n))^2 \to 0$ in $\Omega$ for a subsequence. Then, by dominated convergence, $\int_\Omega y_\varepsilon(\Psi(n_\varepsilon) - \Psi(n))^2|\nabla y|^2dx \to 0$. We have proved that

$$\int_\Omega y_\varepsilon y^{1/2}\nabla\Psi(n_\varepsilon)\cdot\big(y^{1/2}\nabla\Psi(n_\varepsilon) - \xi\big)dx \to 0. \tag{37}$$

Subtracting (36) from (37), we conclude that

$$\int_\Omega y_\varepsilon|y^{1/2}\nabla\Psi(n_\varepsilon) - \xi|^2dx \to 0.$$

Taking into account this convergence and (34), it follows again by the dominated convergence theorem that

$$\int_\Omega yy_\varepsilon|\nabla\Psi(n_\varepsilon)|^2dx = \int_\Omega y_\varepsilon|y^{1/2}\nabla\Psi(n_\varepsilon)-\xi|^2dx \\ + 2\int_\Omega y_\varepsilon\big(y^{1/2}\nabla\Psi(n_\varepsilon)-\xi\big)\cdot\xi dx + \int_\Omega y_\varepsilon|\xi|^2dx \to \int_\Omega y|\xi|^2dx.$$

This shows the first part of the proposition.

*Step 6.* It remains to show that the limit $n$ solves (27). Let $\phi\in W^{1,\infty}(\Omega)$. Since (a subsequence of) $(n_\varepsilon)$ converges weakly* to $n$ in $L^\infty(\Omega)$, we have

$$\frac{1}{\triangle t}\int_\Omega(n_\varepsilon-\overline{n})\phi dx \to \frac{1}{\triangle t}\int_\Omega(n-\overline{n})\phi dx.$$

Furthermore,

$$\int_\Omega y_\varepsilon\nabla\Psi(n_\varepsilon)\cdot\nabla\phi dx = \int_\Omega y^{1/2}y_\varepsilon^{1/2}\nabla\Psi(n_\varepsilon)\cdot\nabla\phi dx \\ + \int_\Omega(y^{1/2}-y_\varepsilon^{1/2})y_\varepsilon^{1/2}\nabla\Psi(n_\varepsilon)\cdot\nabla\phi dx.$$

By Step 5, the first integral converges to

$$\int_\Omega\big(\nabla(y\Psi(n))-\Psi(n)\nabla y\big)\cdot\nabla\phi dx,$$

while the second integral converges to zero since

$$\left|\int_\Omega(y^{1/2}-y_\varepsilon^{1/2})y_\varepsilon^{1/2}\nabla\Psi(n_\varepsilon)\cdot\nabla\phi dx\right| \\ \le \|y_\varepsilon^{1/2}\nabla\Psi(n_\varepsilon)\|_{L^2(\Omega)}\|y^{1/2}-y_\varepsilon^{1/2}\|_{L^2(\Omega)}\|\nabla\phi\|_{L^\infty(\Omega)}\to 0.$$

We conclude that (27) holds in the weak sense for test functions in $W^{1,\infty}(\Omega)$ but a density argument shows that it is sufficient to take test functions in $H^1(\Omega)$. This finishes the proof. ☐

## 5. Existence of solutions to the high-temperature model

We prove the existence of weak solutions to (26) in $\mathbb{T}^d$. We recall the definition of the total ("reverted") energy

$$W^k_{\rm tot} = W^k - \frac{U}{2}(n^k)^2 \tag{38}$$

and introduce the total variance

$$V^k := \int_{\mathbb{T}^d}\left((W^k)^2-\int_{\mathbb{T}^d}W^kdz\right)^2dx + U\int_{\mathbb{T}^d}W^{k-1}dx\int_{\mathbb{T}^d}\left(n^k-\int_{\mathbb{T}^d}n^kdz\right)^2dx. \tag{39}$$

The main result is as follows.

**Theorem 9** (Existence of weak solutions). *Let $\triangle t > 0$, $U > 0$, $\eta \in (0,1]$, $0 < \delta < 1/(1+\eta)$ and let*

$$n^{k-1},\ W^{k-1} \in L^\infty(\mathbb{T}^d), \quad \delta \le n^{k-1} \le \frac{1-\delta}{\eta}, \quad W^{k-1} \ge 0 \text{ in } \mathbb{T}^d.$$

*Then there exists a weak solution $(n^k, W^k)$ to (9)-(10) in the following sense: It holds $\delta \le n^k \le \|n^{k-1}\|_{L^\infty(\mathbb{T}^d)} \le (1-\delta)/\eta$, $0 \le W^k \le \|W^{k-1}\|_{L^\infty(\mathbb{T}^d)}$ in $\mathbb{T}^d$, $W^k n^k$, $W^k \in H^1(\mathbb{T}^d)$, as well as*

$$\frac{1}{\triangle t}\int_{\mathbb{T}^d}(n^k - n^{k-1})\phi_0 dx = -\int_\Omega \frac{\nabla(W^k n^k) - n^k \nabla W^k}{g(n^k)} \cdot \nabla\phi_0 dx, \tag{40}$$

$$\begin{aligned}\frac{1}{\triangle t}\int_{\mathbb{T}^d}(W^k - W^{k-1})W^k\phi_1 dx = &-\frac{2d-1}{2d}\int_{\mathbb{T}^d}\frac{\nabla W^k \cdot \nabla(W^k\phi_1)}{g(n^k)}dx \\ &- U\int_{\mathbb{T}^d}\frac{|\nabla(W^k n^k) - n^k\nabla W^k|^2}{g(n^k)}\phi_1 dx\end{aligned} \tag{41}$$

*for all $\phi_0 \in H^1(\mathbb{T}^d)$ and $\phi_1 \in H^1(\mathbb{T}^d) \cap L^\infty(\mathbb{T}^d)$, where $g(n^k) = n^k(1-\eta n^k)$. For this solution, the following monotonicity properties hold:*

$$\int_{\mathbb{T}^d} W^k_{\rm tot} dx \ge \int_{\mathbb{T}^d} W^{k-1}_{\rm tot} dx, \quad V^k + \triangle t\frac{2d-1}{d}\int_{\mathbb{T}^d}\frac{|\nabla W^k|^2}{g(n^k)}dx \le V^{k-1}, \tag{42}$$

*where $W^k_{\rm tot}$ and $V^k$ are defined in (38) and (39), respectively. Moreover, if*

$$\frac{U}{2}\int_{\mathbb{T}^d}\left(n^{k-1} - \int_{\mathbb{T}^d} n^{k-1}dz\right)^2 dx < \int_{\mathbb{T}^d} W^{k-1}dx \tag{43}$$

*holds then $W^k \not\equiv 0$.*

**Remark 10** (Comments). 1. The existence result holds true for more general functions $g(n)$ under the assumption that $g(n)$ is strictly positive for $\delta \le n \le (1-\delta)/\eta$.

2. One may interpret $W^k$ as a "renormalized" solution since we need test functions of the form $W^k\phi_1$ in order to avoid vacuum sets $W^k = 0$. Such an idea has been used, for instance, for the compressible quantum Navier-Stokes equations to avoid vacuum sets in the particle density [16]. Test functions of the type $W^k\phi$ allow for the trivial solution $n^k = n^{k-1}$ and $W^k = 0$ but assumption (43) excludes this situation. It means that no constant steady state with $W^k = 0$ exists if the variance of $n^{k-1}$ is small compared to the energy $\int_{\mathbb{T}^d} W^{k-1}dx$.

3. The second inequality in (42) involves $W^{k-2}$ which makes sense when the equations are solved iteratively, starting from $k = 1$. Also (43) can be iterated. Indeed, if (43) holds for $(n^{k-1}, W^{k-1})$, the monotonicity property (42) and mass conservation $\int_{\mathbb{T}^d} n^k dx = \int_{\mathbb{T}^d} n^{k-1}dx$ imply that

$$\begin{aligned}\frac{U}{2}\int_{\mathbb{T}^d}\left(n^k - \int_{\mathbb{T}^d} n^k dz\right)^2 &= \frac{U}{2}\int_{\mathbb{T}^d}(n^k)^2 dx - \frac{U}{2}\left(\int_{\mathbb{T}^d} n^k dx\right)^2 \\ &\le \int_{\mathbb{T}^d}(W^k - W^{k-1})dx + \frac{U}{2}\int_{\mathbb{T}^d}(n^{k-1})^2 dx - \frac{U}{2}\left(\int_{\mathbb{T}^d} n^{k-1}dx\right)^2 < \int_{\mathbb{T}^d} W^k dx.\end{aligned}$$

4. We are not able to perform the limit $\triangle t \to 0$. The reason is that we cannot perform the limit in the quadratic gradient term $|\nabla(W^k n^k) - n^k \nabla W^k|^2$, since we cannot prove the strong convergence of $\nabla(W^k n^k) - n^k \nabla W^k$. Proposition 8 provides such a result for the time-discrete elliptic case. The key step is to show that

$$\int_\Omega \frac{\nabla(W^k n^k) - n^k \nabla W^k}{g(n^k)} \cdot \nabla(Wn - W^k n^k)dx$$
$$= \frac{1}{\triangle t}\langle n^k - n^{k-1}, W^k n^k - Wn \rangle \to 0,$$

where $n$, $W$ are the (weak) limits of $(n^k)$, $(W^k)$, respectively, and $\langle \cdot, \cdot \rangle$ is the dual product between $H^1(\mathbb{T}^d)'$ and $H^1(\mathbb{T}^d)$. It is possible to show that $\triangle t^{-1}(n^k - n^{k-1})$ is bounded in $H^1(\mathbb{T}^d)'$, but the limit $W^k n^k - Wn \to 0$ strongly in $H^1(\mathbb{T}^d)$ (more precisely: the limit of the piecewise constant in time construction of $W^k n^k$ in $L^2(0,T;H^1(\mathbb{T}^d))$) cannot be expected. □

In the one-dimensional case and under the smallness condition (44) below, we can show that $W^k$ is positive, which allows us to define the weak solution to (9)-(10) in the standard sense (with test functions $\phi_1$ instead of $W^k\phi_1$). We set

$$\overline{W^{k-1}} = \int_\mathbb{T} W^{k-1}dx, \quad \overline{n^{k-1}} = \int_\mathbb{T} n^{k-1}dx.$$

**Theorem 11** (One-dimensional case)**.** *Let the assumptions of Theorem 9 hold, let $d = 1$, $G = \max_{\delta \le s \le \|n^{k-1}\|_{L^\infty(\mathbb{T})}} g(s)$, and let $(n^k, W^k)$ for $k \ge 0$ be the solution given by Theorem 9. We assume that*

$$\frac{G}{\triangle t}\big\|W^{k-1} - \overline{W^{k-1}}\big\|^2_{L^2(\mathbb{T})} + U\left(\frac{G\overline{W^{k-2}}}{\triangle t} + \frac{1}{2}\right)\big\|n^{k-1} - \overline{n^{k-1}}\big\|^2_{L^2(\mathbb{T})} < \overline{W^{k-1}}, \tag{44}$$

*Then $W^k$ is strictly positive, $n^k \in H^1(\mathbb{T})$, and (9)-(10) hold in the sense of $H^1(\mathbb{T})'$.*

We proceed to the proof of Theorems 9 and 11. In this section, $\varepsilon$ denotes a positive parameter and not the band energy. Since we are not concerned with the kinetic equations, no notational confusion will occur. Let $\triangle t > 0$, $\alpha$, $\gamma$, $\delta$, $\varepsilon > 0$ satisfying $\gamma < 1$ and $\delta < 1/(1+\eta)$. Define the truncations

$$[n]_\delta = \max\big\{\delta, \min\{(1-\delta)/\eta, n\}\big\}, \quad [W]_\gamma = \max\big\{0, \min\{1/\gamma, W\}\big\},$$

and $g_\delta(s) = [s]_\delta(1 - \eta[s]_\delta)$ for $s \in \mathbb{R}$. Then $g_\delta$ is continuous and strictly positive. Given $W^{k-1}$, $n^{k-1} \in L^\infty(\mathbb{T}^d)$ satisfying $\delta \le n^{k-1} \le (1-\delta)/\eta$, we solve the regularized and truncated nonlinear problem in $\mathbb{T}^d$

$$\frac{1}{\triangle t}(n^k - n^{k-1}) = \operatorname{div}\left(\frac{[W^k]_\gamma + \varepsilon}{g_\delta(n^k)}\nabla n^k\right), \tag{45}$$

$$\frac{1}{\triangle t}(W^k - W^{k-1}) = \frac{2d-1}{2d}\operatorname{div}\left(\frac{\nabla W^k}{g_\delta(n^k)}\right) - U\frac{[W^k]_\gamma}{g_\delta(n^k)}\frac{|\nabla n^k|^2}{1 + \alpha|\nabla n^k|^2}. \tag{46}$$

**Remark 12.** Let us explain the approximation (45)-(46). The truncation of $W^k$ with parameter $\gamma$ ensures that the coefficients are bounded, while the truncation $g_\delta(n^k)$ with parameter $\delta > 0$ guarantees that the denominator is always positive. The regularization parameter $\varepsilon$ gives strict ellipticity for (45), since generally the first term on the right-hand side of (45) without $\varepsilon$ is degenerate. Finally, the approximation of the quadratic gradient term with parameter $\alpha$ avoids regularity issues since it holds $|\nabla n^k|^2 \in L^1(\mathbb{T}^d)$ only. □

5.1. **Solution of an approximated problem.** First, we prove the existence of solutions to (45)-(46).

**Lemma 13** (Existence for the approximated problem)**.** *There exists a weak solution* $(n^k, W^k) \in H^1(\mathbb{T}^d)^2$ *to* (45)-(46).

*Proof.* We define the fixed-point operator $S : L^2(\mathbb{T}^d)^2 \times [0,1] \to L^2(\mathbb{T}^d)^2$ by $S(n^*, W^*; \theta) = (n, W)$, where $(n, W) \in H^1(\mathbb{T}^d)^2$ is the unique solution to the linear problem

$$a_0(n, \phi_0) = F_0(\phi_0), \quad a_1(W, \phi_1) = F_1(\phi_1) \quad \text{for all } \phi_0,\ \phi_1 \in H^1(\mathbb{T}^d), \tag{47}$$

where

$$\begin{aligned} a_0(n, \phi_0) &= \int_{\mathbb{T}^d} \frac{[W^*]_\gamma + \varepsilon}{g_\delta(n^*)} s\nabla n \cdot \nabla\phi_0 dx + \frac{1}{\triangle t}\int_{\mathbb{T}^d} n\phi_0 dx, \\ F_0(\phi_0) &= \frac{\theta}{\triangle t}\int_{\mathbb{T}^d} n^{k-1}\phi_0 dx, \\ a_1(W, \phi_1) &= \frac{2d-1}{2d}\int_{\mathbb{T}^d} \frac{\nabla W \cdot \nabla \phi_1}{g_\delta(n^*)} dx + \frac{1}{\triangle t}\int_{\mathbb{T}^d} W\phi_1 dx, \\ F_1(\phi_1) &= \frac{\theta}{\triangle t}\int_{\mathbb{T}^d} W^{k-1}\phi_1 dx - \theta U \int_{\mathbb{T}^d} \frac{[W^*]_\gamma}{g_\delta(n^*)} \frac{|\nabla n|^2}{1+\alpha|\nabla n|^2}\phi_1 dx. \end{aligned}$$

The approximation and truncation ensure that these forms are bounded on $H^1(\mathbb{T}^d)$. The bilinear forms $a_0$ and $a_1$ are coercive. By the Lax-Milgram lemma, there exists a unique solution $(n, W) \in H^1(\mathbb{T}^d)^2$ to (47). Thus, the fixed-point operator is well defined (and has compact range). Furthermore, $S(n^*, W^*; 0) = 0$. Standard arguments show that $S$ is continuous. Let $(n, W)$ be a fixed point of $S(\cdot, \cdot; \theta)$, i.e., $(n, W)$ solves (45)-(46) with $(n^k, W^k)$ replaced by $(n, W)$. With the test functions $\phi_0 = n$ and $\phi_1 = W$ and the inequality

$$\left(\frac{1}{\triangle t} n - \frac{\theta}{\triangle t} n^{k-1}\right) n \geq \frac{1}{2\triangle t}\big(n^2 - (n^{k-1})^2\big),$$

we find that

$$\begin{aligned} \frac{1}{2\triangle t}\int_{\mathbb{T}^d} n^2 dx + \int_{\mathbb{T}^d} \frac{[W]_\gamma + \varepsilon}{g_\delta(n)}|\nabla n|^2 dx &\leq \frac{1}{2\triangle t}\int_{\mathbb{T}^d}(n^{k-1})^2 dx, \\ \frac{1}{2\triangle t}\int_{\mathbb{T}^d} W^2 dx + \frac{2d-1}{2d}\int_{\mathbb{T}^d}\frac{|\nabla W|^2}{g_\delta(n)} dx &\leq \frac{1}{2\triangle t}\int_{\mathbb{T}^d}(W^{k-1})^2 dx \\ &\quad - \theta U \int_{\mathbb{T}^d} \frac{[W]_\gamma W}{g_\delta(n)} \frac{|\nabla n|^2}{1+\alpha|\nabla n|^2} dx. \end{aligned}$$

The last integral is nonnegative since $[W]_\gamma W \geq 0$. Therefore,

$$\|n\|_{H^1(\mathbb{T}^d)} \leq C(\varepsilon), \quad \|W\|_{H^1(\mathbb{T}^d)} \leq C(\delta),$$

where $C(\varepsilon)$ and $C(\delta)$ are positive constants independent of $(n, W)$. This provides the necessary uniform bound for all fixed points of $S(\cdot,\cdot;\theta)$. We can apply the Leray-Schauder fixed-point theorem to infer the existence of a fixed point for $S(\cdot,\cdot;1)$, i.e. of a weak solution to (45)-(46). □

5.2. **Removing the truncation.** The following maximum principle holds.

**Lemma 14** (Maximum principle)**.** *Let $(n^k, W^k)$ be a weak solution to (45)-(46). Then*

$$\delta \leq n^k \leq \|n^{k-1}\|_{L^\infty(\mathbb{T}^d)} \leq \frac{1-\delta}{\eta}, \quad 0 \leq W^k \leq \frac{1}{\gamma} \quad \text{in } \mathbb{T}^d,$$

*where $\gamma \leq 1/\|W^{k-1}\|_{L^\infty(\mathbb{T}^d)}$.*

*Proof.* We choose $(n^k-\delta)_- = \min\{0, n^k-\delta\}$ as a test function in (45):

$$\frac{1}{\triangle t}\int_{\mathbb{T}^d}\big((n^k-\delta)-(n^{k-1}-\delta)\big)(n^k-\delta)_-dx \\ + \int_{\mathbb{T}^d}\frac{[W^k]_\gamma+\varepsilon}{g_\delta(n^k)}\nabla n^k\cdot\nabla(n^k-\delta)_-dx = 0.$$

Since $-(n^{k-1}-\delta)(n^k-\delta)_- \geq 0$, this gives

$$\frac{1}{\triangle t}\int_{\mathbb{T}^d}(n^k-\delta)_-^2dx \leq -\int_{\mathbb{T}^d}\frac{[W^k]_\gamma+\varepsilon}{g_\delta(n^k)}|\nabla(n^k-\delta)_-|^2dx \leq 0,$$

and hence, $n^k \geq \delta$ in $\mathbb{T}^d$. In a similar way, the test function $(n^k-N)_+ = \max\{0, n^k-N\}$ with $N := \|n^{k-1}\|_{L^\infty(\mathbb{T}^d)}$ leads to $n^k - N \leq 0$ in $\mathbb{T}^d$. Next, we use $W_-^k \leq 0$ as a test function in (46):

$$\frac{1}{\triangle t}\int_{\mathbb{T}^d}(W_-^k)^2dx + \frac{2d-1}{2d}\int_{\mathbb{T}^d}\frac{|\nabla W_-|^2}{g_\delta(n^k)}dx \\ = \frac{1}{\triangle t}\int_{\mathbb{T}^d}W^{k-1}W_-^kdx - U\int_{\mathbb{T}^d}\frac{[W^k]_\gamma W_-^k}{g_\delta(n^k)}\frac{|\nabla n^k|^2}{1+\alpha|\nabla n^k|^2}dx \leq 0.$$

We deduce that $W^k \geq 0$. The proof of $W^k \leq \|W^{k-1}\|_{L^\infty(\mathbb{T}^d)} \leq 1/\gamma$ is similar, using the test function $(W^k - \|W^{k-1}\|_{L^\infty(\mathbb{T}^d)})_+$. □

We have shown that $(n^k, W^k)$ solves

$$\frac{1}{\triangle t}(n^k - n^{k-1}) = \operatorname{div}\left(\frac{W^k+\varepsilon}{g(n^k)}\nabla n^k\right), \tag{48}$$

$$\frac{1}{\triangle t}(W^k - W^{k-1}) = \frac{2d-1}{2d}\operatorname{div}\left(\frac{\nabla W^k}{g(n^k)}\right) - U\frac{W^k}{g(n^k)}\frac{|\nabla n^k|^2}{1+\alpha|\nabla n^k|^2}, \tag{49}$$

where $g(n) = n(1-\eta n)$.

5.3. **The limit** $\alpha \to 0$**.** Let $(n^k_\alpha, W^k_\alpha)$ be a weak solution to (48)-(49). We use the test function $n^k_\alpha$ in (48),

$$\frac{1}{2\triangle t}\int_{\mathbb{T}^d}(n^k_\alpha)^2dx + \int_{\mathbb{T}^d}\frac{W^k_\alpha+\varepsilon}{g(n^k_\alpha)}|\nabla n^k_\alpha|^2dx \le \frac{1}{2\triangle t}\int_{\mathbb{T}^d}(n^{k-1})^2dx, \tag{50}$$

and the test function $W^k_\alpha$ in (49),

$$\frac{1}{2\triangle t}\int_{\mathbb{T}^d}(W^k_\alpha)^2dx + \frac{2d-1}{2d}\int_{\mathbb{T}^d}\frac{|\nabla W^k_\alpha|^2}{g(n^k_\alpha)} \le \frac{1}{2\triangle t}\int_{\mathbb{T}^d}(W^{k-1})^2dx, \tag{51}$$

which provides immediately uniform $H^1(\mathbb{T}^d)$ estimates since $g(n^k_\alpha) \ge C(\delta) > 0$:

$$\|n^k_\alpha\|_{H^1(\mathbb{T}^d)} \le C(\delta,\varepsilon,\triangle t), \quad \|W^k_\alpha\|_{H^1(\mathbb{T}^d)} \le C(\delta,\triangle t),$$

where the constants are independent of $\alpha$. By compactness, this implies the existence of a subsequence which is not relabeled such that, as $\alpha \to 0$,

$$\begin{aligned} &n^k_\alpha \to n^k,\ W^k_\alpha \to W^k \quad \text{strongly in } L^2(\mathbb{T}^d), \\ &n^k_\alpha \rightharpoonup n^k,\ W^k_\alpha \rightharpoonup W^k \quad \text{weakly in } H^1(\mathbb{T}^d). \end{aligned}$$

This shows that, maybe for a subsequence, $W^k_\alpha/g(n^k_\alpha) \to W^k/g(n^k)$ and $1/g(n^k_\alpha) \to 1/g(n^k)$ a.e. in $\mathbb{T}^d$, and by dominated convergence, strongly in $L^2(\mathbb{T}^d)$.

We claim that $n^k_\alpha \to n^k$ strongly in $H^1(\mathbb{T}^d)$. Let $y_\alpha := (W^k_\alpha+\varepsilon)/g(n^k_\alpha)$. Then $y_\varepsilon \ge \varepsilon/\sup_{s\in[\delta,N]} g(s) > 0$, where $N = \|n^{k-1}\|_{L^\infty(\mathbb{T}^d)}$, and $y_\alpha \to y := W^k/g(n^k) \ge 0$ strongly in $L^2(\Omega)$. Thus, $y_\alpha^{1/2}\nabla n^k_\alpha \rightharpoonup y^{1/2}\nabla n^k$ weakly in $L^2(\Omega)$, and it follows that

$$\int_{\mathbb{T}^d} y_\alpha \nabla n^k \cdot \nabla(n^k_\alpha - n^k)dx \to 0.$$

Taking $n^k_\alpha - n^k$ as a test function in (48), we obtain

$$\int_{\mathbb{T}^d} y_\alpha \nabla n^k_\alpha \cdot \nabla(n^k_\alpha - n^k)dx = -\frac{1}{\triangle t}\int_\Omega (n^k_\alpha - n^{k-1})(n^k_\alpha - n^k)dx \to 0.$$

Subtraction of these integrals leads to

$$\int_{\mathbb{T}^d} y_\alpha |\nabla(n^k_\alpha - n^k)|^2dx \to 0.$$

Since $y_\alpha \ge \varepsilon/\sup_{s\in[\delta,N]} g(s) > 0$, this proves the claim. In particular, $1/(1+\alpha|\nabla n^k_\alpha|^2) \to 1$ in $L^2(\mathbb{T}^d)$. From this, we can directly deduce that

$$\frac{|\nabla n^k_\alpha|^2}{1+\alpha|\nabla n^k_\alpha|^2} \to |\nabla n^k|^2 \quad \text{in } L^1(\mathbb{T}^d).$$

The above convergence results are sufficient to pass to the limit $\alpha \to 0$ in (48)-(49), showing that $(n^k, W^k)$ solves

$$\frac{1}{\triangle t}(n^k - n^{k-1}) = \operatorname{div}\left(\frac{W^k+\varepsilon}{g(n^k)}\nabla n^k\right), \tag{52}$$

$$\frac{1}{\triangle t}(W^k - W^{k-1}) = \frac{2d-1}{2d}\operatorname{div}\left(\frac{\nabla W^k}{g(n^k)}\right) - U\frac{W^k}{g(n^k)}|\nabla n^k|^2, \tag{53}$$

5.4. **The limit $\varepsilon \to 0$.** This limit is the delicate part of the proof. We first state a lemma concerning weak and strong convergence.

**Lemma 15.** *Let $(f_n)$ be a weakly and $(g_n)$ be a strongly converging sequence in $L^2(\Omega)$ which have the same limit. If $|f_n(x)| \le |g_n(x)|$ for all $n \in \mathbb{N}$ and a.e. $x \in \mathbb{T}^d$, then $(f_n)$ converges strongly in $L^2(\Omega)$.*

*Proof.* Let $f$ denote the weak limit of $(f_n)$ and $(g_n)$. Due to the weak lower semi-continuity of the norm,

$$\begin{aligned}\int_{\mathbb{T}^d}|f(x)|^2dx &\le \liminf_{n\to\infty}\int_{\mathbb{T}^d}|f_n(x)|^2dx \\ &\le \limsup_{n\to\infty}\int_{\mathbb{T}^d}|f_n(x)|^2dx \le \lim_{n\to\infty}\int_{\mathbb{T}^d}|g_n(x)|^2dx = \int_{\mathbb{T}^d}|f(x)|^2dx.\end{aligned}$$

Thus, the limes inferior and superior coincide and $\|f_n\|_{L^2(\Omega)} \to \|f\|_{L^2(\Omega)}$. Together with the weak convergence of $(f_n)$, we deduce the strong convergence. □

Let $(n_\varepsilon^k, W_\varepsilon^k)$ be a weak solution to (52)-(53). Inequalities (50) and (51) show the following bounds uniform in $\varepsilon$:

$$\|n_\varepsilon^k\|_{L^\infty(\mathbb{T}^d)} + \|(W_\varepsilon^k+\varepsilon)^{1/2}\nabla n_\varepsilon^k\|_{L^2(\mathbb{T}^d)} \le C(\delta, \triangle t), \tag{54}$$

$$\|W_\varepsilon^k\|_{L^\infty(\Omega)} + \|W_\varepsilon^k\|_{H^1(\mathbb{T}^d)} \le C(\delta, \triangle t). \tag{55}$$

By compactness, there exists a subsequence (not relabeled) such that, as $\varepsilon \to 0$,

$$n_\varepsilon^k \rightharpoonup^* n^k \quad \text{weakly* in } L^\infty(\mathbb{T}^d), \tag{56}$$

$$W_\varepsilon^k \to W^k \quad \text{strongly in } L^2(\mathbb{T}^d), \quad W_\varepsilon^k \rightharpoonup W^k \quad \text{weakly in } H^1(\mathbb{T}^d). \tag{57}$$

Again, we need strong convergence for $\nabla n_\varepsilon^k$. Since equation (52) is degenerate, we obtain a weaker result. For this, let $\Psi \in C^2(\mathbb{R})$ be strictly monotonically increasing and satisfy $\Psi'(t) = 1/g(t)$ for $\delta \le t \le (1-\delta)/\eta$. Thus, we can apply Proposition 8 for $y_\varepsilon = W_\varepsilon^k + \varepsilon$ and $y = W^k$ to conclude that, up to a subsequence,

$$\begin{aligned}&(W^k)^{1/2}(W_\varepsilon^k+\varepsilon)^{1/2}\nabla\Psi(n_\varepsilon^k) \to \nabla(W^k\Psi(n^k)) - \Psi(n^k)\nabla W^k, \\ &1_{\{W^k>0\}}(n_\varepsilon^k - n^k) \to 0 \quad \text{strongly in } L^2(\Omega).\end{aligned} \tag{58}$$

The latter convergence implies that $n_\varepsilon^k - n^k \to 0$ a.e. in $\{W^k > 0\}$ and, by dominated convergence,

$$\frac{1_{\{W^k>0\}}}{g(n_\varepsilon^k)} \to \frac{1_{\{W^k>0\}}}{g(n^k)} \quad \text{strongly in } L^2(\mathbb{T}^d). \tag{59}$$

Now, let $y_\varepsilon = (W_\varepsilon^k+\varepsilon)/g(n_\varepsilon^k)$ and $y = W^k/g(n^k)$. We know that $y_\varepsilon, y \in L^\infty(\mathbb{T}^d)\cap H^1(\mathbb{T}^d)$ and $y_\varepsilon \to y$ strongly in $L^2(\mathbb{T}^d)$. Thus, we can again apply Proposition 8 with $\Psi = \mathrm{Id}$ and infer that $W^kn^k \in H^1(\mathbb{T}^d)$ as well as

$$\frac{1}{\triangle t}\int_{\mathbb{T}^d}(n^k-n^{k-1})\phi_0 dx + \int_\Omega \frac{\nabla(W^kn^k)-n^k\nabla W^k}{g(n^k)}\cdot\nabla\phi_0 dx = 0$$

for all $\phi_0 \in H^1(\mathbb{T}^d)$.

Let $\phi_1$ be a smooth test function. We use the test function $W^k\phi_1$ in the weak formulation of (53):

$$\begin{aligned}0 &= \frac{1}{\triangle t}\int_{\mathbb{T}^d}(W_\varepsilon^k - W^{k-1})W^k\phi_1 dx + \frac{2d-1}{2d}\int_{\mathbb{T}^d}\frac{\nabla W_\varepsilon^k\cdot\nabla(W^k\phi_1)}{g(n_\varepsilon^k)}dx \\ &\quad + U\int_{\mathbb{T}^d} W^k\frac{W_\varepsilon^k}{g(n_\varepsilon^k)}|\nabla n_\varepsilon^k|^2\phi_1 dx =: I_\varepsilon^1 + I_\varepsilon^2 + I_\varepsilon^3.\end{aligned} \tag{60}$$

We pass to the limit $\varepsilon\to 0$ term by term. By (57),

$$I_\varepsilon^1 \to \frac{1}{\triangle t}\int_{\mathbb{T}^d}(W^k - W^{k-1})W^k\phi_1 dx.$$

For the integral $I_\varepsilon^2$, we use the strong convergence (59) and the weak convergence of $(\nabla W_\varepsilon^k)$ in $L^2(\mathbb{T}^d)$ to infer that

$$\begin{aligned}I_\varepsilon^2 &= \int_{\mathbb{T}^d}\frac{1_{\{W^k>0\}}}{g(n_\varepsilon^k)}W^k\nabla W_\varepsilon^k\cdot\nabla\phi_1 dx + \int_{\mathbb{T}^d}\frac{1_{\{W^k>0\}}}{g(n_\varepsilon^k)}\nabla W_\varepsilon^k\cdot\nabla W^k\phi_1 dx \\ &\to \int_{\mathbb{T}^d}\frac{1_{\{W^k>0\}}}{g(n^k)}W^k\nabla W^k\cdot\nabla\phi_1 dx + \int_{\mathbb{T}^d}\frac{1_{\{W^k>0\}}}{g(n^k)}|\nabla W^k|^2\phi_1 dx \\ &= \int_{\mathbb{T}^d}\frac{\nabla W^k\cdot\nabla(W^k\phi_1)}{g(n^k)}dx.\end{aligned}$$

The remaining integral $I_\varepsilon^3$ requires some work. As a preparation, using Proposition 8, we infer similarly to (58) that

$$h_\varepsilon := (W^k)^{1/2}\left(\frac{W_\varepsilon^k+\varepsilon}{g(n_\varepsilon^k)}\right)^{1/2}\nabla n_\varepsilon^k \to \frac{\nabla(W^kn^k)-n^k\nabla W^k}{g(n^k)^{1/2}} =: h$$

strongly in $L^2(\mathbb{T}^d)$. Let

$$\xi_\varepsilon := \left(\frac{W_\varepsilon^k}{g(n_\varepsilon^k)}\right)^{1/2}\nabla n_\varepsilon^k.$$

Then $(\xi_\varepsilon)$ is bounded in $L^2(\Omega)$ and admits a weakly convergent subsequence, i.e. $\xi_\varepsilon \rightharpoonup \xi$ for some $\xi\in L^2(\mathbb{T}^d)$. Similarly as in the proof of Proposition 8, i.e. with an argument as in (35), we can find that, up to a subsequence, $(W_\varepsilon^k+\varepsilon)^{1/2}\xi_\varepsilon \rightharpoonup h$ weakly in $L^2(\mathbb{T}^d)$ implying $(W^k)^{1/2}\xi = h$. In particular,

$$f_\varepsilon := (W^k)^{1/2}\xi_\varepsilon \rightharpoonup (W^k)^{1/2}\xi = h \quad \text{weakly in } L^2(\mathbb{T}^d).$$

Since $|f_\varepsilon(x)| \le |h_\varepsilon(x)|$ for a.e. $x \in \mathbb{T}^d$ and all $\varepsilon > 0$, we can apply Lemma 15 and obtain that, up to a subsequence, $f_\varepsilon$ converges strongly in $L^2(\Omega)$. Thus,

$$I_\varepsilon^3 = U \int_{\mathbb{T}^d} f_\varepsilon^2 \phi_1 dx \to U \int_{\mathbb{T}^d} \frac{|\nabla(W^k n^k) - n^k \nabla W^k|^2}{g(n^k)} \phi_1 dx.$$

Hence, passing to the limit $\varepsilon \to 0$ in (60), we infer that $(n^k, W^k)$ solves (9)-(10).

5.5. **Energy estimate.** We claim that the total energy $\int_{\mathbb{T}^d} W_{\mathrm{tot}}^k dx$ is nondecreasing in $k$. Let $(n_\varepsilon^k, W_\varepsilon^k)$ be a weak solution to (52)-(53). Then

$$\begin{aligned} \triangle t \int_{\mathbb{T}^d} \Big( W_\varepsilon^k - W^{k-1} - \frac{U}{2}\big((n_\varepsilon^k)^2 - (n^{k-1})^2\big)\Big) dx \\ \ge \triangle t \int_{\mathbb{T}^d} \big( W_\varepsilon^k - W^{k-1} - U(n_\varepsilon^k - n^{k-1}) n_\varepsilon^k \big) dx. \end{aligned}$$

Taking the test functions $\phi_1 = U$ in (53) and $\phi_0 = n_\varepsilon^k$ in (52) and subtracting both equations, the above integral becomes

$$\begin{aligned} \triangle t \int_{\mathbb{T}^d} \Big( W_\varepsilon^k - W^{k-1} - \frac{U}{2}\big((n_\varepsilon^k)^2 - (n^{k-1})^2\big)\Big) dx \\ \ge -U \int_{\mathbb{T}^d} \frac{W_\varepsilon^k}{g(n_\varepsilon^k)} |\nabla n_\varepsilon^k|^2 dx + U \int_{\mathbb{T}^d} \frac{W_\varepsilon^k + \varepsilon}{g(n_\varepsilon^k)} |\nabla n_\varepsilon^k|^2 dx \ge 0. \end{aligned}$$

Thus, with the lower semi-continuity of the norm, we have

$$\begin{aligned} \int_{\mathbb{T}^d} \Big( W^{k-1} - \frac{U}{2}(n^{k-1})^2 \Big) dx \le \liminf_{\varepsilon \to 0} \int_{\mathbb{T}^d} \Big( W_\varepsilon^k - \frac{U}{2}(n_\varepsilon^k)^2 \Big) dx \\ \le \int_{\mathbb{T}^d} \Big( W^k - \frac{U}{2}(n^k)^2 \Big) dx. \end{aligned}$$

In view of mass conservation $\int_{\mathbb{T}^d} n^k dx = \int_{\mathbb{T}^d} n^{k-1} dx$, it follows by Jensen's inequality that

$$\begin{aligned} \int_{\mathbb{T}^d} W^k dx &\ge \int_{\mathbb{T}^d} W^{k-1} dx + \frac{U}{2} \int_{\mathbb{T}^d} \big((n^k)^2 - (n^{k-1})^2\big) dx \\ &\ge \int_{\mathbb{T}^d} \Big( W^{k-1} - \frac{U}{2}(n^{k-1})^2 \Big) dx + \frac{U}{2} \Big( \int_{\mathbb{T}^d} n^{k-1} dx \Big)^2. \end{aligned} \tag{61}$$

This shows the energy inequality in (42). Finally, assumption (43) gives $\int_{\mathbb{T}^d} W^k dx > 0$ and consequently $W^k \not\equiv 0$.

5.6. **An estimate for the variance.** We claim that the total variance

$$V^k := \int_{\mathbb{T}^d} \Big( W^k - \int_{\mathbb{T}^d} W^k dz \Big)^2 dx + U \int_{\mathbb{T}^d} W^{k-1} dx \int_{\mathbb{T}^d} \Big( n^k - \int_{\mathbb{T}^d} n^k dz \Big)^2 dx$$

is nonincreasing in $k$. For the proof, we observe that, taking the test function $\phi_1 = 1$ in the weak formulation of (53) and performing the limit $\varepsilon \to 0$,

$$\int_{\mathbb{T}^d} W^k dx \le \int_{\mathbb{T}^d} W^{k-1} dx. \tag{62}$$

Thus, by the energy estimate (61),

$$\begin{aligned}\frac{U}{2}\int_{\mathbb{T}^d}(n^k)^2dx &\le \frac{U}{2}\int_{\mathbb{T}^d}(n^{k-1})^2dx - \int_{\mathbb{T}^d}(W^{k-1}-W^k)dx \\ &\le \frac{U}{2}\int_{\mathbb{T}^d}(n^{k-1})^2dx. \end{aligned}\tag{63}$$

We employ (61) again to find that

$$\begin{aligned}&\bigg(\int_{\mathbb{T}^d}W^{k-1}dx\bigg)^2 - \bigg(\int_{\mathbb{T}^d}W^k dx\bigg)^2 \\ &\quad = \bigg(\int_{\mathbb{T}^d}W^{k-1}dx + \int_{\mathbb{T}^d}W^k dx\bigg)\bigg(\int_{\mathbb{T}^d}W^{k-1}dx - \int_{\mathbb{T}^d}W^k dx\bigg) \\ &\quad \le \bigg(\int_{\mathbb{T}^d}W^{k-1}dx + \int_{\mathbb{T}^d}W^k dx\bigg)\frac{U}{2}\bigg(\int_{\mathbb{T}^d}(n^{k-1})^2dx - \int_{\mathbb{T}^d}(n^k)^2dx\bigg).\end{aligned}$$

In view of (63), the second bracket on the right-hand side is nonnegative, such that (62) leads to

$$\bigg(\int_{\mathbb{T}^d}W^{k-1}dx\bigg)^2 - \bigg(\int_{\mathbb{T}^d}W^k dx\bigg)^2 \le U\int_{\mathbb{T}^d}W^{k-1}dx\bigg(\int_{\mathbb{T}^d}(n^{k-1})^2dx - \int_{\mathbb{T}^d}(n^k)^2dx\bigg).$$

We take the test function $\phi_1 = 2\triangle t$ in (41):

$$\begin{aligned}0 &\ge 2\int_{\mathbb{T}^d}(W^k - W^{k-1})W^k dx + \triangle t\frac{2d-1}{d}\int_{\mathbb{T}^d}\frac{|\nabla W^k|^2}{g(n^k)}dx \\ &\ge \int_{\mathbb{T}^d}(W^k)^2dx - \int_{\mathbb{T}^d}(W^{k-1})^2dx + \triangle t\frac{2d-1}{d}\int_{\mathbb{T}^d}\frac{|\nabla W^k|^2}{g(n^k)}dx.\end{aligned}$$

Combining the previous two inequalities, we arrive at

$$\begin{aligned}&\int_{\mathbb{T}^d}(W^k)^2dx - \bigg(\int_{\mathbb{T}^d}W^k dx\bigg)^2 + \triangle t\frac{2d-1}{d}\int_{\mathbb{T}^d}\frac{|\nabla W^k|^2}{g(n^k)}dx \\ &\qquad \le \int_{\mathbb{T}^d}(W^{k-1})^2dx - \bigg(\int_{\mathbb{T}^d}W^{k-1}dx\bigg)^2 \\ &\qquad\quad + U\int_{\mathbb{T}^d}W^{k-1}dx\bigg(\int_{\mathbb{T}^d}(n^{k-1})^2dx - \int_{\mathbb{T}^d}(n^k)^2dx\bigg).\end{aligned}$$

Since the measure of $\mathbb{T}^d$ is one, we have

$$\int_{\mathbb{T}^d}(W^k)^2dx - \bigg(\int_{\mathbb{T}^d}W^k dx\bigg)^2 = \int_{\mathbb{T}^d}\bigg((W^k)^2 - \bigg(\int_{\mathbb{T}^d}W^k dz\bigg)\bigg)^2 dx.$$

Thus, taking into account mass conservation $\int_{\mathbb{T}^d} n^k dx = \int_{\mathbb{T}^d} n^{k-1}dx$ and $\int_{\mathbb{T}^d} W^{k-1}dx \leq \int_{\mathbb{T}^d} W^{k-2}dx$ (see (62)),

$$\begin{aligned}\int_{\mathbb{T}^d}\bigg((W^k)^2 - \Big(\int_{\mathbb{T}^d} W^k dz\Big)\bigg)^2 dx + \triangle t\frac{2d-1}{d}\int_{\mathbb{T}^d}\frac{|\nabla W^k|^2}{g(n^k)}dx \\ \leq \int_{\mathbb{T}^d}\bigg((W^{k-1})^2 - \Big(\int_{\mathbb{T}^d} W^{k-1}dz\Big)\bigg)^2 dx \\ + U\int_{\mathbb{T}^d} W^{k-2}dx\int_{\mathbb{T}^d}\bigg((n^{k-1})^2 - \Big(\int_{\mathbb{T}^d} n^{k-1}dz\Big)\bigg)^2 dx \\ - U\int_{\mathbb{T}^d} W^{k-1}dx\int_{\mathbb{T}^d}\bigg((n^k)^2 - \Big(\int_{\mathbb{T}^d} n^k dz\Big)\bigg)^2 dx,\end{aligned}$$

and the claim follows after using the lower-semicontinuity of the $L^2$-norm.

5.7. **Proof of Theorem 11.** Let $d = 1$. The second inequality in (42) implies that

$$\int_{\mathbb{T}^1}|\partial_x W^k|^2 dx \leq \frac{G}{\triangle t}V^{k-1},$$

where $G = \max_{\delta\leq s\leq\{n^{k-1}\|_{L^\infty(\mathbb{T})}\}} g(s) \geq \|g(n^k)\|_{L^\infty(\mathbb{T})}$. By the mean-value theorem, there exists $x_0 \in \mathbb{T}$ such that

$$W^k(x) = W^k(x_0) + \int_{x_0}^x \partial_x W^k(z)dz \geq \int_{\mathbb{T}} W^k(z)dz - \int_{\mathbb{T}}|\partial_x W^k|dz.$$

Then, using Jensen's inequality and the energy estimate in (42),

$$\begin{aligned}W^k(x) &\geq \int_{\mathbb{T}} W^k dx - \int_{\mathbb{T}}|\partial_x W^k|^2 dx \\ &\geq \int_{\mathbb{T}} W^{k-1}dx - \frac{U}{2}\int_{\mathbb{T}}\Big(n^{k-1} - \int_{\mathbb{T}} n^{k-1}dz\Big)^2 dx - \frac{G}{\triangle t}V^{k-1}.\end{aligned}$$

By definition (39) of $V^k$, the right-hand side is positive if

$$\overline{W^{k-1}} > \frac{G}{\triangle t}\int_{\mathbb{T}}(W^{k-1} - \overline{W^{k-1}})^2 dx + U\bigg(\frac{G}{\triangle t}\overline{W^{k-2}} + \frac{1}{2}\bigg)\int_{\mathbb{T}}(n^{k-1} - \overline{n^{k-1}})^2 dx,$$

which is our assumption. Since $W^k \in H^1(\mathbb{T}) \hookrightarrow C^0([0,1])$, we conclude that $W^k > 0$ in $[0,1]$. Then we can use $\phi_1 = \phi/W^k$ as a test function in (41) and obtain the standard weak formulation of (41) for test functions $\phi \in H^1(\mathbb{T})$. Furthermore, for $\phi \in H^1(\mathbb{T})$,

$$\begin{aligned}\int_{\mathbb{T}} n^k\partial_x\phi dx &= \int_{\mathbb{T}} n^k\partial_x\bigg(W^k\partial_x\Big(\frac{\phi}{W^k}\Big) + \partial_x W^k\frac{\phi}{W^k}\bigg)dx \\ &= -\int_{\mathbb{T}}\partial_x(n^kW^k)\frac{\phi}{W^k}dx + \int_{\mathbb{T}} n^k\partial_x W^k\frac{\phi}{W^k}dx,\end{aligned}$$

showing that $n^k \in H^1(\mathbb{T})$ and finishing the proof.

## 6. Numerical simulations

We solve the one-dimensional equations (6) and (8) on the torus in conservative form, i.e. for the variables $n$ and $W_{\rm tot}$. The equations are discretized by the implicit Euler method and solved in a semi-implicit way:

$$\frac{1}{\triangle t}(n^k - n^{k-1}) = \partial_x\left(\frac{W^{k-1}\partial_x n^k}{n^{k-1}(1-\eta n^{k-1})}\right), \tag{64}$$

$$\frac{1}{\triangle t}(W^k_{\rm tot} - W^{k-1}_{\rm tot}) = \partial_x\left(\frac{\partial_x W^k}{2n^k(1-\eta n^k)} + \frac{UW^k}{1-\eta n^k}\partial_x n^k\right), \tag{65}$$

where $W^k_{\rm tot} = W^k - (U/2)(n^k)^2$, $x \in \mathbb{T} = (0,1)$. The spatial derivatives are discretized by centered finite differences with constant space step $\triangle x > 0$. For given $(W^{k-1}, n^{k-1})$, the first equation (64) is solved for $n^k$. This solution is employed in the second equation (65) which is solved for $W^k_{\rm tot}$. Finally, we define $W^k = W^k_{\rm tot} + (U/2)(n^k)^2$. We choose the parameters $U = 10$, $\eta = 1$, $\triangle t = 10^{-5}$, and $\triangle x = 10^{-2}$. The initial energy $W^0$ is constant and the initial density equals

$$n^0(x) = \begin{cases} 3/4 & \text{for } 1/4 \le x \le 3/4, \\ 1/4 & \text{else,} \end{cases} \qquad x \in [0,1].$$

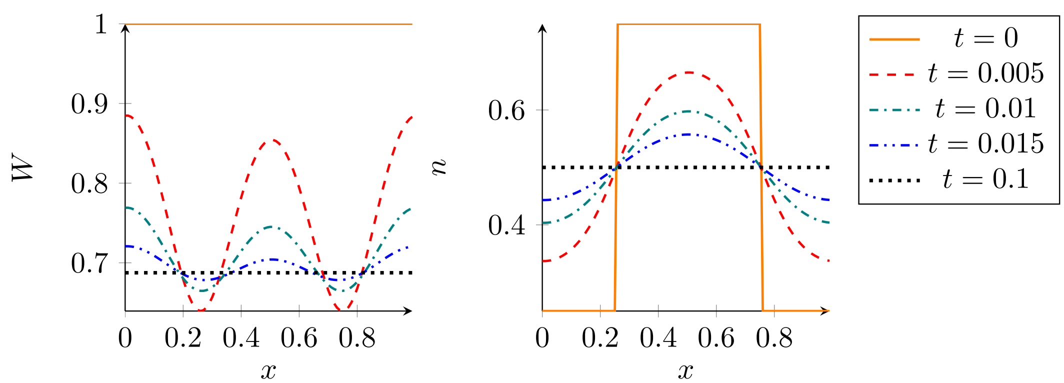


Figure 1. Evolution of $(n, W)$ with initial energy $W^0 = 1$.

The time evolution of the particle density and energy is shown in Figure 1 for initial energy $W^0 = 1$. The variables converge to the constant steady state $(n^\infty, W^\infty)$ as $t \to \infty$, which is almost reached after time $t = 0.1$. Since equations (64)-(65) are conservative, the total particle number $\int_0^1 n(x,t)dx$ and the total energy $\int_0^1 W_{\rm tot}(x,t)dx$ are constant in time. Consequently, the values for the steady state can be computed explicitly. We obtain for $W^0 = 1$,

$$n^\infty = \int_0^1 n^0(x)dx = \frac{1}{2},$$

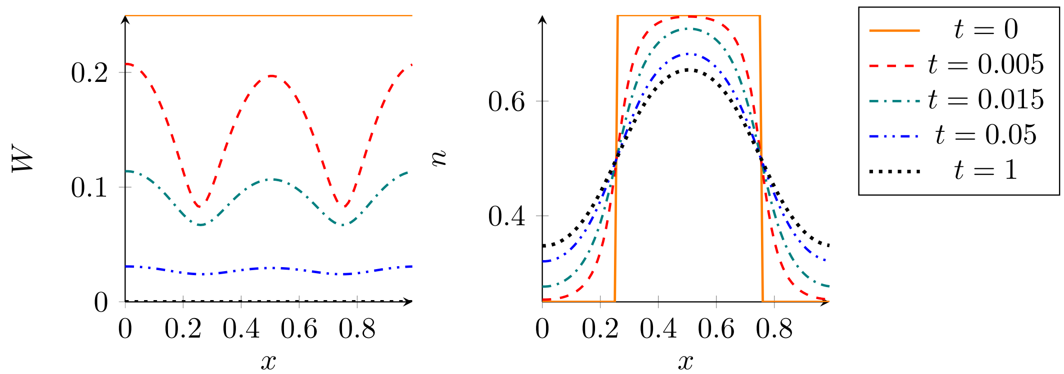


FIGURE 2. Evolution of $(n, W)$ with initial energy $W^0 = 1/4$.

$$W^\infty = W^\infty_{\rm tot} + \frac{U}{2}(n^\infty)^2 = \int_0^1 \left( W^0(x) - \frac{U}{2} n^0(x)^2 \right) dx + \frac{U}{2}(n^\infty)^2 = \frac{11}{16}.$$

The energy stays positive for all times, so the high-temperature equations are strictly parabolic, and the convergence to the (constant) steady state is quite natural.

The situation is different in Figure 2, where the particle density converges to a *nonconstant* steady state $n^\infty$ (we have chosen $W^0 = 1/4$). This can be understood as follows. By contradiction, let both the particle density and energy be converging to a constant steady state. Then $n^\infty = 1/2$ (see the above calculation) and

$$W^\infty = \int_0^1 \left( W^0(x) - \frac{U}{2} n^0(x)^2 \right) dx + \frac{U}{2}(n^\infty)^2 = -\frac{1}{16}.$$

However, this contradicts the fact that the energy $W$ is nonnegative which follows from the maximum principle. Therefore, it is plausible that either $n$ or $W$ cannot converge to a constant. If $n^\infty$ is not constant, $W^\infty \nabla n^\infty$ is constant only if $W^\infty = 0$. Thus, it is reasonable that the energy converges to zero, while $n^\infty$ is not constant. One may say that there is not sufficient initial "reverted" energy to level the particle density.

Another difference between Figure 1 and Figure 2 is the time scale. For larger initial energies, the convergence to equilibrium is faster. In fact, Figure 3 shows that the decay of the $\ell^2$ norm of $n(t) - n^\infty$ and $W(t) - W^\infty$ is exponential. Here, we have chosen the initial particle density $n^0(x) = \frac{1}{4}$ for $0 \le x < \frac{1}{2}$, $n^0(x) = \frac{3}{4}$ for $\frac{1}{2} \le x < 1$, and the initial energy $W^0 \in \{\frac{1}{4}, \frac{1}{2}, 1\}$. For $W^0 \in \{1, \frac{1}{2}\}$, we have $n^\infty = \frac{1}{2}$ and $W^\infty = \max\{0, W^0 - U/32\}$. For $W^0 = \frac{1}{4}$, it holds that $W^\infty = 0$ and we have set $n^\infty(x) = n(x, 2)$.

Finally, we compute the numerical convergence rates for different space and time step sizes $\triangle x$ and $\triangle t$, respectively. Since there is no explicit solution available, we choose as reference solution the solution to (64)-(65) with $\triangle x = 1/1680$ (for the computation of the spatial $\ell^2_x$ error) and $\triangle t = 1/5040$ (for the computation of the $\ell^2_t \ell^2_x$ error). Figure 4 shows that the temporal error is linear in $\triangle t$, and Figure 5 indicates that the spatial error is quadratic in $\triangle x$. These values are expected in view of our finite-difference discretization and they confirm the validity of the numerical scheme.

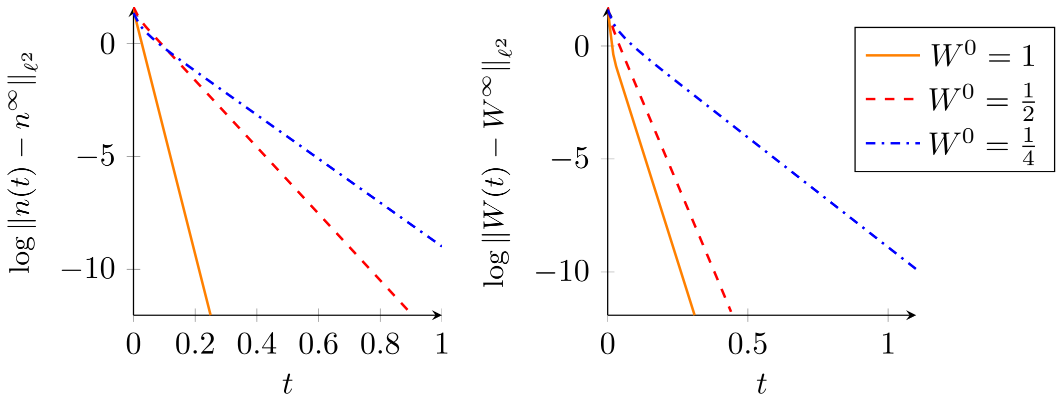


FIGURE 3. Decay rates for various initial energies.

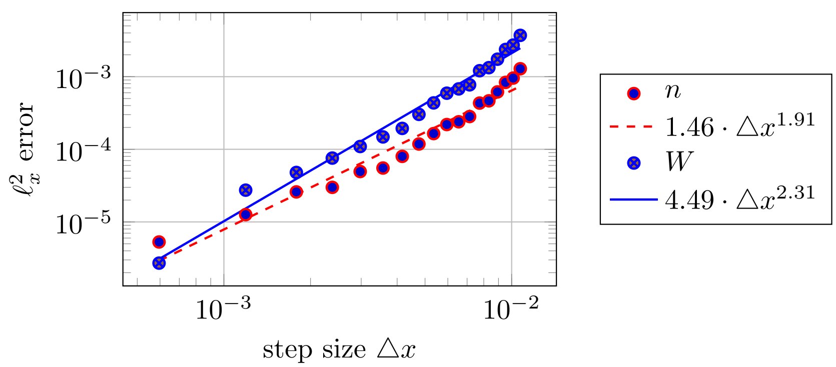


FIGURE 4. Numerical convergence in time.

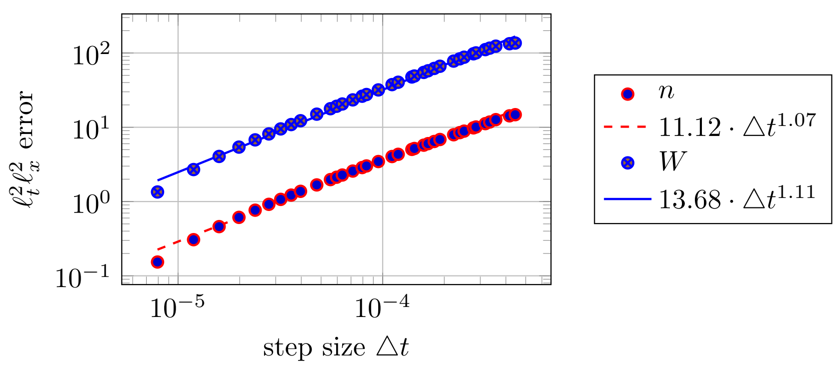


FIGURE 5. Numerical convergence in space.

## APPENDIX A. CALCULATION OF SOME INTEGRALS

We recall that $\varepsilon(p) = -2\varepsilon_0 \sum_{k=1}^d \cos(2\pi p_i)$. Then $u_i(p) = (\partial\varepsilon/\partial p_i)(p) = 4\pi\varepsilon_0 \sin(2\pi p_i)$, and we calculate

$$\int_{\mathbb{T}^2} \varepsilon^2 dp = 4\varepsilon_0^2 \int_0^1 \sum_{k=1}^d \cos^2(2\pi p_k) dp_k = 2d\varepsilon_0^2, \tag{66}$$

$$\int_{\mathbb{T}^d} u_i u_j dp = (4\pi\varepsilon_0)^2\delta_{ij}\int_0^1 \sin^2(2\pi p_i)dp_i = \frac{1}{2}(4\pi\varepsilon_0)^2\delta_{ij}, \tag{67}$$

$$\begin{aligned}\int_{\mathbb{T}^d} \varepsilon u_i u_j dp &= -2\varepsilon_0(4\pi\varepsilon_0)^2\sum_{k=1}^d \int_{\mathbb{T}^d}\cos(2\pi p_k)\sin(2\pi p_i)\sin(2\pi p_j)dp \\ &= -2\varepsilon_0(4\pi\varepsilon_0)^2\delta_{ij}\int_0^1 \cos(2\pi p_i)\sin^2(2\pi p_i)dp_i = 0,\end{aligned} \tag{68}$$

$$\int_{\mathbb{T}^d}\varepsilon^2 u_i u_j dp = \frac{1}{3}\int_{\mathbb{T}^d}\frac{\partial}{\partial p_i}(\varepsilon^3)\frac{\partial\varepsilon}{\partial p_j}dp = -\frac{1}{3}\int_{\mathbb{T}^d}\varepsilon^3\frac{\partial^2\varepsilon}{\partial p_i\partial p_j}dp = 0 \quad \text{if } i\neq j,$$

since $\partial^2\varepsilon/\partial p_i\partial p_j = 0$ for $i\neq j$. We compute the integral $\int_{\mathbb{T}^d}\varepsilon^2 u_i^2 dp$. First, let $d=1$. Then

$$\int_{\mathbb{T}}\varepsilon^2 u_1^2 dp_1 = 4\varepsilon_0^2(4\pi\varepsilon_0)^2\int_0^1\cos^2(2\pi p_1)\sin^2(2\pi p_1)dp_1 = \frac{\varepsilon_0^2}{2}(4\pi\varepsilon_0)^2 = 8\pi^2\varepsilon_0^4.$$

Furthermore, for $d>1$,

$$\begin{aligned}\int_{\mathbb{T}^d}\varepsilon^2 u_i^2 dp &= 4\varepsilon_0^2(4\pi\varepsilon_0)^2\int_{\mathbb{T}^d}\bigg(\sum_{k=1,\,k\neq i}^d\cos(2\pi p_k)+\cos(2\pi p_i)\bigg)^2\sin^2(2\pi p_i)dp \\ &= 4\varepsilon_0^2(4\pi\varepsilon_0)^2\Bigg(\int_{\mathbb{T}^d}\bigg(\sum_{k=1,\,k\neq i}^d\cos(2\pi p_k)\bigg)^2\sin^2(2\pi p_i)dp + \int_{\mathbb{T}}\cos^2(2\pi p_i)\sin^2(2\pi p_i)dp_i\Bigg) \\ &= 4\varepsilon_0^2(4\pi\varepsilon_0)^2\int_{T^d}\sum_{k=1,\,k\neq i}^d\cos^2(2\pi p_k)\sin^2(2\pi p_i)dp + 8\pi^2\varepsilon_0^4 \\ &= 4\varepsilon_0^2(4\pi\varepsilon_0)^2\sum_{k=1,\,k\neq i}^d\int_0^1\cos^2(2\pi p_k)dp_k\int_0^1\sin^2(2\pi p_i)dp_i + 8\pi^2\varepsilon_0^4 \\ &= (d-1)\varepsilon_0^2(4\pi\varepsilon_0)^2 + 8\pi^2\varepsilon_0^4 = 8(2d-1)\pi^2\varepsilon_0^4.\end{aligned}$$

We conclude that

$$\int_{\mathbb{T}^d}\varepsilon^2 u_i u_j dp = 8(2d-1)\pi^2\varepsilon_0^4\delta_{ij}. \tag{69}$$

## References


[1] G. Alì and V. Romano. Existence and uniqueness for a two-temperature energy-transport model for semiconductors. *J. Math. Anal. Appl.* 449 (2017), 1248-1264.

[2] A. Al-Masoudi, S. Dörscher, S. Häfner, U. Sterr, and C. Lisdat. Noise and instability of an optical lattice clock. *Phys. Rev. A* 92 (2015), 063814, 7 pages.

[3] N. Ben Abdallah and P. Degond. On a hierarchy of macroscopic models for semiconductors. *J. Math. Phys.* 37 (1996), 3308-3333.

[4] E. Bloch. Ultracold quantum gases in optical lattices. *Nature Physics* 1 (2005), 23-30.

[5] L. Boccardo and F. Murat. Almost everywhere convergence of the gradients of solutions to elliptic and parabolic equations. *Nonlin. Anal. TMA* 19 (1992), 581-597.

[6] M. Braukhoff. *Effective Equations for a Cloud of Ultracold Atoms in an Optical Lattice.* PhD thesis, Universität zu Köln, Germany, 2017.
[7] G. Croce. An elliptic problem with degenerate coercivity and a singular quadratic gradient lower order term. *Discrete Contin. Dyn. Sys. Ser. S* 5 (2012), 507-530.
[8] A. Dall' Aglio, D. Giachetti, C. Leone, and S. Segura de Leon. Quasi-linear parabolic equations with degenerate coercivity having a quadratic gradient term. *Ann. I. H. Poincaré – AN* 23 (2006), 97-126.
[9] P. Daskalopoulos and M. del Pino. On the Cauchy problem for $u_t = \Delta \log u$ in higher dimensions. *Math. Ann.* 313 (1999), 189-206.
[10] P. Degond, S. Génieys, and A. Jüngel. A system of parabolic equations in nonequilibrium thermodynamics including thermal and electrical effects. *J. Math. Pures Appl.* 76 (1997), 991-1015.
[11] O. Dutta, M. Gajda, P. Hauke, M. Lewenstein, D.-S. Lühmann, B. Malomed, T. Sowinski, and J. Zakrzewski. Non-standard Hubbard models in optical lattices: a review. *Rep. Prog. Phys.* 78 (2015), 066001, 47 pages.
[12] D. Giachetti and G. Maroscia. Existence results for a class of porous medium type equations with a quadratic gradient term. *J. Evol. Eqs.* 8 (2008), 155-188.
[13] A. Griffin, T. Nikuni, and E. Zaremba. *Bose-Condensed Gases at Finite Temperatures.* Cambridge University Press, Cambridge, 2009.
[14] A. Jaksch. Optical lattices, ultracold atoms and quantum information processing. *Contemp. Phys.* 45 (2004), 367-381.
[15] A. Jüngel. *Transport Equations for Semiconductors.* Lect. Notes Phys. 773. Springer, Berlin, 2009.
[16] A. Jüngel. Global weak solutions to compressible Navier-Stokes equations for quantum fluids. *SIAM J. Math. Anal.* 42 (2010), 1025-1045.
[17] A. Jüngel, S. Krause, and P. Pietra. Diffusive semiconductor moment equations using Fermi-Dirac statistics. *Z. Angew. Math. Phys.* 62 (2011), 623-639.
[18] A. Jüngel, R. Pinnau, and E. Röhrig. Existence analysis for a simplified transient energy-transport model for semiconductors. *Math. Meth. Appl. Sci.* 36 (2013), 1701-1712.
[19] C. Liu, Y. Li, and S. Wang. Asymptotic behavior of the solution to a 3-D simplified energy-transport model for semiconductors. *J. Part. Diff. Eqs.* 29 (2016), 71-88.
[20] S. Mandt, A. Rapp, and A. Rosch. Interacting fermionic atoms in optical lattices diffuse symmetrically upwards and downwards in a gravitational potential. *Phys. Rev. Lett.* 106 (2011), 250602, 4 pages.
[21] N. Ramsey. Thermodynamics and statistical mechanics at negative absolute temperature. *Phys. Rev.* 103 (1956), 20-28.
[22] A. Rapp, S. Mandt, and A. Rosch. Equilibration rates and negative absolute temperatures for ultracold atoms in optical lattices. *Phys. Rev. Lett.* 105 (2010), 220405, 4 pages.
[23] U. Schneider, L. Hackermüller, J. Ph. Ronzheimer, S. Will, S. Braun, T. Best, I. Bloch, E. Demler, S. Mandt, D. Rasch, and A. Rosch. Fermionic transport and out-of-equilibrium dynamics in a homogeneous Hubbard model with ultracold atoms. *Nature Physics* 8 (2012), 213-218.
[24] J. L. Vazquez. *The Porous Medium Equation: Mathematical Theory.* Oxford University Press, Oxford, 2006.
[25] N. Zamponi and A. Jüngel. Global existence analysis for degenerate energy-transport models for semiconductors. *J. Diff. Eqs.* 258 (2015), 2339-2363.

Mathematisches Institut, Universität zu Köln, Weyertal 86-90, 50931 Köln, Germany
*E-mail address*: mbraukho@math.uni-koeln.de

Institute for Analysis and Scientific Computing, Vienna University of Technology, Wiedner Hauptstrasse 8–10, 1040 Wien, Austria
*E-mail address*: juengel@tuwien.ac.at