A finite-volume scheme for a spinorial matrix drift-diffusion model for semiconductors

Claire Chainais-Hillairet, Ansgar Jüngel, Polina Shpartko

A FINITE-VOLUME SCHEME FOR A SPINORIAL MATRIX DRIFT-DIFFUSION MODEL FOR SEMICONDUCTORS

CLAIRE CHAINAIS-HILLAIRET, ANSGAR JÜNGEL, AND POLINA SHPARTKO

Abstract. An implicit Euler finite-volume scheme for a spinorial matrix drift-diffusion model for semiconductors is analyzed. The model consists of strongly coupled parabolic equations for the electron density matrix or, alternatively, of weakly coupled equations for the charge and spin-vector densities, coupled to the Poisson equation for the electric potential. The equations are solved in a bounded domain with mixed Dirichlet-Neumann boundary conditions. The charge and spin-vector fluxes are approximated by a Scharfetter-Gummel discretization. The main features of the numerical scheme are the preservation of nonnegativity and L^∞ bounds of the densities and the dissipation of the discrete free energy. The existence of a bounded discrete solution and the monotonicity of the discrete free energy are proved. For undoped semiconductor materials, the numerical scheme is unconditionally stable. The fundamental ideas are reformulations using spin-up and spin-down densities and certain projections of the spin-vector density, free energy estimates, and a discrete Moser iteration. Furthermore, numerical simulations of a simple ferromagnetic-layer field-effect transistor in two space dimensions are presented.

1. Introduction

The exploitation of the electron spin in semiconductor devices is one of the promising trends for future electronics. Since the electron current can be controlled without changing the carrier concentration, this may allow for (almost) energy-conserving and fast-switching devices and, more generally, for electronic devices based on new operating principles. In the literature, several models have been proposed to describe the spin-polarized transport in semiconductor structures [8, 24]. Drift-diffusion approximations are widely employed [18, 23], since they do not require large computational resources but still describe the main transport phenomena. In this paper, we aim to analyze a finite-volume scheme for a spin drift-diffusion system. Before we explain the model equations, we sketch the state of the art in spinorial drift-diffusion modeling.

The existing drift-diffusion models can be classified into two main groups. The first group is given by two-component drift-diffusion equations for the spin-up and spin-down

Date: October 3, 2015.

2000 *Mathematics Subject Classification.* 65M08, 65M12, 82D37.

Key words and phrases. Spinor drift-diffusion equations, semiconductors, finite volumes, energy dissipation, field-effect transistor.

The authors acknowledge partial support from the Austrian-French Project Amadée of the Austrian Exchange Service (ÖAD). The first and last author have been partially supported by the Austrian Science Fund (FWF), grants P24304, P27352, and W1245.

densities. One version of this model was rigorously derived from the spinor Boltzmann equation in the diffusion limit with strong spin-orbit coupling (compared to the mean-free path) [6]. A mathematical analysis of the limit model was performed in [11], proving the global-in-time existence of weak solutions and their equilibration properties in two space dimensions. In three space dimensions, the well-posedness of the stationary system was shown in [10]. A quantum correction of Bohm potential type was derived in [3].

The second group consists of spin-vector drift-diffusion models in which the spin variable is a vector quantity. Combining the charge density with the spin-vector density, we can define the electron density matrix which solves a spinorial matrix drift-diffusion system. These models can be derived from the spinor Boltzmann equation by assuming a moderate spin-orbit coupling [6]. Projecting the spin-vector density in the direction of the precession vector, we recover the two-component drift-diffusion system as a special case. In [6], the scattering rates are supposed to be scalar quantities. Assuming that the scattering rates are positive definite Hermitian matrices, a more general matrix drift-diffusion model was derived in [19]. The global existence of weak solutions to this model was shown in [16].

The aim of this paper is to analyze an implicit Euler finite-volume approximation of the spinorial matrix drift-diffusion model of [19] and to present some numerical simulations in two space dimensions. A numerical analysis of a finite-volume scheme of the stationary two-component drift-diffusion equations was performed in [10]. A finite-element scheme for a spin-vector equation with given electron current density (but coupled to the Landau-Lifshitz-Gilbert equation) was analyzed in [1] and simulated in [2]. However, no numerical analysis seems to be available so far for general spin-vector drift-diffusion models.

1.1. **The model equations.** The spin-vector model of [19], which is analyzed in this paper, consists of the scaled drift-diffusion equation for the (Hermitian) electron density matrix $N \in \mathbb{C}^{2\times 2}$ and the current density matrix $J \in \mathbb{C}^{2\times 2}$,

$$\partial_t N + \operatorname{div} J + i\gamma[N, \vec{m}\cdot\vec{\sigma}] = \frac{1}{\tau}\left(\frac{1}{2}\operatorname{tr}(N)\sigma_0 - N\right), \tag{1}$$

$$J = -D_0 P^{-1/2}(\nabla N + N\nabla V)P^{-1/2} \quad \text{in } \Omega,\ t > 0, \tag{2}$$

where $[A, B] = AB - BA$ is the commutator of two matrices A and B and $\Omega \subset \mathbb{R}^2$ is a bounded domain. The scaled physical parameters are the strength of the effective magnetic field, $\gamma > 0$, the (normalized) direction of the precession vector $\vec{m} = (m_1, m_2, m_3) \in \mathbb{R}^3$, the spin-flip relaxation time $\tau > 0$, and the diffusion coefficient $D_0 > 0$. The precession vector plays the role of the local direction of the magnetization in the ferromagnet. In the analytic part of this paper, we assume for technical reasons that the precession vector $\vec{m}$ is constant. The triple $\vec{\sigma} = (\sigma_1, \sigma_2, \sigma_3)$ are the Pauli matrices and σ_0 is the unit matrix in $\mathbb{C}^{2\times 2}$:

$$\sigma_0 = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, \quad \sigma_1 = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \quad \sigma_2 = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}, \quad \sigma_3 = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix},$$

where i is the imaginary unit. Furthermore, $\operatorname{tr}(N)$ denotes the trace of the matrix N and $P = \sigma_0 + p\vec{m}\cdot\vec{\sigma}$, where $p \in [0, 1)$ represents the spin polarization of the scattering rates.

The product $\vec{m}\cdot\vec{\sigma}$ equals $m_1\sigma_1+m_2\sigma_2+m_3\sigma_3$. The electric potential V is self-consistently given by the Poisson equation

$$-\lambda_D^2\Delta V = \mathrm{tr}(N) - C(x) \quad \text{in } \Omega, \tag{3}$$

where $\lambda_D > 0$ is the scaled Debye length and $C(x) \geq 0$ denotes the doping profile in the n-doped semiconductor [15]. The boundary and initial conditions are specified below. In the semiconductor literature (e.g. [15, 21]), the sign of the electric potential is opposite. We have chosen the above sign convention in order to be close to the notation of [19]. It does not affect the analytical results.

Remark 1. Equations (1)-(2) are scaled using the time scale τ_s and the length scale L, where $L > 0$ a typical length (e.g. the device size). In the numerical part, we choose τ_s to be equal to the physical spin-relaxation time τ^* such that $\tau = \tau^*/\tau_s = 1$. The density matrix and the doping profile are scaled by $\sup_\Omega C$, and $D_0 = D^*\tau_s/L^2$, $\gamma = \gamma^*\tau_s/\hbar$, where $D^* > 0$ is the physical diffusion coefficient, $\gamma^* > 0$ the physical strength of the effective magnetic field, and $\hbar$ the reduced Planck constant. The density matrix is scaled by $\sup_\Omega C$ and the electric potential by the thermal voltage $U_T = 0.026\,\mathrm{V}$ (at room temperature). □

In this paper, we investigate a scalar form of equations (1)-(2). For this, we develop N and J in the Pauli basis via $N = \frac{1}{2}n_0\sigma_0 + \vec{n}\cdot\vec{\sigma}$ and $J = \frac{1}{2}j_0\sigma_0 + \vec{j}\cdot\vec{\sigma}$, where n_0 is the electron charge density and $\vec{n}$ the spin-vector density. Setting $\vec{n} = (n_1, n_2, n_3)$ and $\vec{j} = (j_1, j_2, j_3)$ and defining $\eta = \sqrt{1-p^2}$, system (1)-(2) can be written equivalently (see [19, Remark 1]) as

$$\partial_t n_0 + \mathrm{div}\, j_0 = 0, \tag{4}$$

$$\partial_t n_\ell + \mathrm{div}\, j_\ell - 2\gamma(\vec{n}\times\vec{m})_\ell = -\frac{n_\ell}{\tau}, \quad \ell = 1,2,3, \tag{5}$$

$$j_0 = \frac{D_0}{\eta^2}(J_0 - 2p\vec{J}\cdot\vec{m}), \quad j_\ell = \frac{D_0}{\eta^2}\left(\eta J_\ell + (1-\eta)(\vec{J}\cdot\vec{m})m_\ell - \frac{p}{2}J_0 m_\ell\right), \quad \ell = 1,2,3, \tag{6}$$

$$J_0 = -\nabla n_0 - n_0\nabla V, \quad \vec{J} = (J_1, J_2, J_3) = -\nabla\vec{n} - \vec{n}\nabla V \quad \text{in } \Omega,\ t > 0. \tag{7}$$

Moreover, the Poisson equation (3) rewrites

$$-\lambda_D^2\Delta V = n_0 - C(x) \quad \text{in } \Omega. \tag{8}$$

System (4)-(8) is strongly coupled due to the cross-diffusion terms in (6) and nonlinear due to the Poisson coupling. Note that any solution $(n_0, \vec{n})$ to (4)-(7) defines a solution N to (1)-(2) and vice versa.

The boundary $\partial\Omega = \Gamma^D \cup \Gamma^N$ is assumed to consist of the union of contacts Γ^D and the isolating boundary part Γ^N. Then the boundary and initial data are given by

$$n_0 = n^D, \quad \vec{n} = 0, \quad V = V^D \quad \text{on } \Gamma^D,\ t > 0, \tag{9}$$

$$\nabla n_0\cdot\nu = \nabla n_\ell\cdot\nu = \nabla V\cdot\nu = 0 \quad \text{on } \Gamma^N,\ t > 0,\ \ell = 1,2,3, \tag{10}$$

$$n_0(0) = n_0^0, \quad \vec{n}(0) = \vec{n}^0 \quad \text{in } \Omega, \tag{11}$$

where ν is the exterior unit normal vector to $\partial\Omega$.

1.2. **Mathematical background.** We aim to design a numerical scheme which preserves some qualitative properties of the continuous model, in particular preservation of the positivity of the charge density, boundedness of the density matrix, and dissipation of the free energy. The main difficulty of the analysis is the strong coupling of the equations (the diffusion matrix is not diagonal), since maximum principle or regularity arguments generally do not apply. The key idea is to introduce two transformations of variables which make the diffusion matrix diagonal and thus reduce the level of coupling.

The first transformation is defined by the spin-up and spin-down densities $n_\pm = \frac{1}{2}n_0 \pm \vec{n}\cdot\vec{m}$. Then system (4)-(7) becomes

$$\partial_t n_+ + \operatorname{div}\big(D_0(1+p)(-\nabla n_+ - n_+\nabla V)\big) = -\frac{1}{2\tau}(n_+ - n_-), \tag{12}$$

$$\partial_t n_- + \operatorname{div}\big(D_0(1-p)(-\nabla n_- - n_-\nabla V)\big) = -\frac{1}{2\tau}(n_- - n_+) \tag{13}$$

and the boundary conditions (9), (10) imply

$$n_\pm = \frac{n^D}{2} \text{ on } \Gamma^D \quad \text{and} \quad \nabla n_\pm\cdot\nu = 0 \text{ on } \Gamma^N,\ t>0. \tag{14}$$

We observe that (4)-(7) implies (12)-(13) but not vice versa. Physically this is clear since the spin-up and spin-down densities contain less information than the full density matrix N. By the Stampacchia truncation method, the nonnegativity and boundedness of $n_\pm$ was shown in [16], thus giving the nonnegativity and boundedness of the charge density $n_0 = n_+ + n_-$. Using the notation $\sum_\pm a_\pm = a_+ + a_-$, the (relative) free energy of the above system is given by the sum of the entropy and the electric energy,

$$E(t) = \sum_\pm \int_\Omega \left(n_\pm(\log n_\pm - 1) - n_\pm \log\frac{n^D}{2} + \frac{n^D}{2}\right)dx + \frac{\lambda_D^2}{2}\int_\Omega |\nabla(V-V^D)|^2 dx. \tag{15}$$

Some formal computations show that it is nonnegative and nonincreasing for $t>0$.

The second transformation is given by the decomposition of $\vec{n}$ in the parallel and perpendicular components with respect to $\vec{m}$: $\vec{n}_\parallel = (\vec{n}\cdot\vec{m})\vec{m}$ and $\vec{n}_\perp = \vec{n} - \vec{n}_\parallel$. The equation for $\vec{n}_\perp$ reads as

$$\partial_t \vec{n}_\perp + \operatorname{div}\left(\frac{D_0}{\eta}(-\nabla\vec{n}_\perp - \vec{n}_\perp\nabla V)\right) - 2\gamma(\vec{n}_\perp\times\vec{m}) = -\frac{\vec{n}_\perp}{\tau}. \tag{16}$$

In [16], it was proved by a Moser-type iteration technique that $\vec{n}_\perp$ is bounded. Since $\vec{n}_\parallel = \frac{1}{2}(n_+ - n_-)\vec{m}$ is bounded as well (see above), this implies an L^∞ bound for $\vec{n}$ and consequently for the density matrix $N = \frac{1}{2}n_0\sigma_0 + \vec{n}\cdot\vec{\sigma}$.

The task is to "translate" these ideas to a finite-volume setting. We approximate the diffusive and convective part of the fluxes simultaneously by using a Scharfetter-Gummel discretization. These fluxes were introduced by Il'in [14] and Scharfetter and Gummel [20] for the classical drift-diffusion model (without spin coupling). The discretizations are second-order accurate in space and preserve the steady states. The dissipativity with an implicit Euler discretization was shown in [9]. The discrete steady states were proved to be bounded [10]. Discrete entropy (free energy) estimates and/or the exponential decay

of the free energy along trajectories towards the global equilibrium were investigated in [5, 12] but still without any spin coupling.

Our main results, detailed in Section 2, are the existence of a bounded discrete solution to a fully discrete finite-volume scheme for (4)-(11) and the monotonicity of the discrete free energy for the spin-up and spin-down densities. The mathematical challenge is the proof of lower and upper bounds for the discrete densities. The "translation" of the Stampacchia truncation argument from the continuous to the discret case in, e.g., (16) faces some difficulties due to the drift term. The main difficulty lies in the fact that the monotonicity of the drift term (with respect to the density variable) cannot be exploited. Therefore, a Moser-type iteration method was employed in [16]. The idea is to derive a uniform estimate for $\vec{n}_\perp$ in the L^q norm of the form

$$\frac{d}{dt}\|\vec{n}_\perp\|_q^q \le cq\|\vec{n}_\perp\|_q^q, \quad t>0,$$

where $\|\cdot\|_q$ denotes the $L^q(\Omega)$ norm and $c>0$ does not depend on $q\in(1,\infty)$. By Gronwall's lemma, this implies that

$$\|\vec{n}(t)\|_q \le e^{ct}\|\vec{n}_\perp(0)\|_q,$$

and the limit $q\to\infty$ shows the claim. The discrete equivalent of the estimate is

$$\|\vec{n}_\perp^k\|_q^q - \|\vec{n}_\perp^{k-1}\|_q^q \le cq\triangle t\|\vec{n}_\perp^k\|_q^q,$$

where $\vec{n}_\perp^k$ is an approximation of $\vec{n}_\perp$ at time t^k and $\triangle t$ is the (uniform) time step size. In order to solve this recursion, we require $1-cq\triangle t>0$, thus imposing a condition on the time step size for fixed q. This motivates additional conditions on the model parameters, which are described and discussed in Section 2.3.

The paper is organized as follows. In Section 2, we detail the numerical scheme and present the main results, in particular the existence of discrete solutions (Theorem 1) and the dissipativity of the discrete free energy (Theorem 2). The proofs are given in Sections 3 and 4. Some numerical tests are presented in Section 5.

2. Numerical method and main results

In this section, we specify the numerical discretization of the spin drift-diffusion system (4)-(11) and state the main results of the paper.

2.1. Notations.

Before we state the numerical scheme, we need to define the mesh of the domain Ω and to introduce some notation. We consider the two-dimensional case only but the scheme can be generalized in a straightforward way to higher dimensions.

Let $\Omega\subset\mathbb{R}^2$ be an open bounded polygonal set. The mesh $\mathcal{M}=(\mathcal{T},\mathcal{E},\mathcal{P})$ is given by a family $\mathcal{T}$ of open polygonal control volumes or cells, a family $\mathcal{E}$ of edges, and a family $\mathcal{P}=(x_K)_{K\in\mathcal{T}}$ of points. We assume that the mesh is admissible in the sense of Definition 9.1 in [7]. This definition implies that the straight line between two neighboring centers of cell (x_K,x_L) is orthogonal to the edge $\sigma=K|L$ between two control volumes K and L and therefore collinear to the unit normal vector $\nu_{K,\sigma}$ to σ outward to K. For instance,

triangular meshes satisfy the admissibility condition if all angles of the triangles are smaller than $\pi/2$ [7, Example 9.1]. Voronoi meshes are also admissible meshes [7, Example 9.2].

Each edge $\sigma \in \mathcal{E}$ is either an internal edge, $\sigma = K|L$, or an exterior edge, $\sigma \subset \partial\Omega$, and we set $\mathcal{E} = \mathcal{E}_{\rm int} \cup \mathcal{E}_{\rm ext}$. We assume that each exterior edge is an element of either the Dirichlet or Neumann boundary such that we can set $\mathcal{E}_{\rm ext} = \mathcal{E}_{\rm ext}^D \cup \mathcal{E}_{\rm ext}^N$. For a given control volume $K \in \mathcal{T}$, we define the set $\mathcal{E}_K$ of the edges of K, which can be written as the union of $\mathcal{E}_{K,\rm int}$, $\mathcal{E}_{K,\rm ext}^D$, and $\mathcal{E}_{K,\rm ext}^N$. For every $\sigma \in \mathcal{E}$, there exists at least one cell $K \in \mathcal{T}$ satisfying $\sigma \in \mathcal{E}_K$, and we denote this cell by K_σ. When σ is an interior edge with $\sigma = K|L$, we have $K_\sigma = K$ or $K_\sigma = L$.

For $K \in \mathcal{T}$ and $\sigma \in \mathcal{E}_K$, we denote by $\mathrm{d}_{K,\sigma}$ the distance $\mathrm{d}_{K,\sigma} = d(x_K, \sigma)$. Then, for $\sigma \in \mathcal{E}_{\rm int}$, $\sigma = K|L$, we define $\mathrm{d}_\sigma = \mathrm{d}_{K,\sigma} + \mathrm{d}_{L,\sigma} = d(x_K, x_L)$ and for $\sigma \in \mathcal{E}_{\rm ext}$ with $\sigma \in \mathcal{E}_K$, $\mathrm{d}_\sigma = \mathrm{d}_{K,\sigma}$. Furthermore, the measure of $\sigma \in \mathcal{E}$ or a set $\omega \subset \Omega$ is denoted by $\mathrm{m}(\sigma)$ or $\mathrm{m}(\omega)$, respectively. In the numerical scheme, we need the so-called transmissibility coefficient $\tau_\sigma = \mathrm{m}(\sigma)/\mathrm{d}_\sigma$ for $\sigma \in \mathcal{E}$. We assume that the mesh satisfies the regularity constraint

$$\exists \xi > 0 : \ \forall K \in \mathcal{T} : \ \forall \sigma \in \mathcal{E}_K : \ \mathrm{d}(x_K, \sigma) \geq \xi \mathrm{diam}(K). \tag{17}$$

The finite-volume scheme for a conservation law with unknown u provides a vector $u_\mathcal{T} = (u_K)_{K\in\mathcal{T}}$ of approximate values and the associated piecewise constant function, still denoted by $u_\mathcal{T}$, $u_\mathcal{T} = \sum_{K\in\mathcal{T}} u_K \mathbf{1}_K$, which approximates the unknown u. Here, $\mathbf{1}_K$ denotes the characteristic function of the cell K. The approximate values of the Dirichlet boundary provide a vector $u_{\mathcal{E}^D} = (u_\sigma)_{\sigma\in\mathcal{E}_{\rm ext}^D}$. The vector containing the approximate values in the control volumes and at the Dirichlet boundary edges is denoted by $u_\mathcal{M} = (u_\mathcal{T}, u_{\mathcal{E}^D})$.

The numerical scheme can be formulated in a compact form by introducing the following notation. For any vector $u_\mathcal{M} = (u_\mathcal{T}, u_{\mathcal{E}^D})$, we define, for all $K \in \mathcal{T}$ and $\sigma \in \mathcal{E}_K$,

$$u_{K,\sigma} = \begin{cases} u_L & \text{if } \sigma = K|L, \\ u_\sigma & \text{if } \sigma = \mathcal{E}_{K,\rm ext}^D, \\ u_K & \text{if } \sigma = \mathcal{E}_{K,\rm ext}^N, \end{cases}$$

and we set $\mathrm{D}u_{K,\sigma} = u_{K,\sigma} - u_K$. We remark that the definition of $u_{K,\sigma}$ ensures that $\mathrm{D}u_{K,\sigma} = 0$ on the Neumann boundary edges. Then the discrete H^1 seminorm for $u_\mathcal{M}$ can be defined by

$$|u_\mathcal{M}|_{1,\mathcal{M}} = \left(\sum_{\sigma\in\mathcal{E}} \tau_\sigma |\mathrm{D}u_{K,\sigma}|^2 \right)^{1/2},$$

where the summation is over all edges $\sigma \in \mathcal{E}$ with $K = K_\sigma$. The L^p norm of $u_\mathcal{T}$ reads as

$$\|u_\mathcal{T}\|_p = \left(\sum_{K\in\mathcal{T}} \mathrm{m}(K)|u_K|^p \right)^{1/p} \quad \text{for } 1 \leq p < \infty \quad \text{and} \quad \|u_\mathcal{T}\|_\infty = \max_{K\in\mathcal{T}} |u_K|.$$

When formulating a finite-volume scheme, we have to define some numerical fluxes $J_{K,\sigma}$ which are consistent approximations of the exact fluxes through the edges $\int_\sigma J \cdot \nu_{K,\sigma} ds$. We impose the conservation of the numerical fluxes $J_{K,\sigma} + J_{L,\sigma}$ for $\sigma = K|L$, requiring that

they vanish on the Neumann boundary edges, $J_{K,\sigma} = 0$ for $\sigma \in \mathcal{E}_{K,\mathrm{ext}}^N$. Then the discrete integration-by-parts formula becomes

$$\sum_{K\in\mathcal{T}} \sum_{\sigma\in\mathcal{E}_K} J_{K,\sigma} u_K = -\sum_{\sigma\in\mathcal{E}} J_{K,\sigma} \mathrm{D}u_{K,\sigma} + \sum_{\sigma\in\mathcal{E}_{\mathrm{ext}}^D} J_{K,\sigma} u_{K,\sigma}. \tag{18}$$

2.2. **Numerical scheme.** At each time step $k \geq 0$, we define the approximate solution $u_{\mathcal{T}}^k = (u_K^k)_{K\in\mathcal{T}}$ for $u \in \{n_0, \vec{n}, V\}$ and the approximate values at the Dirichlet boundary, $u_{\mathcal{E}^D}^k = (u_\sigma^k)_{\sigma\in\mathcal{E}_{\mathrm{ext}}^D}$ (which in fact does not depend on k since the boundary data is time-independent). We first define the initial and boundary conditions corresponding to (11) and (9). We set

$$\begin{aligned} (n_{0,K}^0, \vec{n}_K^0) &= \frac{1}{\mathrm{m}(K)} \int_K (n_0^0, \vec{n}^0) dx \quad \text{for all } K \in \mathcal{T}, \\ (n_{0,\sigma}^D, \vec{n}_\sigma^D, V_\sigma^D) &= \frac{1}{\mathrm{m}(\sigma)} \int_\sigma (n^D, \vec{0}, V^D) ds \quad \text{for all } \sigma \in \mathcal{E}_{\mathrm{ext}}^D. \end{aligned} \tag{19}$$

Note that $\vec{n}_\sigma^D = 0$ for $\sigma \in \mathcal{E}_{\mathrm{ext}}^D$. We may define similarly the quantities $\vec{m}_K, C_K, D_{0,K}, p_K$ for a given $K \in \mathcal{T}$.

We consider a temporal implicit Euler and spatial finite-volume discretization. The scheme for (4), (5), (8) writes, for all $K \in \mathcal{T}$ and $k \geq 1$, as

$$\mathrm{m}(K) \frac{n_{0,K}^k - n_{0,K}^{k-1}}{\triangle t} + \sum_{\sigma\in\mathcal{E}_K} j_{0,K,\sigma}^k = 0, \tag{20}$$

$$\mathrm{m}(K) \frac{\vec{n}_K^k - \vec{n}_K^{k-1}}{\triangle t} + \sum_{\sigma\in\mathcal{E}_K} \vec{j}_{K,\sigma}^k - 2\gamma\, \mathrm{m}(K)(\vec{n}_K^k \times \vec{m}_K) = -\frac{\mathrm{m}(K)}{\tau} \vec{n}_K^k, \tag{21}$$

$$-\lambda_D^2 \sum_{\sigma\in\mathcal{E}_K} \tau_\sigma \mathrm{D}V_{K,\sigma}^k = \mathrm{m}(K)(n_{0,K}^k - C_K), \tag{22}$$

where the discrete counterpart to (6) is, for all $K \in \mathcal{T}$, $\sigma \in \mathcal{E}_K$, $k \geq 0$,

$$j_{0,K,\sigma}^k = \frac{D_\sigma}{\eta_\sigma^2} \big(J_{0,K,\sigma}^k - 2p_\sigma \vec{J}_{K,\sigma}^k \cdot \vec{m}_\sigma \big), \tag{23}$$

$$\vec{j}_{K,\sigma}^k = \frac{D_\sigma}{\eta_\sigma^2} \Big(\eta_\sigma \vec{J}_{K,\sigma}^k + (1-\eta_\sigma)(\vec{J}_{K,\sigma}^k \cdot \vec{m}_\sigma)\vec{m}_\sigma - \frac{p_\sigma}{2} J_{0,K,\sigma}^k \vec{m}_\sigma \Big). \tag{24}$$

The numerical fluxes $J_{0,K,\sigma}^k$ and $J_{\ell,K,\sigma}^k$ are approximations of the integrals $\int_\sigma J_0 \cdot \nu_{K,\sigma} ds$ and $\int_\sigma J_\ell \cdot \nu_{K,\sigma} ds$ at time $k\triangle t$, and we set $\vec{J}_{K,\sigma}^k = (J_{\ell,K,\sigma}^k)_{\ell=1,2,3}$. We recall that J_0 and $\vec{J}$ are defined by (7). We use a Scharfetter-Gummel approximation for the definition of the numerical fluxes. For given $K \in \mathcal{T}$ and $\sigma \in \mathcal{E}_K$, we set

$$J_{\ell,K,\sigma}^k = \tau_\sigma \big(B(\mathrm{D}V_{K,\sigma}^k) n_{\ell,K}^k - B(-\mathrm{D}V_{K,\sigma}^k) n_{\ell,K,\sigma}^k \big), \quad \ell = 0,1,2,3, \tag{25}$$

where B is the Bernoulli function defined by

$$B(x) = \frac{x}{\exp(x) - 1} \text{ for } x \neq 0 \quad \text{and} \quad B(0) = 1.$$

It remains to define the quantities D_σ, $\vec{m}_\sigma$, p_σ and η_σ appearing in (23) and (24). We use a weighted harmonic average on the interior edges and a classical mean value on the boundary edges,

$$D_\sigma = \frac{\mathrm{d}_\sigma\, D_{0,K} D_{0,L}}{\mathrm{d}_{K,\sigma}\, D_{0,L} + \mathrm{d}_{L,\sigma}\, D_{0,K}} \text{ for } \sigma \in \mathcal{E}_{\mathrm{int}},\ \sigma = K|L,\ D_\sigma = \frac{1}{\mathrm{m}(\sigma)} \int_\sigma D_0(s)ds \text{ for } \sigma \in \mathcal{E}^D_{\mathrm{ext}},$$

and similar definitions for $\vec{m}_\sigma$ and p_σ. Furthermore, we set $\eta_\sigma = \sqrt{1-p_\sigma^2}$.

Finally, the boundary conditions are

$$n^k_{0,\sigma} = n^D_{0,\sigma}, \quad \vec{n}^k_\sigma = 0, \quad V^k_\sigma = V^D_\sigma \quad \text{for } \sigma \in \mathcal{E}^D_{\mathrm{ext}}, \tag{26}$$

$$\mathrm{D}n^k_{\ell,K,\sigma} = \mathrm{D}V^k_{K,\sigma} = 0 \quad \text{for } \sigma \in \mathcal{E}^N_{K,\mathrm{ext}},\ \ell = 0,1,2,3,\ k \geq 0. \tag{27}$$

We remark that they imply $J^k_{l,K,\sigma} = 0$ for $\sigma \in \mathcal{E}^N_{K,\mathrm{ext}}$, $\ell = 0,1,2,3$, and $k \geq 0$.

For later use, we note that, using the elementary property $B(x) - B(-x) = -x$ for $x \in \mathbb{R}$, the numerical fluxes can be reformulated in two different manners:

$$J^k_{\ell,K,\sigma} = \tau_\sigma\big(- \mathrm{D}V^k_{K,\sigma} n^k_{\ell,K} - B(-\mathrm{D}V^k_{K,\sigma})\mathrm{D}n^k_{\ell,K,\sigma}\big) \tag{28}$$

$$= \tau_\sigma\big(- \mathrm{D}V^k_{K,\sigma} n^k_{\ell,K,\sigma} - B(\mathrm{D}V^k_{K,\sigma})\mathrm{D}n^k_{\ell,K,\sigma}\big), \quad \ell = 0,1,2,3, \tag{29}$$

and adding these expressions leads to a third formulation:

$$J^k_{\ell,K,\sigma} = \tau_\sigma \left(-\frac{1}{2}(n^k_{\ell,K} + n^k_{\ell,K,\sigma})\mathrm{D}V^k_{K,\sigma} - B^s(\mathrm{D}V^k_{K,\sigma})\mathrm{D}n^k_{\ell,K,\sigma} \right), \tag{30}$$

where

$$B^s(x) = \frac{x}{2}\coth\left(\frac{x}{2}\right) = \frac{B(x)+B(-x)}{2}. \tag{31}$$

2.3. **Main results.** We impose the following assumptions on the domain and the data:

$$\Omega \subset \mathbb{R}^2 \text{ bounded domain},\ \partial\Omega = \Gamma^D \cup \Gamma^N,\ \Gamma^D \cap \Gamma^N = \emptyset,\ \mathrm{m}(\Gamma^D) > 0,\ \Gamma^N \text{ open}, \tag{32}$$

$$D_0,\ p,\ \lambda_D,\ \vec{m} \text{ are constant and } |\vec{m}| = 1,\ C \in L^\infty(\Omega),\ C(x) \geq 0, \tag{33}$$

$$n^0_0,\ \vec{n}^0,\ n^D \in L^\infty(\Omega),\ \frac{1}{2}n^0_0 \pm \vec{n}^0 \cdot m \geq 0,\ n^D \geq 0,\ n^D,\ V^D \in H^1(\Omega), \tag{34}$$

$$\mathcal{M} = (\mathcal{T}, \mathcal{E}, \mathcal{P}) \text{ is an admissible mesh satisfying (17).} \tag{35}$$

We first remark that if $(n^k_{0,\mathcal{T}}, \vec{n}^k_{\mathcal{T}}, V^k_{\mathcal{T}})$ is a solution to scheme (20)-(27) for a given $k \geq 1$ ($(n^0_{0,\mathcal{T}}, \vec{n}^0_{\mathcal{T}})$ are defined as the discretization of the initial conditions), we can define $n^k_{\pm,\mathcal{T}} = \frac{1}{2}n^k_{0,\mathcal{T}} \pm \vec{n}^k_{\mathcal{T}} \cdot \vec{m}$, $\vec{n}^k_{\perp,\mathcal{T}} = \vec{n}^k_{\mathcal{T}} - (\vec{n}^k_{\mathcal{T}} \cdot \vec{m})\vec{m}$. Moreover, as n^D and V^D are defined on the whole domain Ω, we can define $n^D_{\mathcal{T}}$ and $V^D_{\mathcal{T}}$ by taking the mean value of n^D and V^D on each control volume $K \in \mathcal{T}$.

Then the following existence result holds.

Theorem 1 (Existence of a solution to the numerical scheme and L^∞ bounds)**.** *Let assumptions* (32)-(35) *hold. We impose the following constraints:*

$$\triangle t \leq \frac{1}{\alpha} := \frac{\lambda_D^2}{D_0(1+p)\|C\|_\infty}, \quad \tau \leq \frac{\eta\lambda_D^2}{D_0\|C\|_\infty}. \tag{36}$$

Then for $k \geq 1$, there exists a solution $(n_{0,\mathcal{T}}^k, \vec{n}_{\mathcal{T}}^k, V_{\mathcal{T}}^k)$ to scheme (20)-(27) satisyfing

$$0 \leq n_{0,\mathcal{T}}^k \leq 2M^0, \quad 0 \leq n_{\pm,\mathcal{T}}^k \leq M^0, \quad |\vec{n}_{\mathcal{T}}^k| \leq 2M^k \quad \text{in } \Omega,$$

where $M^k = M^0(1-\alpha\triangle t)^{-k}$ and

$$M^0 = \max\left(\frac{1}{2}\sup_{\partial\Omega} n^D,\ \sup_{\Omega}\left(\frac{1}{2}n_0^0 + |\vec{n}^0\cdot\vec{m}|\right),\ \sup|\vec{n}_\perp^0|,\ \sup_{\Omega} C\right).$$

In the continuous case, similar L^∞ bounds for the spin-up and spin-down densities, and therefore for the electron charge density, were shown in [16]. These bounds do not depend on time. The mixing of the spin-vector components prevents the use of the monotonicity argument for $\vec{n}_\perp$, solving (16). Therefore, both in the continuous and discrete situations, the L^∞ bound for the spin-vector density depends on time.

The constraint on $\triangle t$ is needed in the definition of M^k. Furthermore, the condition on τ is necessary to prove the L^∞ bound for $\vec{n}_{\perp,\mathcal{T}}^k$. The numerical results presented in Section 5 indicate that the latter restriction is technical. We stress the fact that our scheme is unconditionally stable if the semiconductor is undoped, i.e. $C = 0$. In this situation, $\triangle t$ and τ can be chosen arbitrarily.

Let us discuss the conditions under which the constraint on τ in (36) is satisfied. Choosing $\tau_s = \tau^*$ (see Remark 1) and observing that the scaled doping profile satisfies $\|C\|_\infty = 1$, we obtain

$$\frac{D^*\tau^*}{L^2} = D_0 \leq \eta\lambda_D^2 = \frac{\eta\varepsilon_r\varepsilon_0 U_T}{q_e L^2 \sup_\Omega C},$$

where ε_r is the relative permittivity of the material, ε_0 the permittivity of vacuum, and q_e the elementary charge. This gives a bound on the spin-flip relaxation time or the maximal physical doping value:

$$\tau^* \sup_\Omega C \leq \frac{\eta\varepsilon_r\varepsilon_0 U_T}{q_e D^*}.$$

For silicon at room temperature and with the choice $D^* = 10^{-3}\,\mathrm{m^2s^{-1}}$ (see Section 5), we obtain $\tau^*\sup_\Omega C \leq 10^7\eta\,\mathrm{m^3s}$. With the relaxation time $\tau^* = 10^{-12}\,\mathrm{s}$, a small spin polarization (such that $\eta \approx 1$), the above bound is satisfied for lowly doped semiconductors, $\sup_\Omega C \lesssim 10^{19}\,\mathrm{m^{-3}}$.

We note that Stampacchia's method applied to the discrete Poisson equation (22) gives an L^∞ bound for the electric potential $V_{\mathcal{T}}^k$. This bound depends on M^0 and is uniform in time. The proof follows the lines of the proof for the continuous equation; see e.g. [22, Section 2.3].

Next, we prove that the scheme dissipates the discrete free energy, defined by

$$E^k = \sum_{\pm}\sum_{K\in\mathcal{T}} \mathrm{m}(K)\left(n_{\pm,K}^k(\log n_{\pm,K}^k - 1) - n_{\pm,K}^k \log\frac{n_K^D}{2} + \frac{n_K^D}{2}\right)$$
$$+ \frac{\lambda_D^2}{2}\sum_{\sigma\in\mathcal{E}} \tau_\sigma(\mathrm{D}(V^k - V^D)_{K,\sigma})^2. \tag{37}$$

Theorem 2 (Dissipation of the discrete free energy)**.** *Let assumptions* (32)-(35) *hold and let* $(n_{0,\mathcal{T}}^k, \vec{n}_{\mathcal{T}}^k, V_{\mathcal{T}}^k)_{k\geq 0}$ *be a solution to scheme* (20)-(27) *satisyfing* $0 \leq n_{\pm,\mathcal{T}}^k \leq M^0$. *We further assume that* $n^D \geq n_* > 0$ *and that* $\log(n^D/2) + V^D$ *is constant in* $\overline{\Omega}$. *Then the mapping* $k \mapsto E^k$ *is nonincreasing, i.e., the scheme dissipates the free energy* (37)*:*

(38)
$$E^k + \frac{\triangle t}{2} \sum_{\pm} D_0(1 \pm p) \sum_{\sigma\in\mathcal{E}} \tau_\sigma \min\{n_{\pm,K}^k, n_{\pm,K,\sigma}^k\} \big(\mathrm{D}(\log n_\pm^k + V^k)_{K,\sigma}\big)^2 \leq E^{k-1}, \quad k \geq 1.$$

The above dissipation inequality for the free energy is the discrete counterpart of the continuous estimate for the free energy (15) [16, Formula (28)]:

$$E(t_2) + \frac{1}{2}\int_{t_1}^{t_2}\int_\Omega \sum_{\pm} D_0(1\pm p) n_\pm |\nabla(\log n_\pm + V)|^2 dxds \leq E(t_1), \quad 0 \leq t_1 \leq t_2.$$

The energy dissipation vanishes if $\log n_\pm + V = \text{const.}$. This equation coincides with the definition of the thermodynamic equilibrium (together with $\vec{n} = 0$ and V solving (22)). Consequently, the assumption $\log(n^D/2) + V^D = \text{const.}$ in Theorem 2, imposed on the Dirichlet boundary, means that we require that the Dirichlet boundary data is compatible with the thermodynamic equilibrium.

Since our estimates are local in time, we may also use nonconstant time step sizes $\triangle t^k$ as long as condition (36) is satisfied.

One may ask if the discrete solution converges to the continuous one when the approximation parameters tend to zero. However, it seems to be difficult to extract a discrete gradient estimate for $n_{\pm,\mathcal{T}}^k$ from the discrete free energy estimate in Theorem 2 since we do not have a suitable discrete version of the chain rule $n_\pm|\nabla \log n_\pm|^2 = 4|\nabla\sqrt{n_\pm}|^2$.

3. Proof of Theorem 1

The proof of Theorem 1 will be presented in two subsections. We first establish the existence of a solution $(n_{0,\mathcal{T}}^k, \vec{n}_{\mathcal{T}}^k, V_{\mathcal{T}}^k)$ at each time step $k \geq 1$ by an induction argument. The proof is based on the fixed-point theorem of Brouwer. In this subsection, we also show L^∞ bounds on $n_{0,\mathcal{T}}^k$ and $n_{\pm,\mathcal{T}}^k$ which depend on k. Then, in the second subsection, we prove that these bounds are in fact uniform with respect to k.

3.1. **Existence of a solution to the scheme.** We first note that the initial condition $(n_{0,\mathcal{T}}^0, \vec{n}_{\mathcal{T}}^0)$ is well-defined by (19). Moreover, the definition of M^0 ensures that $|n_{\perp,\mathcal{T}}^0| \leq M^0$, $0 \leq n_{\pm,\mathcal{T}}^0 \leq M^0$ and therefore $0 \leq n_{0,\mathcal{T}}^0 \leq 2M^0$ and $|\vec{n}_{\pm,\mathcal{T}}^0| \leq 2M^0$.

The proof is done by induction. Let $k \geq 1$. Assuming that $(n_{0,\mathcal{T}}^{k-1}, \vec{n}_{\mathcal{T}}^{k-1}, V_{\mathcal{T}}^{k-1})$ is given and verifies $|n_{\perp,\mathcal{T}}^{k-1}| \leq M^{k-1}$, $0 \leq n_{\pm,\mathcal{T}}^{k-1} \leq M^{k-1}$, we will prove the existence of $(n_{0,\mathcal{T}}^k, \vec{n}_{\mathcal{T}}^k, V_{\mathcal{T}}^k)$, solution to (20)-(27), satisfying these bounds with k instead of $k-1$. Scheme (20)-(27) is a nonlinear system of equations. We prove the existence of a solution by using a fixed-point theorem. Let us denote by θ the cardinality of the mesh $\mathcal{T}$ (the number of control volumes) and let $\mu > 0$. We define an application $F_\mu^k : \mathbb{R}^{4\theta} \to \mathbb{R}^{4\theta}$ such that $F_\mu^k(\rho_\mathcal{T}) = n_\mathcal{T}$, where

$\rho_{\mathcal{T}} = (\rho_{0,\mathcal{T}}, \vec{\rho}_{\mathcal{T}})$ and $n_{\mathcal{T}} = (n_{0,\mathcal{T}}, \vec{n}_{\mathcal{T}})$. It is based on a linearization of the scheme and defined in two steps:

- First, we define $V_{\mathcal{T}} \in \mathbb{R}^{\theta}$ as the solution to the linear system

$$-\lambda_D^2 \sum_{\sigma\in\mathcal{E}_K} \tau_\sigma \mathrm{D}V_{K,\sigma} = \mathrm{m}(K)(\rho_{0,K} - C_K) \quad \text{for } K \in \mathcal{T}, \tag{39}$$

$$V_\sigma = V_\sigma^D \text{ for } \sigma \in \mathcal{E}_{\text{ext}}^D, \quad \mathrm{D}V_{K,\sigma} = 0 \text{ for } \sigma \in \mathcal{E}_{K,\text{ext}}^N.$$

- Second, we construct $n_{\mathcal{T}} = (n_{0,\mathcal{T}}, \vec{n}_{\mathcal{T}}) \in \mathbb{R}^{4\theta}$ as the solution to

$$\frac{\mathrm{m}(K)}{\triangle t}(n_{0,K} - n_{0,K}^{k-1}) + \mu\frac{\mathrm{m}(K)}{\triangle t}(n_{0,K} - \rho_{0,K}) + \sum_{\sigma\in\mathcal{E}_K} j_{0,K,\sigma} = 0 \quad \text{for } K \in \mathcal{T}, \tag{40}$$

$$\begin{aligned}\frac{\mathrm{m}(K)}{\triangle t}(n_{\ell,K} - n_{\ell,K}^{k-1}) + \mu\frac{\mathrm{m}(K)}{\triangle t}(n_{\ell,K} - \rho_{\ell,K}) + \sum_{\sigma\in\mathcal{E}_K} j_{\ell,K,\sigma} \\ - 2\gamma\,\mathrm{m}(K)(\vec{n}_K \times \vec{m})_\ell = -\frac{\mathrm{m}(K)}{\tau} n_{\ell,K} \quad \text{for } K \in \mathcal{T},\ \ell = 1,2,3,\end{aligned} \tag{41}$$

where $j_{0,K,\sigma}$ and $j_{\ell,K,\sigma}$ are defined in (23) and (24), with $J_{\ell,K,\sigma}$ defined in (25), but without the superindex k. The boundary conditions read as

$$n_{0,\sigma} = n_\sigma^D,\ n_{\ell,\sigma} = 0 \quad \text{for } \sigma \in \mathcal{E}_{\text{ext}}^D,\ \ell = 1,2,3, \tag{42}$$

$$\mathrm{D}n_{\ell,K,\sigma} = 0 \quad \text{for } \sigma \in \mathcal{E}_{K,\text{ext}}^N,\ K \in \mathcal{T},\ \ell = 0,1,2,3. \tag{43}$$

The parameter $\mu > 0$ allows us to prove unconditional stability for the linearized problem; see e.g. [4]. The corresponding term vanishes for fixed points $\rho_{\mathcal{T}} = n_{\mathcal{T}}$, so that a fixed point for F_μ^k is a solution to scheme (20)-(27). We choose

$$\mu \geq \frac{D_0\|C\|_\infty}{\lambda_D^2} \max\left\{\frac{1}{\eta^2}, \frac{1+p}{2}\right\} \triangle t. \tag{44}$$

The existence and uniqueness of $V_{\mathcal{T}}$, solution to (39), are obvious since the corresponding matrix is positive definite. As this matrix does not depend on $\rho_{\mathcal{T}}$ and the right-hand side is continuous with respect to $\rho_{\mathcal{T}}$, the first mapping $\rho_{\mathcal{T}} \mapsto V_{\mathcal{T}}$ is continuous from $\mathbb{R}^{4\theta}$ to $\mathbb{R}^{\theta}$. This property is not so obvious for the second mapping, based on the linear system of equations (40)-(43). We will prove this property below (Step 1), in order to guarantee that the mapping F_μ^k is well-defined and continuous.

Then, in order to apply Brouwer's fixed-point theorem, we will prove that F_μ^k preserves the set

$$\mathcal{S}^k = \left\{n_{\mathcal{T}} = (n_{0,\mathcal{T}}, \vec{n}_{\mathcal{T}}) \in \mathbb{R}^{4\theta} :\ 0 \leq n_{\pm,\mathcal{T}} \leq M^k,\ |\vec{n}_{\perp,\mathcal{T}}| \leq M^k\right\}. \tag{45}$$

It is a bounded set because each element $n_{\mathcal{T}} \in \mathcal{S}^k$ verifies $0 \leq n_{0,\mathcal{T}} \leq 2M^k$ and $|\vec{n}_{\mathcal{T}}| \leq 2M^k$. This part of the proof is the most challenging one. Given $\rho_{\mathcal{T}} \in \mathcal{S}^k$ and $n_{\mathcal{T}} = F_\mu^k(\rho_{\mathcal{T}})$, we will first establish the nonnegativity of $n_{\pm,\mathcal{T}}$ (Step 2), then the upper bounds for $n_{\pm,\mathcal{T}}$ (Step 3), and finally the L^∞ bound for $\vec{n}_{\perp,\mathcal{T}}$ (Step 4).

Step 1: Existence and uniqueness of solutions to (40)-(43). The linear system of equations (40)-(43) is a square system of size 4θ. The existence of a solution is equivalent to the uniqueness of a solution and to the invertibility of the corresponding matrix. Therefore, we just have to prove that if the right-hand side to the system is zero then the solution is zero. Thus, we may work with the original linear system assuming homogeneous Dirichlet boundary conditions and setting $n_{0,K}^{k-1} = \rho_{0,K} = 0$ and $\vec{n}_K^{k-1} = \vec{\rho}_K = 0$, in order to set the right-hand side to zero.

We multiply the corresponding equation (40) by $\frac{1}{4}n_{0,K}$ and (41) by $n_{\ell,K}$, sum these four equations, and sum over all control volumes $K \in \mathcal{T}$:

$$\begin{aligned}
0 &= (1+\mu)\sum_{K\in\mathcal{T}} \frac{\mathrm{m}(K)}{4\triangle t} n_{0,K}^2 + (1+\mu)\sum_{K\in\mathcal{T}} \frac{\mathrm{m}(K)}{\triangle t}|\vec{n}_K|^2 + \frac{1}{4}\sum_{K\in\mathcal{T}}\sum_{\sigma\in\mathcal{E}_K} j_{0,K,\sigma} n_{0,K} \\
&\quad + \sum_{K\in\mathcal{T}}\sum_{\sigma\in\mathcal{E}_K} \vec{j}_{K,\sigma}\cdot\vec{n}_K - 2\gamma\sum_{K\in\mathcal{T}} \mathrm{m}(K)(\vec{n}_K\times\vec{m})\cdot\vec{n}_K + \frac{1}{\tau}\sum_{K\in\mathcal{T}} \mathrm{m}(K)|\vec{n}_K|^2 \\
&= T_1 + \cdots + T_6.
\end{aligned}$$

Note that $T_5 = 0$ and T_1, T_2, and T_6 are nonnegative. Thus, it remains to estimate the terms T_3 and T_4. By discrete integration by parts (note that the problem is homogeneous) and the definitions (23)-(24) of $j_{\ell,K,\sigma}$ (omitting the superindex k),

$$\begin{aligned}
T_3 &= -\frac{D_0}{4\eta^2}\sum_{\sigma\in\mathcal{E}} (J_{0,K,\sigma} - 2p\vec{J}_{K,\sigma}\cdot\vec{m})\mathrm{D}n_{0,K,\sigma} =: T_{31} + T_{32}, \\
T_4 &= -\frac{D_0}{\eta^2}\sum_{\sigma\in\mathcal{E}} \Big(\eta\vec{J}_{K,\sigma} + (1-\eta)(\vec{J}_{K,\sigma}\cdot\vec{m})\vec{m} - \frac{p}{2}J_{0,K,\sigma}\vec{m}\Big)\cdot\mathrm{D}\vec{n}_{K,\sigma} =: T_{41} + T_{42} + T_{43}.
\end{aligned}$$

With formulation (30), definition (31) of B^s, and the discrete chain rule $(n_{\ell,K} + n_{\ell,K,\sigma})$ $\times \mathrm{D}n_{\ell,K,\sigma} = \mathrm{D}(n_\ell^2)_{K,\sigma}$, we have

$$\begin{aligned}
T_{31} &= \frac{D_0}{8\eta^2}\sum_{\sigma\in\mathcal{E}} \tau_\sigma\left(2B^s(\mathrm{D}V_{K,\sigma})(\mathrm{D}n_{0,K,\sigma})^2 + \mathrm{D}(n_0^2)_{K,\sigma}\mathrm{D}V_{K,\sigma}\right), \\
T_{32} &= -\frac{pD_0}{4\eta^2}\sum_{\sigma\in\mathcal{E}} \tau_\sigma\big(2B^s(\mathrm{D}V_{K,\sigma})\mathrm{D}\vec{n}_{K,\sigma}\cdot\vec{m} + (\vec{n}_K + \vec{n}_{K,\sigma})\cdot\vec{m}\mathrm{D}V_{K,\sigma}\big)\mathrm{D}n_{0,K,\sigma}, \\
T_{41} &= \frac{D_0}{2\eta}\sum_{\sigma\in\mathcal{E}} \tau_\sigma\left(2B^s(\mathrm{D}V_{K,\sigma})|\mathrm{D}\vec{n}_{K,\sigma}|^2 + \mathrm{D}(|\vec{n}|^2)_{K,\sigma}\mathrm{D}V_{K,\sigma}\right), \\
T_{42} &= \frac{(1-\eta)D_0}{2\eta^2}\sum_{\sigma\in\mathcal{E}} \tau_\sigma\big(2B^s(\mathrm{D}V_{K,\sigma})(\mathrm{D}\vec{n}_{K,\sigma}\cdot\vec{m})^2 + \mathrm{D}((\vec{n}\cdot\vec{m})^2)_{K,\sigma}\mathrm{D}V_{K,\sigma}\big), \\
T_{43} &= -\frac{pD_0}{4\eta^2}\sum_{\sigma\in\mathcal{E}} \tau_\sigma\big(2B^s(\mathrm{D}V_{K,\sigma})\mathrm{D}n_{0,K,\sigma} + (n_{0,K} + n_{0,K,\sigma})\mathrm{D}V_{K,\sigma}\big)\mathrm{D}\vec{n}_{K,\sigma}\cdot\vec{m}.
\end{aligned}$$

We collect all terms from T_3+T_4 involving the function B^s:

$$\begin{aligned}
I_1 &:= \frac{D_0}{\eta^2}\sum_{\sigma\in\mathcal{E}}\tau_\sigma B^s(\mathrm{D}V_{K,\sigma})\bigg(\frac{1}{4}(\mathrm{D}n_{0,K,\sigma})^2+\eta|\mathrm{D}\vec{n}_{K,\sigma}|^2+(1-\eta)(\mathrm{D}\vec{n}_{K,\sigma}\cdot\vec{m})^2 \\
&\quad - p\mathrm{D}n_{0,K,\sigma}\mathrm{D}\vec{n}_{K,\sigma}\cdot\vec{m}\bigg) \\
&= \frac{D_0}{\eta^2}\sum_{\sigma\in\mathcal{E}}\tau_\sigma B^s(\mathrm{D}V_{K,\sigma})\bigg[\begin{pmatrix}\mathrm{D}n_{0,K,\sigma}\\ \mathrm{D}\vec{n}_{K,\sigma}\cdot\vec{m}\end{pmatrix}^\top\begin{pmatrix}1/4 & -p/2\\ -p/2 & 1-\eta^2/2\end{pmatrix}\begin{pmatrix}\mathrm{D}n_{0,K,\sigma}\\ \mathrm{D}\vec{n}_{K,\sigma}\cdot\vec{m}\end{pmatrix} \\
&\quad + \frac{\eta^2}{2}|\mathrm{D}\vec{n}_{K,\sigma}|^2+\eta\left(1-\frac{\eta}{2}\right)\left(|\mathrm{D}\vec{n}_{K,\sigma}|^2-|\mathrm{D}\vec{n}_{K,\sigma}\cdot\vec{m}|^2\right)\bigg].
\end{aligned}$$

The eigenvalues of the 2×2 matrix appearing in I_1 are

$$\lambda_\pm = \frac{1}{8}(5-2\eta^2)\pm\frac{1}{8}\sqrt{(5-2\eta^2)^2-8\eta^2}\ge\frac{1}{4}>0.$$

Then, using the inequalities $B^s(z)\ge 1$ for all $z\in\mathbb{R}$ and $|\mathrm{D}\vec{n}_{K,\sigma}|^2\ge|\mathrm{D}\vec{n}_{K,\sigma}\cdot\vec{m}|^2$ (since $|\vec{m}|^2=1$), it follows that

$$I_1\ge\frac{D_0}{\eta^2}\sum_{\sigma\in\mathcal{E}}\tau_\sigma\left(\lambda_-\big((\mathrm{D}n_{0,K,\sigma})^2+(\mathrm{D}\vec{n}_{K,\sigma}\cdot\vec{m})^2\big)+\frac{\eta^2}{2}|\mathrm{D}\vec{n}_{K,\sigma}|^2\right)\ge 0.$$

Next, we collect in I_2 the remaining terms from T_3+T_4 involving the discrete gradient $\mathrm{D}V_{K,\sigma}$. Taking into account that

$$(\vec{n}_K+\vec{n}_{K,\sigma})\cdot\vec{m}\mathrm{D}n_{0,K,\sigma}+(n_{0,K}+n_{0,K,\sigma})\mathrm{D}\vec{n}_{K,\sigma}\cdot\vec{m}=2\mathrm{D}((\vec{n}\cdot\vec{m})n_0)_{K,\sigma},$$

integrating by parts, and employing the discrete Poisson equation (39), we infer that

$$\begin{aligned}
I_2 &:= \frac{D_0}{2\eta^2}\sum_{\sigma\in\mathcal{E}}\tau_\sigma\mathrm{D}V_{K,\sigma}\bigg(\frac{1}{4}\mathrm{D}(n_0^2)_{K,\sigma}+\eta\mathrm{D}(|\vec{n}|^2)_{K,\sigma}+(1-\eta)\mathrm{D}((\vec{n}\cdot\vec{m})^2)_{K,\sigma} \\
&\quad - p\mathrm{D}((\vec{n}\cdot\vec{m})n_0)_{K,\sigma}\bigg) \\
&= \frac{D_0}{2\eta^2\lambda_D^2}\sum_{K\in\mathcal{T}}\mathrm{m}(K)(\rho_{0,K}-C_K)\bigg(\frac{1}{4}n_{0,K}^2+\eta|\vec{n}_K|^2+(1-\eta)(\vec{n}_K\cdot\vec{m})^2 \\
&\quad - p(\vec{n}_K\cdot\vec{m})n_{0,K}\bigg).
\end{aligned}$$

The sum of the terms in the brackets is nonnegative since

$$\begin{aligned}
&\frac{1}{4}n_{0,K}^2+\eta|\vec{n}_K|^2+(1-\eta)(\vec{n}_K\cdot\vec{m})^2-p(\vec{n}_K\cdot\vec{m})n_{0,K} \\
&\qquad = \left(\frac{1}{2}n_{0,K}-p\vec{n}_K\cdot\vec{m}\right)^2+\eta^2(\vec{n}_K\cdot\vec{m})^2+\eta\big(|\vec{n}_K|^2-(\vec{n}_K\cdot\vec{m})^2\big)\ge 0.
\end{aligned}$$

Therefore, the term involving $\rho_{0,K} \geq 0$ can be omitted, giving

$$\begin{aligned} I_2 &\geq -\frac{D_0}{2\eta^2\lambda_D^2}\sum_{K\in\mathcal{T}} \mathrm{m}(K)C_K\left(\frac{1}{4}n_{0,K}^2 + \eta|\vec{n}_K|^2 + (1-\eta)(\vec{n}_K\cdot\vec{m})^2 - p(\vec{n}_K\cdot\vec{m})n_{0,K}\right) \\ &\geq -\frac{D_0}{2\eta^2\lambda_D^2}\|C\|_\infty \sum_{K\in\mathcal{T}} \mathrm{m}(K)\left(\frac{1}{2}n_{0,K}^2 + 2|\vec{n}_K|^2\right) = -\frac{D_0}{\eta^2\lambda_D^2}\|C\|_\infty\left(\frac{1}{4}\|n_{0,\mathcal{T}}\|_2^2 + \|\vec{n}_\mathcal{T}\|_2^2\right). \end{aligned}$$

This shows finally that

$$T_3 + T_4 \geq -\frac{D_0\|C\|_\infty}{\eta^2\lambda_D^2}\left(\frac{1}{4}\|n_{0,\mathcal{T}}\|_2^2 + \|\vec{n}_\mathcal{T}\|_2^2\right),$$

and summarizing the estimates for $T_1,\ldots,T_6$, we conclude that

$$\left(\frac{1+\mu}{\triangle t} - \frac{D_0\|C\|_\infty}{\eta^2\lambda_D^2}\right)\left(\frac{1}{4}\|n_{0,\mathcal{T}}\|_2^2 + \|\vec{n}_\mathcal{T}\|_2^2\right) \leq 0.$$

Hence, choosing μ as in (44), the first bracket is positive, showing that $n_{0,\mathcal{T}} = 0$ and $\vec{n}_\mathcal{T} = 0$, which proves the invertibility of the linear system of equations (40)-(43). The second step involved in the definition of F_μ^k, $(V_\mathcal{T}, \rho_\mathcal{T}) \to n_\mathcal{T}$, is a well-defined mapping. Moreover, the matrix and the right-hand side of the linear system of equations are continuous with respect to $(V_\mathcal{T}, \rho_\mathcal{T})$ so that the mapping is continuous.

Step 2: Nonnegativity of $n_{\pm,\mathcal{T}}$. We will prove that $n_{\pm,K} \geq 0$ for all $K \in \mathcal{T}$. Multiplying (41) by $\vec{m}$ and adding or subtracting it from (40), multiplied by $\frac{1}{2}$, we find that

$$\begin{aligned} &\frac{\mathrm{m}(K)}{\triangle t}(n_{\pm,K} - n_{\pm,K}^{k-1}) + \mu\frac{\mathrm{m}(K)}{\triangle t}(n_{\pm,K} - \rho_{\pm,K}) + D_0(1\pm p)\sum_{\sigma\in\mathcal{E}_K} J_{\pm,K,\sigma} \\ &\qquad = \mp\frac{\mathrm{m}(K)}{2\tau}(n_{+,K} - n_{-,K}), \end{aligned} \tag{46}$$

where $\rho_{\pm,K} = \frac{1}{2}\rho_{0,K} \pm \vec{\rho}_K\cdot\vec{m}_K$ and $J_{\pm,K,\sigma} = \frac{1}{2}J_{0,K,\sigma} \pm \vec{J}_{K,\sigma}\cdot\vec{m}$, i.e.

$$J_{\pm,K,\sigma} = \tau_\sigma\big(B(\mathrm{D}V_{K,\sigma})n_{\pm,K} - B(-\mathrm{D}V_{K,\sigma})n_{\pm,K,\sigma}\big). \tag{47}$$

Then, multiplying (46) by $n_{\pm,K}^- = \min\{0, n_{\pm,K}\}$, summing over all control volumes $K \in \mathcal{T}$, and adding both equations, it follows that

$$\begin{aligned} 0 &= \sum_\pm\sum_{K\in\mathcal{T}}\frac{\mathrm{m}(K)}{\triangle t}(n_{\pm,K} - n_{\pm,K}^{k-1})n_{\pm,K}^- + \mu\sum_\pm\sum_{K\in\mathcal{T}}\frac{\mathrm{m}(K)}{\triangle t}(n_{\pm,K} - \rho_{\pm,K})n_{\pm,K}^- \\ &\quad + D_0\sum_\pm(1\pm p)\sum_{K\in\mathcal{T}}\sum_{\sigma\in\mathcal{E}_K} J_{\pm,K,\sigma}n_{\pm,K}^- + \sum_{K\in\mathcal{T}}\frac{\mathrm{m}(K)}{2\tau}(n_{+,K} - n_{-,K})(n_{+,K}^- - n_{-,K}^-) \\ &=: T_7 + T_8 + T_9 + T_{10}, \end{aligned} \tag{48}$$

Since $n_{\pm,K}n_{\pm,K}^- = (n_{\pm,K}^-)^2$, $n_{\pm,K}^{k-1} \geq 0$, and $\rho_{\pm,K} \geq 0$, the first two terms in (48) can be estimated as

$$T_7 \geq \sum_{\pm}\sum_{K\in\mathcal{T}} \frac{\mathrm{m}(K)}{\triangle t}(n_{\pm,K}^-)^2, \quad T_8 \geq \mu \sum_{\pm}\sum_{K\in\mathcal{T}} \frac{\mathrm{m}(K)}{\triangle t}(n_{\pm,K}^-)^2.$$

The monotonicity of the mapping $z \mapsto z^-$ shows that T_{10} is nonnegative. By the discrete integration-by-parts formula (18), the third term in (48) becomes

$$T_9 = -D_0 \sum_{\pm}(1 \pm p) \sum_{\sigma\in\mathcal{E}} J_{\pm,K,\sigma}\mathrm{D}(n_{\pm}^-)_{K,\sigma}.$$

The sum over the boundary edges vanishes since $n_{\pm,K,\sigma}^- = 0$ for all $\sigma \in \mathcal{E}_{\mathrm{ext}}^D$. We claim that

$$-J_{\pm,K,\sigma}\mathrm{D}(n_{\pm}^-)_{K,\sigma} \geq \frac{1}{2}\tau_\sigma \mathrm{D}V_{K,\sigma}\big((n_{\pm,K,\sigma}^-)^2 - (n_{\pm,K}^-)^2\big), \quad \text{for } K \in \mathcal{T} \text{ and } \sigma \in \mathcal{E}_K, \tag{49}$$

such that

$$T_9 \geq \frac{D_0}{2}\sum_{\pm}(1 \pm p)\sum_{\sigma\in\mathcal{E}} \tau_\sigma \mathrm{D}V_{K,\sigma}\mathrm{D}\big((n_{\pm}^-)^2\big)_{K,\sigma}. \tag{50}$$

To prove (49), we distinguish the cases $\mathrm{D}V_{K,\sigma} \geq 0$ and $\mathrm{D}V_{K,\sigma} \leq 0$. If $\mathrm{D}V_{K,\sigma} \geq 0$, we apply a formulation similar to (29), leading to

$$-J_{\pm,K,\sigma}\mathrm{D}(n_{\pm}^-)_{K,\sigma} = \tau_\sigma\big(\mathrm{D}V_{K,\sigma}n_{\pm,K,\sigma} + B(\mathrm{D}V_{K,\sigma})\mathrm{D}n_{\pm,K,\sigma}\big)\mathrm{D}(n_{\pm}^-)_{K,\sigma}.$$

Then, using the nonnegativity of the function B, the monotonicity of the mapping $z \mapsto z^-$, and the inequality $z(z^- - y^-) \geq \frac{1}{2}((z^-)^2 - (y^-)^2)$, we obtain (49). If $\mathrm{D}V_{K,\sigma} \leq 0$, we employ formulation (28), so that

$$-J_{\pm,K,\sigma}\mathrm{D}(n_{\pm}^-)_{K,\sigma} = \tau_\sigma\big(\mathrm{D}V_{K,\sigma}n_{\pm,K} + B(-\mathrm{D}V_{K,\sigma})\mathrm{D}n_{\pm,K,\sigma}\big)\mathrm{D}(n_{\pm}^-)_{K,\sigma}$$

and similar arguments lead to (49).

Applying discrete integration by parts to the right-hand side of (50) (the boundary term vanishes since the boundary data is nonnegative) and employing the discrete Poisson equation (39), we find that

$$\begin{aligned}
T_9 &\geq -\frac{D_0}{2}\sum_{K\in\mathcal{T}}\sum_{\sigma\in\mathcal{E}_K} \tau_\sigma \mathrm{D}V_{K,\sigma}\big((1+p)(n_{+,K}^-)^2 + (1-p)(n_{-,K}^-)^2\big) \\
&= \frac{D_0}{2\lambda_D^2}\sum_{K\in\mathcal{T}} \mathrm{m}(K)(\rho_{0,K} - C_K)\big((1+p)(n_{+,K}^-)^2 + (1-p)(n_{-,K}^-)^2\big) \\
&\geq -\frac{D_0}{2\lambda_D^2}\|C\|_\infty \sum_{K\in\mathcal{T}} \mathrm{m}(K)\big((1+p)(n_{+,K}^-)^2 + (1-p)(n_{-,K}^-)^2\big).
\end{aligned}$$

Summarizing the above estimates, we conclude from (48) that

$$\left(\frac{1+\mu}{\triangle t} - \frac{D_0(1+p)}{2\lambda_D^2}\|C\|_\infty\right)\sum_{K\in\mathcal{T}} \mathrm{m}(K)\big((n_{+,K}^-)^2 + (n_{-,K}^-)^2\big) \leq 0.$$

By the choice of μ in (44), we deduce that

$$\sum_{K\in\mathcal{T}} \mathrm{m}(K)\big((n_{+,K}^-)^2 + (n_{-,K}^-)^2\big) \le 0,$$

which implies that $n_{\pm,K}^- = 0$ and hence $n_{\pm,K} \ge 0$ for all $K \in \mathcal{T}$.

Step 3: Upper bounds for $n_{\pm,\mathcal{T}}$. The goal is to show that $n_{\pm,K} \le M^k$ for all $K \in \mathcal{T}$, where M^k is defined in Theorem 1. We multiply (46) by $(n_{\pm,K} - M^k)^+ = \max\{0, n_{\pm,K} - M^k\}$, sum over all $K \in \mathcal{T}$, and add both equations:

$$\begin{aligned}
0 &= \sum_{\pm}\sum_{K\in\mathcal{T}} \frac{\mathrm{m}(K)}{\triangle t}\big((n_{\pm,K} - M^k) - (n_{\pm,K}^{k-1} - M^{k-1})\big)(n_{\pm,K} - M^k)^+ \\
&\quad + \mu \sum_{\pm}\sum_{K\in\mathcal{T}} \frac{\mathrm{m}(K)}{\triangle t}\big((n_{\pm,K} - M^k) - (\rho_{\pm,K} - M^k)\big)(n_{\pm,K} - M^k)^+ \\
&\quad + \sum_{\pm}\sum_{K\in\mathcal{T}} \frac{\mathrm{m}(K)}{\triangle t}(M^k - M^{k-1})(n_{\pm,K} - M^k)^+ \\
&\quad + D_0 \sum_{\pm}(1\pm p)\sum_{K\in\mathcal{T}}\sum_{\sigma\in\mathcal{E}_K} J_{\pm,K,\sigma}(n_{\pm,K} - M^k)^+ \\
&\quad + \frac{1}{2}\sum_{\pm}\sum_{K\in\mathcal{T}} \mathrm{m}(K)\big((n_{+,K} - M^k) - (n_{-,K} - M^k)\big)(\pm(n_{\pm,K} - M^k)^+) \\
&=: T_{11} + T_{12} + T_{13} + T_{14} + T_{15}.
\end{aligned} \tag{51}$$

Using the inequality $(z-y)z^+ \ge \frac{1}{2}((z^+)^2 - (y^+)^2)$, the first two terms are estimated by

$$\begin{aligned}
T_{11} &\ge \frac{1}{2\triangle t}\sum_{\pm}\sum_{K\in\mathcal{T}} \mathrm{m}(K)\big((n_{\pm,K} - M^k)^+\big)^2, \\
T_{12} &\ge \frac{\mu}{2\triangle}\sum_{\pm}\sum_{K\in\mathcal{T}} \mathrm{m}(K)\big((n_{\pm,K} - M^k)^+\big)^2,
\end{aligned}$$

since $n_{\pm,K}^{k-1} \le M^{k-1}$ and $\rho_{\pm,K} \le M^k$ by assumption. By definition of M^k, the third term T_{13} becomes

$$T_{13} = \alpha M^k \sum_{\pm}\sum_{K\in\mathcal{T}} \mathrm{m}(K)(n_{\pm,K} - M^k)^+,$$

and the last term T_{15} is nonnegative.

It remains to estimate T_{14}. By discrete integration by parts (the boundary term vanishes in view of $n_{\pm,\sigma}^D \le M^k$ for $\sigma \in \mathcal{E}_{\mathrm{ext}}^D$), we find that

$$T_{14} = -D_0 \sum_{\pm}(1\pm p)\sum_{\sigma\in\mathcal{E}} J_{\pm,K,\sigma}\mathrm{D}\big((n_\pm - M^k)^+\big)_{K,\sigma}.$$

Similarly as in Step 2, we claim that the following estimate holds:

$$T_{14} \geq D_0 \sum_{\pm}(1 \pm p) \sum_{\sigma\in\mathcal{E}} \tau_\sigma \mathrm{D}V_{K,\sigma} \mathrm{D}\left(\frac{1}{2}((n_\pm - M^k)^+)^2 + M^k(n_\pm - M^k)^+\right)_{K,\sigma}.$$

Indeed, let first $\mathrm{D}V_{K,\sigma} \geq 0$. Using the inequalities $\mathrm{D}(n_\pm - M^k)_{K,\sigma}\mathrm{D}((n_\pm - M^k)^+)_{K,\sigma} \geq 0$ and $(n_{\pm,K,\sigma} - M^k)\mathrm{D}((n_\pm - M^k)^+)_{K,\sigma} \geq \frac{1}{2}\mathrm{D}(((n_\pm - M^k)^+)^2)_{K,\sigma} \geq 0$, it follows from (29) that

$$\begin{aligned} -J_{\pm,K,\sigma} &\geq \tau_\sigma \mathrm{D}V_{K,\sigma}\big((n_{\pm,K,\sigma} - M^k) + M^k\big)\mathrm{D}((n_\pm - M)^+)_{K,\sigma} \\ &\geq \tau_\sigma \mathrm{D}V_{K,\sigma}\left(\frac{1}{2}\mathrm{D}(((n_\pm - M^k)^+)^2)_{K,\sigma} + M^k\mathrm{D}((n_\pm - M^k)^+)_{K,\sigma}\right). \end{aligned}$$

The proof for $\mathrm{D}V_{K,\sigma} \leq 0$ is similar, employing formulation (28). Then, integrating by parts and employing the Poisson equation and $\rho_{0,K} \geq 0$,

$$\begin{aligned} T_{14} &\geq \frac{D_0}{\lambda_D^2}\sum_{\pm}(1 \pm p)\sum_{K\in\mathcal{T}} \mathrm{m}(K)(\rho_{0,K} - C_K)\left(\frac{1}{2}((n_{\pm,K} - M^k)^+)^2 + M^k(n_{\pm,K} - M^k)^+\right) \\ &\geq -\frac{D_0}{\lambda_D^2}\|C\|_\infty(1+p)\sum_{\pm}\sum_{K\in\mathcal{T}} \mathrm{m}(K)\left(\frac{1}{2}((n_{\pm,K} - M^k)^+)^2 + M^k(n_{\pm,K} - M^k)^+\right). \end{aligned}$$

Summarizing the above estimates, we infer from (51) that

$$\begin{aligned} &\left(\frac{1+\mu}{2\triangle t} - \frac{D_0}{\lambda_D^2}\|C\|_\infty(1+p)\right)\sum_{\pm}\sum_{K\in\mathcal{T}} \mathrm{m}(K)\big((n_{\pm,K} - M^k)^+\big)^2 \\ &\quad + \left(\alpha - \frac{D_0}{\lambda_D^2}\|C\|_\infty(1+p)\right) M^k \sum_{\pm}\sum_{K\in\mathcal{T}} \mathrm{m}(K)(n_{\pm,K} - M^k)^+ \leq 0. \end{aligned}$$

Then, choosing μ as in (44) and taking into account the definition of α, we infer that $n_{\pm,K} \leq M^k$ for $K \in \mathcal{T}$.

Step 4: L^∞ bound for $\vec{n}^k_{\perp,\mathcal{T}}$. We prove a uniform L^{2q} bound for $\vec{n}_{\perp,K} = \vec{n}_K - (\vec{n}_K \cdot \vec{m})\vec{m}$. For this, we multiply the vector version of (41) (omitting the superindex k) by $\vec{m}$ twice, and taking the difference of the equations for $\vec{n}_K$ and $(\vec{n}_K \cdot \vec{m})\vec{m}$, we obtain

$$\begin{aligned} &\frac{\mathrm{m}(K)}{\triangle t}\big(\vec{n}_{\perp,K} - \vec{n}^{k-1}_{\perp,K} + \mu(\vec{n}_{\perp,K} - \vec{\rho}_{\perp,K})\big) + \frac{D_0}{\eta}\sum_{\sigma\in\mathcal{E}_K} \vec{J}_{\perp,K,\sigma} - 2\gamma\,\mathrm{m}(K)(\vec{n}_{\perp,K} \times \vec{m}) \\ &\qquad = -\frac{\mathrm{m}(K)}{\tau}\vec{n}_{\perp,K}, \end{aligned} \tag{52}$$

where $\vec{\rho}_{\perp,K} = \vec{\rho}_K - (\vec{\rho}_K \cdot \vec{m})\vec{m}$, and $\vec{J}_{\perp,K,\sigma}$ is given by

$$\vec{J}_{\perp,K,\sigma} = \tau_\sigma\big(-\mathrm{D}V_{K,\sigma}\vec{n}_{\perp,K} - B(-\mathrm{D}V_{K,\sigma})\mathrm{D}\vec{n}_{\perp,K,\sigma}\big) \tag{53}$$

$$= \tau_\sigma\big(-\mathrm{D}V_{K,\sigma}\vec{n}_{\perp,K,\sigma} - B(\mathrm{D}V_{K,\sigma})\mathrm{D}\vec{n}_{\perp,K,\sigma}\big). \tag{54}$$

Then, multiplying (52) by $|\vec{n}_{\perp,K}|^{2(q-1)}\vec{n}_{\perp,K}$ (where $q\in\mathbb{N}$) and summing over $K\in\mathcal{T}$, we arrive at $T_{16}+T_{17}+T_{18}+T_{19}=0$ with

$$\begin{aligned}
T_{16}&=\frac{1}{\triangle t}\sum_{K\in\mathcal{T}}\mathrm{m}(K)(\vec{n}_{\perp,K}-\vec{n}_{\perp,K}^{k-1})\cdot\vec{n}_{\perp,K}|\vec{n}_{\perp,K}|^{2(q-1)},\\
T_{17}&=\frac{\mu}{\triangle t}\sum_{K\in\mathcal{T}}\mathrm{m}(K)(\vec{n}_{\perp,K}-\vec{\rho}_{\perp,K})\cdot\vec{n}_{\perp,K}|\vec{n}_{\perp,K}|^{2(q-1)},\\
T_{18}&=\frac{D_0}{\eta}\sum_{K\in\mathcal{T}}\sum_{\sigma\in\mathcal{E}_K}\vec{J}_{\perp,K,\sigma}\cdot\vec{n}_{\perp,K}|\vec{n}_{\perp,K}|^{2(q-1)},\\
T_{19}&=\sum_{K\in\mathcal{T}}\mathrm{m}(K)|\vec{n}_{\perp,K}|^{2q}.
\end{aligned}$$

The elementary inequality $|\vec{a}|^{2(q-1)}\vec{a}\cdot(\vec{a}-\vec{b})\ge(|\vec{a}|^{2q}-|\vec{b}|^{2q})/(2q)$ for $\vec{a},\vec{b}\in\mathbb{R}^3$ shows that

$$\begin{aligned}
T_{16}&\ge\frac{1}{2q\triangle t}\left(\sum_{K\in\mathcal{T}}\mathrm{m}(K)|\vec{n}_{\perp,K}|^{2q}-\sum_{K\in\mathcal{T}}\mathrm{m}(K)|\vec{n}_{\perp,K}^{k-1}|^{2q}\right),\\
T_{17}&\ge\frac{\mu}{2q\triangle t}\left(\sum_{K\in\mathcal{T}}\mathrm{m}(K)|\vec{n}_{\perp,K}|^{2q}-\sum_{K\in\mathcal{T}}\mathrm{m}(K)|\vec{\rho}_{\perp,K}|^{2q}\right).
\end{aligned}$$

By discrete integration by parts (observe that $\vec{n}_{\perp,K,\sigma}=0$ for $\sigma\in\mathcal{E}_{\mathrm{ext}}^D$),

$$T_{18}=-\frac{D_0}{\eta}\sum_{\sigma\in\mathcal{E}}\vec{J}_{\perp,K,\sigma}\cdot\big(\vec{n}_{\perp,K,\sigma}|\vec{n}_{\perp,K,\sigma}|^{2(q-1)}-\vec{n}_{\perp,K}|\vec{n}_{\perp,K}|^{2(q-1)}\big).$$

Again, we distinguish the cases $\mathrm{D}V_{K,\sigma}\ge0$ and $\mathrm{D}V_{K,\sigma}<0$ for given $K\in\mathcal{T}$ and $\sigma\in\mathcal{E}_K$. First, let $\mathrm{D}V_{K,\sigma}\ge0$ and use formulation (54) of the numerical flux. This gives

$$\begin{aligned}
&-\vec{J}_{\perp,K,\sigma}\cdot\big(\vec{n}_{\perp,K,\sigma}|\vec{n}_{\perp,K,\sigma}|^{2(q-1)}-\vec{n}_{\perp,K}|\vec{n}_{\perp,K}|^{2(q-1)}\big)\\
&\quad=\tau_\sigma\mathrm{D}V_{K,\sigma}\vec{n}_{\perp,K,\sigma}\cdot\big(\vec{n}_{\perp,K,\sigma}|\vec{n}_{\perp,K,\sigma}|^{2(q-1)}-\vec{n}_{\perp,K}|\vec{n}_{\perp,K}|^{2(q-1)}\big)\\
&\qquad+\tau_\sigma B(\mathrm{D}V_{K,\sigma})(\vec{n}_{\perp,K,\sigma}-\vec{n}_{\perp,K})\cdot\big(\vec{n}_{\perp,K,\sigma}|\vec{n}_{\perp,K,\sigma}|^{2(q-1)}-\vec{n}_{\perp,K}|\vec{n}_{\perp,K}|^{2(q-1)}\big).
\end{aligned}$$

Because of

$$\begin{aligned}
\vec{a}\cdot(\vec{a}|\vec{a}|^{2(q-1)}-\vec{b}|\vec{b}|^{2(q-1)})&=|\vec{a}|^{2q}-\vec{a}\cdot\vec{b}|\vec{b}|^{2(q-1)}\ge|\vec{a}|^{2q}-\frac{1}{2q}|\vec{a}|^{2q}-\left(1-\frac{1}{2q}\right)|\vec{b}|^{2q}\\
&\ge\left(1-\frac{1}{2q}\right)(|\vec{a}|^{2q}-|\vec{b}|^{2q})\quad\text{for all }\vec{a},\vec{b}\in\mathbb{R}^3,
\end{aligned}$$

applied to $\vec{a}=\vec{n}_{\perp,K,\sigma}$ and $\vec{b}=\vec{n}_{\perp,K}$, and the monotonicity of the mapping $\vec{a}\mapsto\vec{a}|\vec{a}|^{2(q-1)}$, we find that

$$\begin{aligned}
&-\vec{J}_{\perp,K,\sigma}\cdot\big(\vec{n}_{\perp,K,\sigma}|\vec{n}_{\perp,K,\sigma}|^{2(q-1)}-\vec{n}_{\perp,K}|\vec{n}_{\perp,K}|^{2(q-1)}\big)\\
&\qquad\ge\tau_\sigma\left(1-\frac{1}{2q}\right)\mathrm{D}V_{K,\sigma}\big(|\vec{n}_{\perp,K,\sigma}|^{2q}-|\vec{n}_{\perp,K}|^{2q}\big).
\end{aligned}$$

This result still holds if $\mathrm{D}V_{K,\sigma} < 0$, thanks to formulation (53). Therefore

$$T_{18} \geq \frac{D_0}{\eta}\left(1 - \frac{1}{2q}\right)\sum_{\sigma\in\mathcal{E}} \tau_\sigma \mathrm{D}V_{K,\sigma}\mathrm{D}(|\vec{n}_\perp|^{2q})_{K,\sigma}.$$

Using discrete integration by parts and the Poisson equation (22) leads to

$$T_{18} \geq \frac{D_0}{\eta\lambda_D^2}\left(1 - \frac{1}{2q}\right)\sum_{K\in\mathcal{T}} \mathrm{m}(K)(\rho_{0,K} - C_K)|\vec{n}_{\perp,K}|^{2q} \geq -\frac{D_0}{\eta\lambda_D^2}\|C\|_\infty \sum_{K\in\mathcal{T}} \mathrm{m}(K)|\vec{n}_{\perp,K}|^{2q}.$$

Summarizing the above estimates, we obtain

$$\left(1 + \mu + 2q\triangle t\Big(\frac{1}{\tau} - \frac{D_0\|C\|_\infty}{\eta\lambda_D^2}\Big)\right)\sum_{K\in\mathcal{T}} \mathrm{m}(K)|\vec{n}_{\perp,K}|^{2q} \leq \sum_{K\in\mathcal{T}} \mathrm{m}(K)|\vec{n}_{\perp,K}^{k-1}|^{2q} + \mu\sum_{K\in\mathcal{T}} \mathrm{m}(K)|\vec{\rho}_{\perp,K}|^{2q}.$$

Condition (36) on τ, the induction hypothesis $\|\vec{n}_{\perp,\mathcal{T}}^{k-1}\|_\infty \leq M^{k-1} \leq M^k$, and the fact that $\rho \in \mathcal{S}^k$ (see (45) for the definition of $\mathcal{S}^k$), such that $\|\vec{\rho}_{\perp,\mathcal{T}}\|_\infty \leq M^k$, imply that

$$\|\vec{n}_{\perp,\mathcal{T}}\|_{2q} \leq \mathrm{meas}(\Omega)^{1/(2q)}M^k \quad \text{for } q \geq 1.$$

Passing to the limit $q \to +\infty$, we deduce that $\|n_{\perp,\mathcal{T}}\|_\infty \leq M^k$.

Conclusion. In Step 1, we have proved that the mapping F_μ^k is well-defined and continuous. In Steps 2-4, we have proved that F_μ^k preserves the bounded set $\mathcal{S}^k$. Thus, the fixed-point theorem of Brouwer shows the existence of a fixed point to F_μ^k, belonging to $\mathcal{S}^k$. Let us denote this fixed point by $n_\mathcal{T}^k = (n_{0,\mathcal{T}}^k, \vec{n}_\mathcal{T}^k)$. It is a solution to scheme (20)–(27) at step k and satisfies

$$0 \leq n_{\pm,K}^k \leq M^k \quad \text{and} \quad |\vec{n}_{\perp,K}^k| \leq M^k, \quad \text{for } K \in \mathcal{T}.$$

3.2. **Uniform bounds for the spin-up and spin-down densities.** In order to conclude the proof of Theorem 1, it remains to prove that the upper bounds on the spin-up and spin-down densities in fact do not depend on k. The negativity of these densities is already proved above.

We assume as induction hypothesis that $n_{\pm,K}^{k-1} \leq M^0$ for all $K \in \mathcal{T}$ (this property is ensured for $k = 1$ by the definition of M^0). Scheme (20)-(21) implies that

$$\frac{\mathrm{m}(K)}{\triangle t}(n_{\pm,K}^k - n_{\pm,K}^{k-1}) + D_0(1 \pm p)\sum_{\sigma\in\mathcal{E}_K} J_{\pm,K,\sigma}^k = \mp\frac{\mathrm{m}(K)}{2\tau}(n_{+,K}^k - n_{-,K}^k). \tag{55}$$

As in Step 3 above, we multiply (55) by $(n_{\pm,K}^k - M^0)^+$, sum over all $K \in \mathcal{T}$ and add both equations. This yields $S_1 + S_2 + S_3 = 0$, where

$$S_1 = \sum_\pm \sum_{K\in\mathcal{T}} \frac{\mathrm{m}(K)}{\triangle t}\big((n_{\pm,K}^k - M^0) - (n_{\pm,K}^{k-1} - M^0)\big)(n_{\pm,K}^k - M^0)^+,$$

$$S_2 = D_0 \sum_{\pm} (1 \pm p) \sum_{K \in \mathcal{T}} \sum_{\sigma \in \mathcal{E}_K} J^k_{\pm,K,\sigma} (n^k_{\pm,K} - M^0)^+,$$

$$S_3 = \frac{1}{2\tau} \sum_{K \in \mathcal{T}} \mathrm{m}(K)(n^k_{+,K} - n^k_{-,K}) \left((n^k_{+,K} - M^0)^+ - (n^k_{-,K} - M^0)^+ \right).$$

It is clear that $S_3 \geq 0$ and, by the induction hypothesis, that

$$S_1 \geq \frac{1}{2\triangle t} \sum_{\pm} \sum_{K \in \mathcal{T}} \mathrm{m}(K) \left((n^k_{\pm,K} - M^0)^+ \right)^2.$$

The term S_2 is the analogue of T_{14}. Following the same ideas as in Step 3, we obtain

$$S_2 \geq \frac{D_0}{\lambda_D^2} \sum_{\pm} (1 \pm p) \sum_{K \in \mathcal{T}} \mathrm{m}(K)(n^k_{0,K} - C_K) \left(\frac{1}{2}((n^k_{\pm,K} - M^0)^+)^2 + M^0 (n^k_{\pm,K} - M^0)^+ \right).$$

But, as $n^k_{0,K} = n^k_{+,K} + n^k_{-,K}$, the negativity of $n^k_{+,K}$ and $n^k_{-,K}$ and the definition of M^0 ensure that $n^k_{0,K} - C_K \geq n^k_{+,K} - M^0$ and $n^k_{0,K} - C_K \geq n^k_{-,K} - M^0$, leading to $S_2 \geq 0$. Therefore, we infer that

$$\sum_{\pm} \sum_{K \in \mathcal{T}} \mathrm{m}(K) \left((n^k_{\pm,K} - M^0)^+ \right)^2 \leq 0,$$

which yields the expected result.

4. Proof of Theorem 2

Let $(n^k_{\pm,\mathcal{T}}, V^k_{\mathcal{T}})_{k \geq 0}$ be a solution to (22), (55) with the corresponding Dirichlet-Neumann boundary conditions. Since we have to deal with the logarithm of the densities $n^k_{\pm,K}$, which may vanish, we introduce a regularization of the discrete free energy. For $\delta > 0$, we set $n^{k,\delta}_{\pm,K} = n^k_{\pm,K} + \delta$ and define

$$E^k_\delta = \sum_{\pm} \sum_{K \in \mathcal{T}} \mathrm{m}(K) \left(n^{k,\delta}_{\pm,K} (\log(n^{k,\delta}_{\pm,K}) - 1) - n^k_{\pm,K} \log\left(\frac{n^D_K}{2} + \delta \right) + \frac{n^D_K}{2} \right) + \frac{\lambda_D^2}{2} \sum_{\sigma \in \mathcal{E}} \tau_\sigma (\mathrm{D}(V^k - V^D)_{K,\sigma})^2. \tag{56}$$

Therefore, we have $E^k_\delta - E^{k-1}_\delta = U_1 + U_2$, where

$$U_1 = \sum_{\pm} \sum_{K \in \mathcal{T}} \mathrm{m}(K) \Bigg(n^{k,\delta}_{\pm,K} (\log n^{k,\delta}_{\pm,K} - 1) - n^{k-1,\delta}_{\pm,K} (\log n^{k-1,\delta}_{\pm,K} - 1) - (n^k_{\pm,K} - n^{k-1}_{\pm,K}) \log\left(\frac{n^D_K}{2} + \delta \right) \Bigg),$$

$$U_2 = \frac{\lambda_D^2}{2} \sum_{\sigma \in \mathcal{E}} \tau_\sigma \big((\mathrm{D}(V^k - V^D)_{K,\sigma})^2 - (\mathrm{D}(V^{k-1} - V^D)_{K,\sigma})^2 \big).$$

The convexity of $x \mapsto x(\log x - 1)$ shows that $x(\log x - 1) - y(\log y - 1) \le (x-y)\log x$ for all x, $y > 0$. Hence,

$$U_1 \le \sum_{\pm} \sum_{K\in\mathcal{T}} \mathrm{m}(K)(n_{\pm,K}^k - n_{\pm,K}^{k-1}) \left(\log n_{\pm,K}^{k,\delta} - \log\left(\frac{n_K^D}{2} + \delta\right)\right).$$

Using the elementary inequality $\frac{1}{2}(x^2 - y^2) \le (x-y)x$ for all x, $y \in \mathbb{R}$, integrating by parts, and employing the discrete Poisson equation (22), it follows that

$$\begin{aligned} U_2 &\le \lambda_D^2 \sum_{\sigma\in\mathcal{E}} \tau_\sigma \mathrm{D}(V^k - V^{k-1})_{K,\sigma} \mathrm{D}(V^k - V^D)_{K,\sigma} \\ &= -\lambda_D^2 \sum_{K\in\mathcal{T}} \sum_{\sigma\in\mathcal{E}_K} \tau_\sigma \mathrm{D}(V^k - V^{k-1})_{K,\sigma}(V_K^k - V_K^D) \\ &= \sum_{\pm} \sum_{K\in\mathcal{T}} \mathrm{m}(K)\big(n_{\pm,K}^k - n_{\pm,K}^{k-1}\big)(V_K^k - V_K^D). \end{aligned}$$

We summarize the above inequalities and use scheme (55) to find that

$$\begin{aligned} \frac{1}{\triangle t}(E_\delta^k - E_\delta^{k-1}) &\le -\frac{1}{\tau} \sum_{K\in\mathcal{T}} \mathrm{m}(K)(n_{+,K}^k - n_{-,K}^k)\big(\log n_{+,K}^{k,\delta} - \log n_{-,K}^{k,\delta}\big) \\ &\quad - \sum_{\pm} D_0(1\pm p) \sum_{K\in\mathcal{T}} \sum_{\sigma\in\mathcal{E}_K} J_{\pm,K,\sigma}^k \left(\log n_{\pm,K}^{k,\delta} + V_K^k - \log\left(\frac{n_K^D}{2} + \delta\right) - V_K^D \right). \end{aligned}$$

The first term on the right-hand side is clearly nonpositive. We apply the discrete integration-by-parts formula (18) to the second term. Then, with the hypothesis on the boundary data (i.e. $\log(n^D/2) + V^D$ is constant in $\overline{\Omega}$ such that $\mathrm{D}V_{K,\sigma}^D = -\mathrm{D}(\log n^D)_{K,\sigma}$ for all $K \in \mathcal{T}$ and $\sigma \in \mathcal{E}_K$), we infer that

$$\begin{aligned} \frac{1}{\triangle t}(E_\delta^k - E_\delta^{k-1}) &\le \sum_{\pm} D_0(1\pm p) \sum_{\sigma\in\mathcal{E}} J_{\pm,K,\sigma}^k \mathrm{D}(\log n_\pm^{k,\delta} + V^k)_{K,\sigma} \\ &\quad + \sum_{\pm} D_0(1\pm p) \sum_{\sigma\in\mathcal{E}} J_{\pm,K,\sigma}^k \mathrm{D}\big(\log n^D - \log(n^D + 2\delta)\big)_{K,\sigma}. \end{aligned}$$

Introducing the numerical fluxes associated to the regularized densities,

$$J_{\pm,K,\sigma}^{k,\delta} = \tau_\sigma\big(B(\mathrm{D}V_{K,\sigma}^k) n_{\pm,K}^{k,\delta} - B(-\mathrm{D}V_{K,\sigma}^k) n_{\pm,K,\sigma}^{k,\delta}\big) = J_{\pm,K,\sigma}^k - \delta\tau_\sigma \mathrm{D}V_{K,\sigma}^k,$$

we can write

$$\frac{1}{\triangle t}(E_\delta^k - E_\delta^{k-1}) \le U_3 + U_4 + U_5 + U_6,$$

where

$$\begin{aligned} U_3 &= \sum_{\pm} D_0(1\pm p) \sum_{\sigma\in\mathcal{E}} J_{\pm,K,\sigma}^{k,\delta} \mathrm{D}\big(\log n_\pm^{k,\delta} + V^k\big)_{K,\sigma}, \\ U_4 &= \sum_{\pm} D_0(1\pm p) \sum_{\sigma\in\mathcal{E}} J_{\pm,K,\sigma}^{k,\delta} \mathrm{D}(\log n^D - \log(n^D + 2\delta))_{K,\sigma}, \end{aligned}$$

$$U_5 = \delta \sum_{\pm} D_0(1 \pm p) \sum_{\sigma \in \mathcal{E}} \tau_\sigma \mathrm{D}V_{K,\sigma}^k \mathrm{D}(\log n_\pm^{k,\delta} + V^k)_{K,\sigma},$$
$$U_6 = \delta \sum_{\pm} D_0(1 \pm p) \sum_{\sigma \in \mathcal{E}} \tau_\sigma \mathrm{D}V_{K,\sigma}^k \mathrm{D}\big(\log n^D - \log(n^D + 2\delta)\big)_{K,\sigma}.$$

Now, we employ the following inequalities, which are proved in [4, Appendix A]:

$$J_{\pm,K,\sigma}^{k,\delta} \mathrm{D}(\log n_\pm^{k,\delta} + V^k)_{K,\sigma} \le -\tau_\sigma \min(n_{\pm,K}^{k,\delta}, n_{\pm,K,\sigma}^{k,\delta})(\mathrm{D}(\log n_\pm^{k,\delta} + V^k)_{K,\sigma})^2,$$
$$\big|J_{\pm,K,\sigma}^{k,\delta}\big| \le \tau_\sigma \max(n_{\pm,K}^{k,\delta}, n_{\pm,K,\sigma}^{k,\delta})\big|\mathrm{D}(\log n_\pm^{k,\delta} + V^k)_{K,\sigma}\big|.$$

The first inequality yields

$$U_3 \le -\sum_{\pm} D_0(1 \pm p) \sum_{\sigma \in \mathcal{E}} \tau_\sigma \min(n_{\pm,K}^{k,\delta}, n_{\pm,K,\sigma}^{k,\delta})\big(\mathrm{D}(\log n_\pm^{k,\delta} + V^k)_{K,\sigma}\big)^2,$$

while the second one, together with Young's inequality, gives $U_4 \le U_{41} + U_{42}$, where

$$U_{41} = \frac{1}{4} \sum_{\pm} D_0(1 \pm p) \sum_{\sigma \in \mathcal{E}} \tau_\sigma \min(n_{\pm,K}^{k,\delta}, n_{\pm,K,\sigma}^{k,\delta})\big(\mathrm{D}(\log n_\pm^{k,\delta} + V^k)_{K,\sigma}\big)^2,$$
$$U_{42} = \sum_{\pm} D_0(1 \pm p) \sum_{\sigma \in \mathcal{E}} \tau_\sigma \big(\max(n_{\pm,K}^{k,\delta}, n_{\pm,K,\sigma}^{k,\delta})\big)^2 \frac{(\mathrm{D}(\log n^D - \log(n^D + 2\delta))_{K,\sigma})^2}{\min(n_{\pm,K}^{k,\delta}, n_{\pm,K,\sigma}^{k,\delta})},$$
$$\le 2D_0 \frac{(M^0 + \delta)^2}{\delta} \big|\log(n_{\mathcal{M}}^D + 2\delta) - \log n_{\mathcal{M}}^D\big|_{1,\mathcal{M}}^2,$$

since $\min(n_{\pm,K}^{k,\delta}, n_{\pm,K,\sigma}^{k,\delta}) \ge \delta$ for all $K \in \mathcal{T}$ and $\sigma \in \mathcal{E}_K$. Applying Young's inequality again, we obtain $U_5 \le U_{51} + U_{52}$ with

$$U_{51} = \frac{1}{4} \sum_{\pm} D_0(1 \pm p) \sum_{\sigma \in \mathcal{E}} \tau_\sigma \min(n_{\pm,K}^{k,\delta}, n_{\pm,K,\sigma}^{k,\delta})\big(\mathrm{D}(\log n_\pm^{k,\delta} + V^k)_{K,\sigma}\big)^2,$$
$$U_{52} = \delta^2 \sum_{\pm} D_0(1 \pm p) \sum_{\sigma \in \mathcal{E}} \tau_\sigma \frac{(\mathrm{D}V_{K,\sigma}^k)^2}{\min(n_{\pm,K}^{k,\delta}, n_{\pm,K,\sigma}^{k,\delta})} \le 2D\delta \sum_{\sigma \in \mathcal{E}} \tau_\sigma (\mathrm{D}V_{K,\sigma}^k)^2$$

and

$$U_6 \le 2D_0\delta \left(\big|\log(n_{\mathcal{M}}^D + 2\delta) - \log n_{\mathcal{M}}^D\big|_{1,\mathcal{M}}^2 + \sum_{\sigma \in \mathcal{E}} \tau_\sigma (\mathrm{D}V_{K,\sigma}^k)^2 \right).$$

Summarizing the above inequalities, we deduce that

$$\frac{1}{\triangle t}(E_\delta^k - E_\delta^{k-1}) + \frac{1}{2} \sum_{\pm} D_0(1 \pm p) \sum_{\sigma \in \mathcal{E}} \tau_\sigma \min(n_{\pm,K}^{k,\delta}, n_{\pm,K,\sigma}^{k,\delta})\big(\mathrm{D}(\log n_\pm^{k,\delta} + V^k)_{K,\sigma}\big)^2$$
$$\le 4D_0\delta \sum_{\sigma \in \mathcal{E}} \tau_\sigma (\mathrm{D}V_{K,\sigma}^k)^2 + 2D_0 \left(\delta + \frac{(M^0 + \delta)^2}{\delta}\right) \big|\log(n_{\mathcal{M}}^D + 2\delta) - \log n_{\mathcal{M}}^D\big|_{1,\mathcal{M}}^2. \tag{57}$$

On the one hand, the term $\sum_{\sigma \in \mathcal{E}} \tau_\sigma (\mathrm{D}V_{K,\sigma}^k)^2$ does not depend on δ and is bounded (this can be seen by using scheme (22) and the L^∞ bound on $n_{\pm,\mathcal{T}}^k$). On the other hand, we

rewrite

$$\begin{aligned}\big|\log(n_{\mathcal{M}}^D+2\delta)-\log n_{\mathcal{M}}^D\big|_{1,\mathcal{M}}^2 &= \left|\log\left(1+\frac{2\delta}{n_{\mathcal{M}}^D}\right)\right|_{1,\mathcal{M}}^2 \\ &= \sum_{\sigma\in\mathcal{E}}\tau_\sigma\left(\log\left(1+\frac{2\delta}{n_{K,\sigma}^D}\right)-\log\left(1+\frac{2\delta}{n_K^D}\right)\right)^2.\end{aligned}$$

Employing the inequality $|\log y-\log x|\le|y-x|/\min(x,y)$ for x, $y>0$, and the fact that $n^D\ge n_*>0$, we obtain

$$\big|\log(n_{\mathcal{M}}^D+2\delta)-\log n_{\mathcal{M}}^D\big|_{1,\mathcal{M}}^2\le\frac{4\delta^2}{n_*^2}|n_{\mathcal{M}}^D|_{1,\mathcal{M}}^2.$$

Thanks to hypothesis (33), $n^D\in H^1(\Omega)$, and Lemma 9.4 in [7], we conclude that $|n_{\mathcal{M}}^D|_{1,\mathcal{M}}\le K\|n^D\|_{H^1(\Omega)}$ with K depending only on the regularity of the mesh $\mathcal{M}$. Therefore, the right-hand side in (57) tends to zero when $\delta\to0$. Passing to the limit $\delta\to0$ in (57) then leads to (38). This concludes the proof of Theorem 2.

5. Numerical simulations

As an illustration of the numerical scheme, analyzed in the previous sections, we present two-dimensional simulations of a simple double-gate ferromagnetic MESFET (metal semiconductor field-effect transistor). This device is composed of a semiconductor region which is sandwiched between two ferromagnetic contact regions (see Figure 1). The idea of such devices is that the source region plays the role of a spin polarizer. The non-zero spin-orbit interaction causes the electrons to precess during the propagation through the middle channel region. At the drain contact, only those electrons with spin aligned to the drain magnetization can leave the channel and contribute to the current flow. Here, we focus on the feasibility of our numerical scheme and the verification of the properties of the numerical solution and less on the physical properties. Therefore, the physical setting considered here is strongly simplified. In particular, we just modify the standard MESFET setup by allowing for ferromagnetic regions. For a more detailed modeling, we refer e.g. to [17].

In the following, we describe the geometry of the device in the (x,y) plane (see Figure 1). The total length is $L=0.6\,\mu$m and the height equals $H=0.2\,\mu$m. The source and drain regions are highly doped with doping $C_+=3\cdot10^{23}\mathrm{m}^{-3}$. The doping in the channel region is $C_0=10^{23}\mathrm{m}^{-3}$. The length of the source and drain regions are $\ell=0.1\,\mu$m. The gate contacts are attached at the middle of the device with a length of $L_G=0.2\,\mu$m.

The values of the physical parameters are given in Table 5. They are similar to those used in [13] (there is a small difference in the relaxation time value). The (squared) scaled Debye length becomes $\lambda_D^2=\varepsilon_0\varepsilon_rU_T/(q_eC_+L^2)\approx1.6\cdot10^{-4}$. Note that condition (36) on τ is not satisfied with these physical values but it turns out that the numerical solution is still bounded (even uniformly in time). This may indicate that condition (36) on τ is technical only.

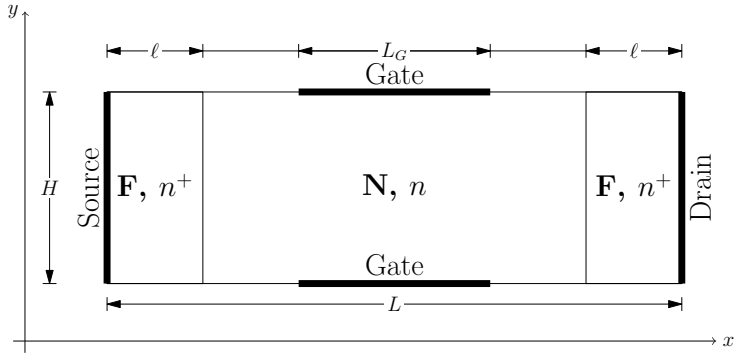

FIGURE 1. Geometry of a MESFET with ferromagnetic (F) source and drain regions and nonmagnetic (N) channel region.

Name	Description	Value
D^*	Diffusion coefficient	$10^{-3}\,\mathrm{m^2s^{-1}}$
ε_r	Relative permittivity of silicon	11.7
ε_0	Permittivity of the vacuum	$8.9 \cdot 10^{-12}\,\mathrm{Fm^{-1}}$
q_e	Elementary charge	$1.6 \cdot 10^{-10}\,\mathrm{C}$
τ^*	Spin-flip relaxation time	$10^{-12}\,\mathrm{s}$
U_T	Thermal voltage at room temperature	$0.0026\,\mathrm{V}$

The gate contact is considered as a Schottky contact. Usually, Robin-type boundary conditions are prescribed at a Schottky contact but also Dirichlet conditions involving the Schottky barrier height have been used to simplify the modeling [21, Section 5.1]. This simplification is possible for Schottky contacts on n-doped materials as it is the case here. We choose the barrier potential $V_S = 0.8\,\mathrm{V}$. The total voltage between source and gate is $V_G + V_S$, where V_G is the voltage applied at the gate. The density boundary value at the gate contact is calclulated according to [21, Formula (5.1-19)], and the potential of the closed state is taken from [13]. This gives

- at the source: $n_0 = C_+$, $\vec{n} = 0$, potential: $0\,\mathrm{V}$,
- at the drain: $n_0 = C_+$, $\vec{n} = 0$, potential: V_D,
- at the gate:
  open state: $n_0 = 3.9 \cdot 10^{11}\,\mathrm{m^{-3}}$, $\vec{n} = 0$, potential: V_S,
  closed state: $n_0 = 3.2 \cdot 10^{9}\,\mathrm{m^{-3}}$, $\vec{n} = 0$, potential: $V_S + 1.2\,\mathrm{V}$,
- for the other segments: homogeneous Neumann boundary conditions.

The magnetic field is caused by the local orientation of the electron spin in the crystal and is predetermined by the ferromagnetic properties of the material. We consider a constant magnetic field, oriented along the z-axis (perpendicular to the device). The electron spin may be also changed under the influence of the spin current, but we do not consider this effect here. In our model, $\vec{m}$ corresponds to the direction of the local magnetic field, and the parameter γ describes the intensity of the spin precession around this field. We choose

$\vec{m} = 0$ in the channel region and

$$\vec{m} = \begin{cases} (0,0,1) & \text{for } x < L/3 \text{ or } x \geq 2L/3, \\ (0,0,0) & \text{for } L/3 \leq x < 2L/3. \end{cases}$$

The value for γ is taken from [19], i.e. $\gamma = \hbar/\tau$, with $\hbar$ being the reduced Planck constant. The spin polarization is nonzero only in the highly doped source and drain regions, and we take $p = 0.9$.

For the numerical discretization, we have chosen an admissible triangular mesh. Equations (4)-(8) are approximated by scheme (20)-(25), with the corresponding boundary conditions. The nonlinear system is solved at each time step by Newton's method. The time step size is $\triangle t = 0.05$. The computations are continued until a steady state is reached or, more precisely, until the difference of the solutions at two consecutive time steps in the ℓ^2 norm falls below a threshold (typically, 10^{-5}).

Simulations for a one-dimensional multilayer structure were presented in [16]. Here, we consider also a multi-layer structure but for a simple MESFET model. Furthermore, the authors of [16] employed a standard finite-volume discretization together with a Gummel iteration method, while we employed here a Scharfetter-Gummel discretization, which is better adapted to large electric fields than a standard technique, and a full Newton method.

Figure 2 illustrates the scaled steady-state charge density n_0 and the spin density n_3 (note that $n_1 = n_2 = 0$) in the open state. The densities are scaled by the doping concentration C_+, the spatial variable by the device length L. Compared to the closed state in Figure 3, the charge density is rather large in the channel region, which can be also observed in standard MESFET devices. The charge current density in the closed state is by the factor $10^5 \ldots 10^6$ smaller than in the open state. The spin density is (almost) zero in the closed state. Furthermore, the electrostatic potential for the open- and closed-state MESFET from Figures 2 and 3 is presented in Figure 4.

Current-voltage characteristics for MESFETs with and without ferromagnetic regions are shown in Figure 5. We observe that in the open state (nonpositive gate potentials), the current densities in the ferromagnetic MESFET are slightly larger than in the standard device, which allows for an improved device performance. When the transistor is closed ($V_G = 1.2\,\mathrm{V}$), the current densities are (almost) zero for both transistor types.

In the left panel of Figure 6, we present the transient behavior of the charge density when switching from the open to the closed state ($V_D = -2\,\mathrm{V}$; dotted line). The current values stabilize after about 1 ps. This justifies to define the numerical solution after 12 ps as the "steady-state solution". We compare these values with those computed from a standard MESFET (solid line). The stabilization in the ferromagnetic case is slightly faster which allows for faster devices.

Finally, we illustrate the free energy decay in Figure 6 (right) for various relaxation times τ. In this experiment, we have set $V_D = 0$ (source-drain voltage) and $V_G = 0$ (source-gate voltage). It turns out that the free energy decays with an exponential rate. For times larger than about 18 ps, the steady state is almost attained, and the numerical oscillations are caused by the finite machine precision. We observe that the decay is faster for smaller relaxation times which is expected. The decay rates are approximately 0.2/ps

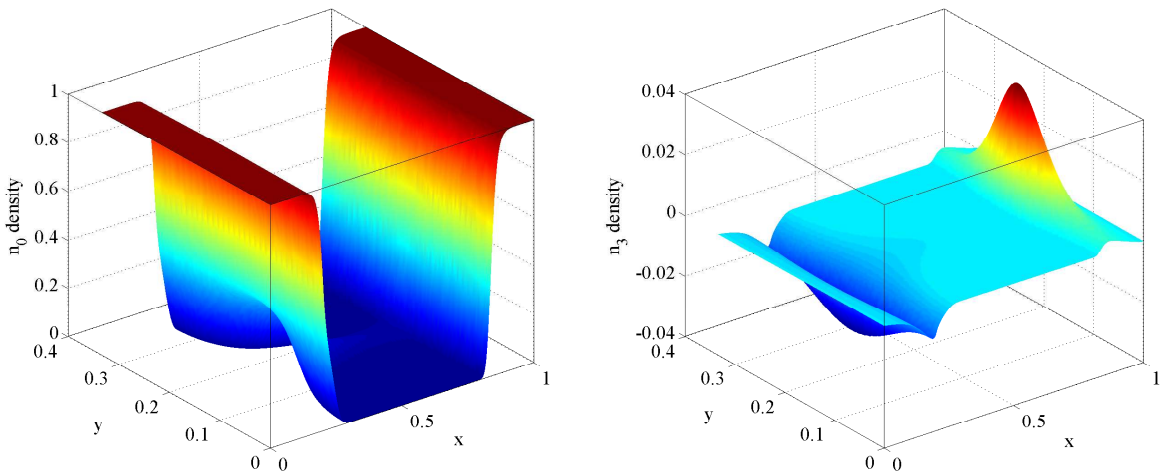

FIGURE 2. Scaled stationary charge density (left) and spin density n_3 (right) in an open-state MESFET with $V_D = -2\,\mathrm{V}$ and $V_G = 0\,\mathrm{V}$.

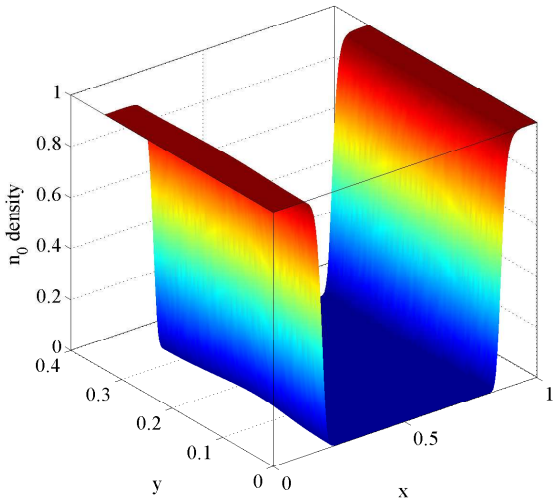

FIGURE 3. Scaled stationary charge density in a closed-state MESFET with $V_D = -2\,\mathrm{V}$ and $V_G = 1.2\,\mathrm{V}$.

for $\tau_s = 100\,\mathrm{ps}$, 0.4/ps for $\tau_s = 10\,\mathrm{ps}$, and 1.7/ps for $\tau_s = 1\,\mathrm{ps}$. The nonlinear dependence of the decay rates on τ_s may be caused by the influence of the energy dissipation.

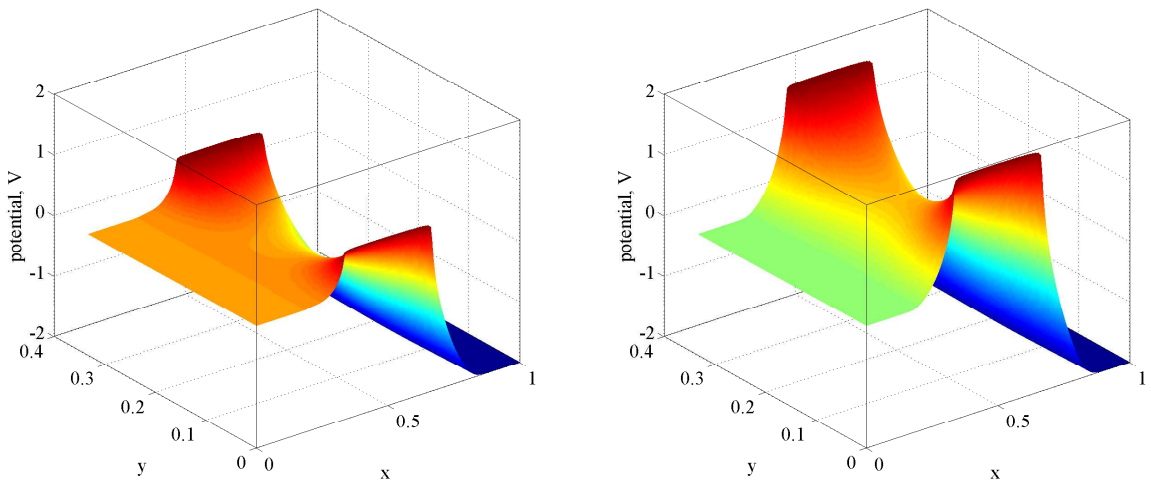

FIGURE 4. Electrostatic potential in a MESFET with $V_G = 0\,\mathrm{V}$ (open state; left) and $V_G = 1.2\,\mathrm{V}$ (closed state; right).

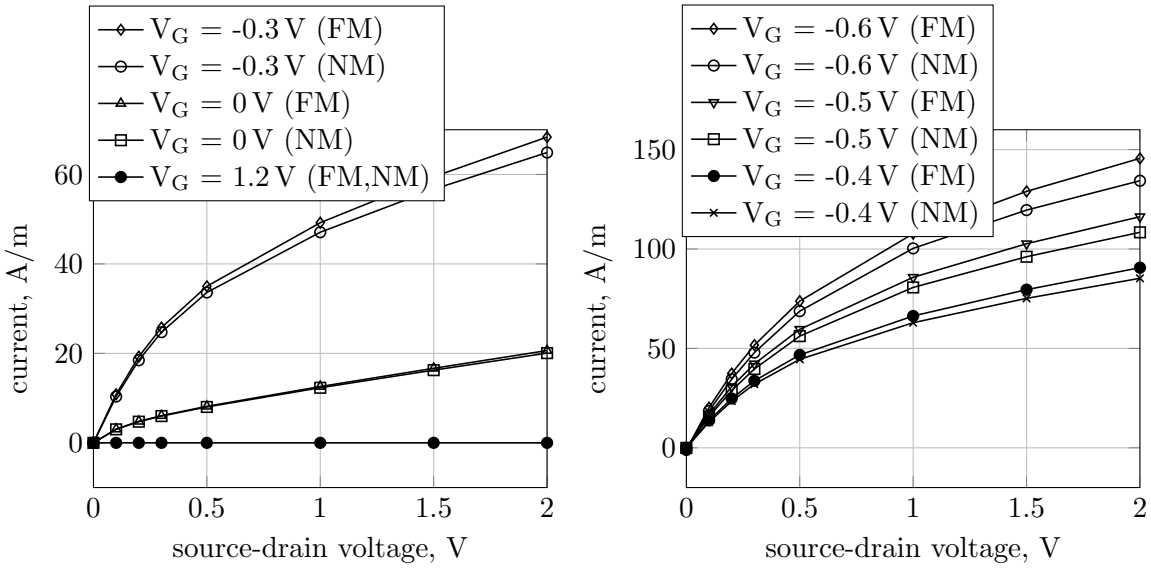

FIGURE 5. Current-voltage characteristics for the ferromagnetic (FM) and standard (NM) MESFET for various gate voltages V_G. For convenience, the source-drain voltages are given by their absolute values.

REFERENCES

[1] C. Abert, G. Hrkac, M. Page, D. Praetorius, M. Ruggeri, and D. Suess. Spin-polarized transport in ferromagnetic multilayers: An unconditionally convergent FEM integrator. *Computers Math. Appl.* 68 (2014), 639-654.

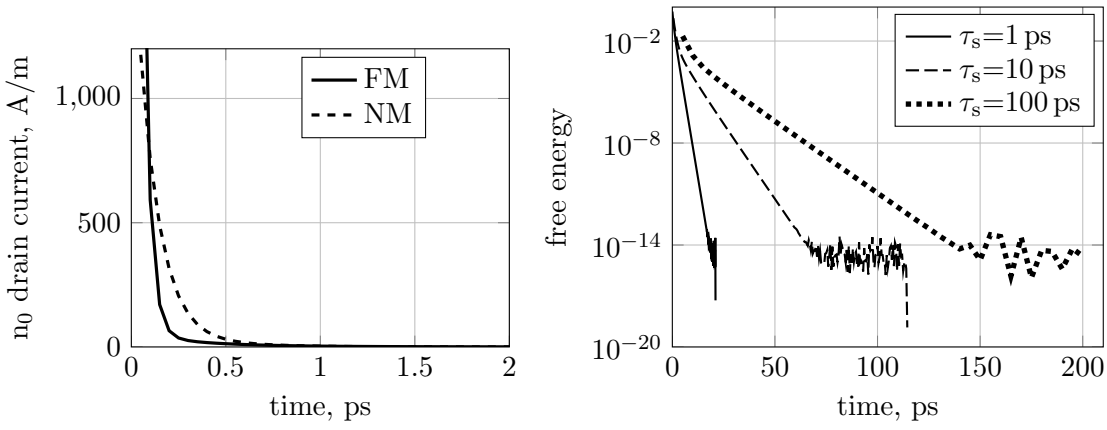

FIGURE 6. Left: Change of the electron drain current in the ferromagnetic (FM) and standard MESFET (NM), switching from open to closed state. Right: Semilogarithmic plot of the free energy versus time for various relaxation times.

[2] C. Abert, M. Ruggeri, F. Bruckner, C. Vogler, G. Hrkac, D. Praetorius, and D. Suess. Self-consistent micromagnetic simulations including spin-diffusion effects. Preprint, 2014. `arXiv:1410:6067`.

[3] L. Barletti and F. Méhats. Quantum drift-diffusion modeling of spin transport in nanostructures. *J. Math. Phys.* 51 (2010), 053304, 20 pages.

[4] M. Bessemoulin-Chatard, C. Chainais-Hillairet, and M.-H. Vignal. Study of a finite volume scheme for the drift-diffusion system. Asymptotic behavior in the quasi-neutral regime. *SIAM J. Numer. Anal.* 52-4 (2014), 1666-1691.

[5] M. Chatard. Asymptotic behavior of the Scharfetter-Gummel scheme for the drift-diffusion model. Proceedings of the conference "Finite Volumes for Complex Applications. VI. Problems and Perspectives". *Springer Proc. Math.* 4 (2011), 235-243.

[6] R. El Hajj. Diffusion models for spin transport derived from the spinor Boltzmann equation. *Commun. Math. Sci.* 12 (2014), 565-592.

[7] R. Eymard, T. Gallouët, and R. Herbin. Finite volume methods. In: P. G. Ciarlet and J. L. Lions (eds.). *Handbook of Numerical Analysis*, Vol. 7. North-Holland, Amsterdam (2000), 713-1020.

[8] J. Fabian, A. Matos-Abiague, C. Ertler, P. Stano, and I. Žutić. Semiconductor spintronics. *Acta Phys. Slovava* 57 (2007), 565-907.

[9] H. Gajewski and K. Gärtner. On the discretization of van Roosbroeck's equations with magnetic field. *Z. Angew. Math. Mech.* 76 (1996), 247-264.

[10] K. Gärtner and A. Glitzky. Existence of bounded steady state solutions to spin-polarized drift-diffusion systems. *SIAM J. Math. Anal.* 41 (2010), 2489-2513.

[11] A. Glitzky. Analysis of a spin-polarized drift-diffusion model. *Adv. Math. Sci. Appl.* 18 (2008), 401-427.

[12] A. Glitzky. Uniform exponential decay of the free energy for Voronoi finite volume discretized reaction-diffusion systems. *Math. Nachr.* 284 (2011), 2159-2174.

[13] S. Holst, A. Jüngel and P. Pietra. An adaptive mixed scheme for energy-transport simulations of field-effect transistors. *SIAM J. Sci. Comp.* 25 (2004), 1698-1716.

[14] A.M. Il'in. A difference scheme for a differential equation with a small parameter affecting the highest derivative. *Mat. Zametki* 6 (1969), 237-248 (in Russian).

[15] A. Jüngel. *Transport Equations for Semiconductors.* Lecture Notes in Physics 773. Springer, Berlin, 2009.

[16] A. Jüngel, C. Negulescu, and P. Shpartko. Bounded weak solutions to a matrix drift-diffusion model for spin-coherent electron transport in semiconductors. To appear in *Math. Models Meth. Appl. Sci.*, 2015.

[17] T. Low, M. Lundstrom, and D. Nikonov. Modeling of spin metal-oxide-semiconductor field-effect transistor: A nonequilibrium Green's function approach with spin relaxation. *J. Appl. Phys.* 104 (2008), 094511, 10 pages.

[18] Y. Pershin, S. Saikin, and V. Privman. Semiclassical transport models for semiconductor spintronics. *Electrochem. Soc. Proc.* 2004-13 (2005), 183-205.

[19] S. Possanner and C. Negulescu. Diffusion limit of a generalized matrix Boltzmann equation for spin-polarized transport. *Kinetic Related Models* 4 (2011), 1159-1191.

[20] D.I. Scharfetter and H.K. Gummel, Large-signal analysis of a silicon Read diode oscillator. *IEEE Trans. Electron Dev.* ED-16 (1969), 64-77.

[21] S. Selberherr. *Analysis and Simulation of Semiconductor Devices.* Springer, Vienna, 1984.

[22] G. Troianiello. *Elliptic Differential Equations and Obstacle Problems.* Plenum Press, New York, 1987.

[23] I. Žutić, J. Fabian, and S. Das Sarma. Spin-polarized transport in inhomogeneous magnetic semiconductors: theory of magnetic/nonmagnetic p-n junctions. *Phys. Rev. Lett.* 88 (2002), 066603.

[24] I. Žutić, J. Fabian, and S. Das Sarma. Spintronics: Fundamentals and applications. *Rev. Modern Phys.* 76 (2004), 323-410.

Laboratoire Paul Painlevé, Université Lille 1 Sciences et Technologies, Cité Scientifique, 59655 Villeneuve d'Ascq Cedex, France
E-mail address: `Claire.Chainais@math.univ-lille1.fr`

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

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