

# Analysis of numerical schemes for semiconductors energy-transport models

Marianne Bessemoulin-Chatard, Claire Chainais-Hillairet, Hélène Mathis

▶ **To cite this version:**

Marianne Bessemoulin-Chatard, Claire Chainais-Hillairet, Hélène Mathis. Analysis of numerical schemes for semiconductors energy-transport models. Applied Numerical Mathematics, 2021, 168, pp.143-169. 10.1016/j.apnum.2021.05.030 . hal-02940224v2

**HAL Id: hal-02940224**
**https://hal.science/hal-02940224v2**

Submitted on 10 Jun 2021

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

# ANALYSIS OF NUMERICAL SCHEMES FOR SEMICONDUCTOR ENERGY-TRANSPORT MODELS

MARIANNE BESSEMOULIN-CHATARD, CLAIRE CHAINAIS-HILLAIRET, AND HÉLÈNE MATHIS

Abstract. Some finite volume schemes for unipolar energy-transport models are introduced. Using a reformulation in dual entropy variables, we can show the decay of a discrete entropy with control of the discrete entropy dissipation. We establish *a priori* estimates which lead to the existence of a solution to the scheme. Similarly to the continuous framework, we prove the exponential decay of the discrete relative entropy towards the thermal equilibrium. Numerical results assess the good behaviour of the whole numerical scheme.

**Key-words.** Energy-transport model, finite volume schemes, entropy method.

**2020 MCS.** 65M08, 65M12, 35K20

Contents

1. Introduction 2
1.1. The energy-transport models 2
1.2. Entropy structure 4
1.3. Goal and outline of the paper 5
2. Numerical schemes 5
2.1. Mesh notations 5
2.2. Schemes in primal and dual entropy variables 6
2.3. Equivalence of the schemes in the primal and dual entropy variables 8
3. Discrete entropy inequality and *a priori* estimates 10
3.1. Discrete entropy inequality 10
3.2. *A priori* estimates on the approximate solutions 12
4. Existence of a solution to the schemes 14
5. Long time behavior 16
5.1. Definition of the thermal equilibrium 16
5.2. Exponential decay towards equilibrium 17
6. Numerical experiments 19
6.1. Physical models 19
6.2. Implementation of the whole method 20
6.3. Test cases 20
7. Conclusion 27
References 28

MBC and HM were partially funded by the Centre Henri Lebesgue (ANR-11-LABX-0020-01) and ANR Project MoHyCon (ANR-17-CE40-0027-01). CCH was partially funded by ANR Project MoHyCon (ANR-17-CE40-0027-01) and was supported by the ANR Labex CEMPI grant ANR-11-LABX-0007-01.

## 1. Introduction

In this article, we are interested in the numerical analysis of unipolar energy-transport models for semiconductor devices. Such models describe the flow of electrons through a semiconductor crystal, influenced by diffusive, electrical and thermal effects. Contrary to classical drift-diffusion systems (see [34, 21, 22]), they take into account the temperature and permit to observe hot electrons effects. As recalled in [29], these models can be derived either from hydrodynamic models by neglecting some convection terms [23] or from the Boltzmann equation by the moment method [2]. There exists several other models which allow to take into account thermal effects. For instance non-isothermal drift-diffusion system, considered in [32], couples the drift-diffusion semiconductor equations to the heat flow equation for the temperature distribution in the device. We refer to [32] and references therein for details about this approach. In the sequel we focus on the energy-transport model with cross-diffusion structure, obtained while considering another moment of the Boltzmann equation [29].

### 1.1. The energy-transport models.

The primal energy-transport system consists in two continuity equations for the electron density $\rho_1$ and the internal energy density $\rho_2$, coupled with a Poisson equation for the electrical potential $V$. Following the framework adopted in [14], we consider that the electron and the energy densities are defined as functions of the entropy variables $\mathbf{u} = (u_1, u_2)$ with

$$u_1 = \frac{\mu}{T}, \quad u_2 = -\frac{1}{T}, \tag{1}$$

where $\mu$ denotes the chemical potential and $T$ the temperature. Then the electron density $\rho_1$ and the internal energy density $\rho_2$ are assumed to depend on $\mathbf{u}$. Let $\Omega$ be an open bounded subset of $\mathbb{R}^d$, $d \geq 1$, describing the geometry of the considered semiconductor device and let $T_{max} > 0$ be a finite time horizon. The energy-transport model writes in $\Omega \times (0, T_{max})$

$$\partial_t \rho_1(\mathbf{u}) + \text{div } J_1(\mathbf{u}) = 0, \tag{2}$$

$$\partial_t \rho_2(\mathbf{u}) + \text{div } J_2(\mathbf{u}) = \nabla_x V \cdot J_1(\mathbf{u}) + W(\mathbf{u}), \tag{3}$$

$$-\lambda^2 \Delta V = C(x) - \rho_1(\mathbf{u}), \tag{4}$$

where the electron and energy current densities are given by:

$$\begin{aligned} J_1(\mathbf{u}) &= -L_{11}(\mathbf{u})(\nabla u_1 + u_2 \nabla V) - L_{12}(\mathbf{u}) \nabla u_2, \\ J_2(\mathbf{u}) &= -L_{21}(\mathbf{u})(\nabla u_1 + u_2 \nabla V) - L_{22}(\mathbf{u}) \nabla u_2, \end{aligned} \tag{5}$$

where $\mathbb{L}(\mathbf{u}) = (L_{ij}(\mathbf{u}))_{1 \leq i,j \leq 2}$ is a symmetric uniformly positive definite matrix in the sense that there exists a constant $\alpha > 0$ such that

$$(\mathbb{L}\mathbf{u}, \mathbf{u}) \geq \alpha |\mathbf{u}|^2, \quad \forall \mathbf{u} \in \mathbb{R}^2. \tag{6}$$

The term $\nabla V \cdot J_1$ corresponds to a Joule heating term and $W(\mathbf{u})$ is an energy relaxation term. The doping profile $C \in L^\infty(\Omega)$ describes the fixed charged background and $\lambda$ is the rescaled Debye length.

The system (2)-(5) is supplemented with an initial condition $\mathbf{u}_0 = (u_{1,0}, u_{2,0})$ and with mixed boundary conditions. There are Dirichlet boundary conditions on the ohmic contacts and homogeneous Neumann boundary conditions on insulating

segments. More precisely the domain $\Omega$ is assumed to be an open bounded polygonal (or polyhedral) subset of $\mathbb{R}^d$ ($d \geq 1$) and $\partial\Omega$ is split into $\partial\Omega = \Gamma^D \cup \Gamma^N$, with $\Gamma^D \cap \Gamma^N = \emptyset$ and $\mathrm{m}_{d-1}(\Gamma^D) > 0$. We denote by $\mathbf{n}$ the normal to $\partial\Omega$ outward $\Omega$.

The boundary conditions write

$$u_1 = u_1^D, \quad u_2 = u_2^D, \quad V = V^D, \quad \text{on } \Gamma^D \times (0, T_{max}), \tag{7}$$

$$J_1 \cdot \mathbf{n} = J_2 \cdot \mathbf{n} = \nabla V \cdot \mathbf{n} = 0, \quad \text{on } \Gamma^N \times (0, T_{max}). \tag{8}$$

We assume that the Dirichlet boundary conditions $u_1^D$, $u_2^D$ and $V^D$ do not depend on time and are the traces of some functions defined on the whole domain $\Omega$, still denoted $u_1^D$, $u_2^D$ and $V^D$, and such that

$$u_1^D,\, u_2^D \in H^1(\Omega), \quad V^D \in H^1(\Omega) \cap L^\infty(\Omega). \tag{9}$$

Moreover we assume that

$$u_2^D \text{ is a given negative constant on } \Gamma^D \tag{10}$$

and that the energy relaxation term $W(\mathbf{u})$ verifies, for all $\mathbf{u} \in \mathbb{R}^2$ and $u_2^D < 0$, that

$$W(\mathbf{u})(u_2 - u_2^D) \leq 0. \tag{11}$$

The main results on the energy-transport model (2)-(5) are presented in [14, 28]: existence of weak solutions to the system, regularity, uniqueness and long-time convergence towards thermal equilibrium. They are based on a reformulation of the system in terms of dual entropy variables (electrochemical potentials). This change of variables symmetrizes the equations and allows to derive a crucial entropy–entropy dissipation estimate.

The proofs established in [14] rely on some assumptions on $\boldsymbol{\rho}$ recalled below (Assumptions 1) and on the assumption of uniform definiteness of the diffusion matrix $\mathbb{L}$. This is the framework we will adopt here. However, these assumptions are hardly satisfied by practical test cases, since they require that the electron density $\rho_1$ and the temperature $T$ have positive lower bounds, which is not the case in general. Existence results for physically more realistic diffusion matrices (only positive semi-definite) are established in [18, 24] for the stationary model and in [12, 13] for the transient system, but only in the case of data close to thermal equilibrium. More recently, existence of solutions has been proved in simplified degenerate cases, namely for a model with a simplified temperature equation in [30] and for vanishing electric fields (avoiding the coupling with Poisson equation) in [35].

We consider the following assumptions on the function $\mathbf{u} \mapsto \boldsymbol{\rho}(\mathbf{u}) = (\rho_1(\mathbf{u}), \rho_2(\mathbf{u}))$.

**Assumptions 1.** The function $\boldsymbol{\rho}$ is such that

(i) $\boldsymbol{\rho} \in W^{1,\infty}(\mathbb{R}^2; \mathbb{R}^2)$, in particular there exists $C_\rho > 0$ such that $\|\boldsymbol{\rho}(\mathbf{u})\|_\infty \leq C_\rho$ for all $\mathbf{u} \in \mathbb{R}^2$,

(ii) there exists a constant $C_0$ such that

$$(\boldsymbol{\rho}(\mathbf{u}) - \boldsymbol{\rho}(\mathbf{v})) \cdot (\mathbf{u} - \mathbf{v}) \geq C_0 |\mathbf{u} - \mathbf{v}|^2, \quad \forall \mathbf{u}, \mathbf{v} \in \mathbb{R}^2, \tag{12}$$

(iii) there exists $\chi \in \mathcal{C}^1(\mathbb{R}^2; \mathbb{R})$ convex such that $\boldsymbol{\rho}(\mathbf{u}) = \nabla_{\mathbf{u}} \chi(\mathbf{u})$.

The key point of the analysis of the primal model (2)-(5) is to use another set of variables which symmetrizes the problem, see [14]. Let us define the so-called dual

entropy variables $\mathbf{w} = (w_1, w_2)$ ($w_1$ is actually an eletrochemical potential):

$$\begin{cases} w_1 = u_1 + u_2 V, \\ w_2 = u_2. \end{cases} \tag{13}$$

Through this change of variables, the problem (2)-(5) is equivalent to

$$\partial_t b_1(\mathbf{w}, V) + \text{div } I_1(\mathbf{w}, V) = 0, \tag{14}$$

$$\partial_t b_2(\mathbf{w}, V) + \text{div } I_2(\mathbf{w}, V) = \widetilde{W}(\mathbf{w}) - \partial_t V b_1(\mathbf{w}, V), \tag{15}$$

$$-\lambda^2 \Delta V = C - b_1(\mathbf{w}, V), \tag{16}$$

where the function $\mathbf{b}(\mathbf{w}, V) = (b_1(\mathbf{w}, V), b_2(\mathbf{w}, V))$ is related to $\boldsymbol{\rho}$ and $V$ by

$$\begin{cases} b_1(\mathbf{w}, V) = \rho_1(\mathbf{u}), \\ b_2(\mathbf{w}, V) = \rho_2(\mathbf{u}) - V\rho_1(\mathbf{u}), \end{cases} \tag{17}$$

and the new energy relaxation term is defined by $\widetilde{W}(\mathbf{w}) = W(\mathbf{u})$. The symmetrized currents are given by

$$\begin{aligned} I_1 &= J_1, \\ I_2 &= J_2 - V J_1, \end{aligned} \tag{18}$$

which leads to

$$\begin{aligned} I_1(\mathbf{w}, V) &= -D_{11}(\mathbf{w}, V)\nabla w_1 - D_{12}(\mathbf{w}, V)\nabla w_2, \\ I_2(\mathbf{w}, V) &= -D_{21}(\mathbf{w}, V)\nabla w_1 - D_{22}(\mathbf{w}, V)\nabla w_2. \end{aligned} \tag{19}$$

The new diffusion matrix $\mathbb{D}(\mathbf{w}, V) = (D_{ij}(\mathbf{w}, V))_{1 \le i,j \le 2}$ is defined by

$$\mathbb{D}(\mathbf{w}, V) = \mathbb{P}(V)^T \mathbb{L}(\mathbf{u}) \mathbb{P}(V), \quad \text{with } \mathbb{P}(V) = \begin{pmatrix} 1 & -V \\ 0 & 1 \end{pmatrix}, \tag{20}$$

so that

$$\begin{aligned} D_{11}(\mathbf{w}, V) &= L_{11}(\mathbf{u}), \\ D_{12}(\mathbf{w}, V) &= D_{21}(\mathbf{w}, V) = L_{12}(\mathbf{u}) - V L_{11}(\mathbf{u}), \\ D_{22}(\mathbf{w}, V) &= L_{22}(\mathbf{u}) - 2V L_{12}(\mathbf{u}) + V^2 L_{11}(\mathbf{u}). \end{aligned} \tag{21}$$

It is clear that the diffusion matrix $\mathbb{D}$ is also symmetric and uniformly positive definite.

1.2. **Entropy structure.** The existence results for model (2)-(5) given in [14] are mainly based on entropy estimates obtained thanks to the dual model (14)-(19).

We recall in this section the entropy/entropy-dissipation property. The entropy function is defined by

$$S(t) = \int_\Omega \left[ \boldsymbol{\rho}(\mathbf{u}) \cdot (\mathbf{u} - \mathbf{u}^D) - (\chi(\mathbf{u}) - \chi(\mathbf{u}^D)) \right] dx - \frac{\lambda^2}{2} u_2^D \int_\Omega |\nabla(V - V^D)|^2 dx. \tag{22}$$

Since $u_2^D < 0$ and $\chi$ is a convex function, see Assumption 1-(iii), $S(t)$ is nonnegative for all $t \ge 0$.

Moreover, if the boundary conditions are at thermal equilibrium, namely

$$\nabla w_1^D = \nabla w_2^D = 0, \tag{23}$$

one can prove, under Assumptions 1, that the entropy satisfies the following identity:

$$\frac{d}{dt}S(t) = -\int_\Omega \nabla \mathbf{w}^T \mathbb{D} \nabla \mathbf{w} + \int_\Omega W(\mathbf{u})(u_2 - u_2^D). \tag{24}$$

Thanks to the uniform positivity of the matrix $\mathbb{D}$ and the dissipation of the relaxation term $W$, see hypothesis (11), one deduces that the entropy function decreases along time. The identity (24) gives also *a priori* estimates on $\mathbf{w}$.

More general boundary conditions are considered in [14], but they involve the control of several additional terms.

1.3. **Goal and outline of the paper.** Different kind of numerical schemes have already been designed for the energy–tranport systems, mainly for the stationary systems: finite difference schemes in [19, 33], finite element schemes in [15, 20], mixed finite element schemes in [25, 26]. We also refer to [9] for DDFV (Discrete Duality Finite Volume) schemes for the evolutive case. Up to our knowledge, there exists no analysis of these numerical schemes.

The purpose of this paper is to design some finite volume schemes for the systems (2)–(4) (and (14)–(16)) with a two-point flux approximation (TPFA) of the numerical fluxes. We pay attention while building the scheme on the possibility of adapting the entropy method (briefly described in Section 1.2) to the discrete setting. This will be crucial in order to adapt the proof of [14] to the discrete level. The *a priori* estimates obtained for the approximate solution thanks to the entropy method are sufficient to establish the existence of a solution to the scheme *via* a fixed point theorem.

In Section 2, we introduce the finite volume schemes for the system written in the primal entropy variables (2)–(5) and the system written in the dual entropy variables (14)–(19). We also establish the equivalence between the different schemes. Section 3 is devoted to the proof of a discrete entropy-entropy dissipation estimate, which yields *a priori* estimates on the approximate solutions. In Section 4, we establish the existence of a solution to the schemes and in Section 5, we study the long time behavior of the numerical solutions. Finally, Section 6 is devoted to numerical experiments.

## 2. Numerical schemes

2.1. **Mesh notations.** The mesh of the domain $\Omega$ is given by a family $\mathcal{T}$ of open polygonal (or polyhedral in 3D) control volumes, a family $\mathcal{E}$ of edges (or faces), and a family $\mathcal{P} = (x_K)_{K\in\mathcal{T}}$ of points. As it is classical in the finite volume discretization of diffusive terms with two-point flux approximations, we assume that the mesh is admissible in the sense of [17, Definition 9.1]. It implies that the straight line between two neighboring centers of cells $(x_K, x_L)$ is orthogonal to the edge $\sigma = K|L$.

In the set of edges $\mathcal{E}$, we distinguish the interior edges $\sigma = K|L \in \mathcal{E}_{int}$ and the boundary edges $\sigma \in \mathcal{E}_{ext}$. Due to the mixed boundary conditions, we have to distinguish the Dirichlet boundary edges included in $\Gamma^D$ from the Neumann boundary edges included in $\Gamma^N$: $\mathcal{E}_{ext} = \mathcal{E}_{ext}^D \cup \mathcal{E}_{ext}^N$. For a control volume $K \in \mathcal{T}$, we define $\mathcal{E}_K$ the set of its edges, which is also split into $\mathcal{E}_K = \mathcal{E}_{K,int} \cup \mathcal{E}_{K,ext}^D \cup \mathcal{E}_{K,ext}^N$.

In the sequel, we denote by $d$ the distance in $\mathbb{R}^d$ and m the measure in $\mathbb{R}^d$ or $\mathbb{R}^{d-1}$. For all $\sigma \in \mathcal{E}$, we define $d_\sigma = d(x_K, x_L)$ if $\sigma = K|L \in \mathcal{E}_{int}$ and $d_\sigma = d(x_K, \sigma)$

if $\sigma \in \mathcal{E}_{ext}$, with $\sigma \in \mathcal{E}_K$. Then the transmissibility coefficient is defined by $\tau_\sigma = \mathrm{m}(\sigma)/d_\sigma$, for all $\sigma \in \mathcal{E}$.

We assume that the mesh fulfills the following regularity constraint: there exists $\xi > 0$ such that

$$d(x_K, \sigma) \geq \xi d_\sigma, \quad \forall K \in \mathcal{M}, \forall \sigma \in \mathcal{E}_K. \tag{25}$$

Concerning the time discretization, let $\Delta t > 0$ be the time step. For $T_{max} > 0$, we set $N_T$ the integer part of $T_{max}/\Delta t$ and $t^n = n\Delta t$ for all $n = 0, \dots, N_T$. We denote by $\delta = \max(\Delta t, \mathrm{size}(\mathcal{T}))$ the size of the space–time discretization.

A TPFA finite volume scheme provides, for an unknown $v$, a vector $v_\mathcal{T} = (v_K)_{K\in\mathcal{T}} \in \mathbb{R}^\theta$ (with $\theta = \mathrm{Card}(\mathcal{T})$) of approximate values on each cells. Then, a piecewise constant function, still denoted $v_\mathcal{T}$, can be defined by

$$v_\mathcal{T} = \sum_{K\in\mathcal{T}} v_K \mathbf{1}_K,$$

where $\mathbf{1}_K$ denotes the characteristic function of the cell $K$. However, since there are Dirichlet boundary conditions on a part of the boundary, we also need to define approximate values for $v$ at the corresponding boundary edges: $v_{\mathcal{E}^D} = (v_\sigma)_{\sigma\in\mathcal{E}^D_{ext}} \in \mathbb{R}^{\theta^D}$ (with $\theta^D = \mathrm{Card}(\mathcal{E}^D_{ext})$). Therefore, the vector containing the approximate values both in the control volumes and at the Dirichlet boundary edges is denoted by $v_\mathcal{M} = (v_\mathcal{T}, v_{\mathcal{E}^D})$. We denote by $X(\mathcal{M})$ the set of the discrete functions $v_\mathcal{M}$. For any vector $v_\mathcal{M} = (v_\mathcal{T}, v_{\mathcal{E}^D})$, we define for all $K \in \mathcal{T}$ and all $\sigma \in \mathcal{E}_K$:

$$v_{K,\sigma} = \begin{cases} v_L & \text{if } \sigma = K|L \in \mathcal{E}_{K,int}, \\ v_\sigma & \text{if } \sigma \in \mathcal{E}^D_{K,ext}, \\ v_K & \text{if } \sigma \in \mathcal{E}^N_{K,ext}, \end{cases} \tag{26}$$

and

$$D_{K,\sigma} v_\mathcal{M} = v_{K,\sigma} - v_K, \qquad D_\sigma v_\mathcal{M} = |D_{K,\sigma} v_\mathcal{M}|.$$

Then, we can define the discrete $H^1$-seminorm $|\cdot|_{1,\mathcal{M}}$ on $X(\mathcal{M})$ by

$$|v_\mathcal{M}|^2_{1,\mathcal{M}} = \sum_{\sigma\in\mathcal{E}} \tau_\sigma (D_\sigma v_\mathcal{M})^2, \quad \forall v_\mathcal{M} \in X(\mathcal{M}).$$

We also define the discrete $H^1$-norm $\|\cdot\|_{1,\mathcal{M}}$ on $X(\mathcal{M})$ by

$$\|v_\mathcal{M}\|^2_{1,\mathcal{M}} = \|v_\mathcal{T}\|^2_{L^2(\Omega)} + |v_\mathcal{M}|^2_{1,\mathcal{M}}.$$

2.2. **Schemes in primal and dual entropy variables.** Our aim is to design a scheme for the energy-transport model in the primal entropy variables (2)-(5). This scheme must lead to an equivalent scheme for the system written in the dual entropy variables (14)-(19). Indeed, in this case, it will be possible to apply the entropy method at the discrete level. This step is crucial since it brings *a priori* estimates on the sequences of approximate solutions, which allows to prove existence of a solution to the scheme.

One main difficulty in writing a TPFA scheme for the energy-transport model (2)-(5) comes from the approximation of the Joule heating term $\nabla V \cdot J_1$, see equation (3). One possibility would be to apply the technique developed in [5], and further used in [32, 16], to discretize de Joule heating term. However, with such

discretization, the rewriting of the scheme in dual entropy variables is not straightforward. Therefore, following [7], we propose an approximation of the Joule heating term which is based on its following reformulation:

$$\nabla V \cdot J_1 = \text{div}\ (VJ_1) - V\text{div}\ J_1. \tag{27}$$

Let us now turn to the definition of the scheme for the model (2)-(5). Initial and Dirichlet boundary conditions are discretized as usual by taking the mean values of the data on the cells or on the Dirichlet boundary edges:

$$\left(u^0_{1,K}, u^0_{2,K}\right) = \frac{1}{\text{m}(K)} \int_K \left(u_{1,0}(x), u_{2,0}(x)\right)\, dx, \quad \forall K \in \mathcal{T}, \tag{28}$$

$$\left(u^D_{1,\sigma}, u^D_{2,\sigma}, V^D_\sigma\right) = \frac{1}{\text{m}(\sigma)} \int_\sigma \left(u^D_1(\gamma), u^D_2(\gamma), V^D(\gamma)\right) d\gamma, \quad \forall \sigma \in \mathcal{E}^D_{ext}, \tag{29}$$

and we set

$$u^n_{1,\sigma} = u^D_{1,\sigma}, \quad u^n_{2,\sigma} = u^D_{2,\sigma}, \quad V^n_\sigma = V^D_\sigma, \quad \forall \sigma \in \mathcal{E}^D_{ext}, \quad \forall n \geq 0. \tag{30}$$

We propose a backward Euler in time and finite volume in space scheme with a two-point flux approximation. It reads for all $n \geq 0$, for all $K \in \mathcal{T}$:

$$\text{m}(K)\frac{\rho^{n+1}_{1,K} - \rho^n_{1,K}}{\Delta t} + \sum_{\sigma \in \mathcal{E}_K} \mathcal{F}^{n+1}_{1,K,\sigma} = 0, \tag{31}$$

$$\begin{aligned} \text{m}(K)\frac{\rho^{n+1}_{2,K} - \rho^n_{2,K}}{\Delta t} &+ \sum_{\sigma \in \mathcal{E}_K} \mathcal{F}^{n+1}_{2,K,\sigma} = \sum_{\sigma \in \mathcal{E}_K} V^{n+1}_\sigma \mathcal{F}^{n+1}_{1,K,\sigma} \\ &- V^{n+1}_K \sum_{\sigma \in \mathcal{E}_K} \mathcal{F}^{n+1}_{1,K,\sigma} + \text{m}(K) W^{n+1}_K, \end{aligned} \tag{32}$$

$$-\lambda^2 \sum_{\sigma \in \mathcal{E}_K} \tau_\sigma D_{K,\sigma} V^{n+1}_{\mathcal{M}} = \text{m}(K)(C_K - \rho^{n+1}_{1,K}), \tag{33}$$

where

$$\rho^{n+1}_{i,K} = \rho_i(\mathbf{u}^{n+1}_K), \quad i = 1,\, 2, \quad \text{and } W^{n+1}_K = W(\mathbf{u}^{n+1}_K), \text{for all } K \in \mathcal{T}.$$

The numerical fluxes are given by

$$\begin{aligned} \mathcal{F}^{n+1}_{1,K,\sigma} &= -\tau_\sigma(L^n_{11,\sigma}(D_{K,\sigma}u^{n+1}_{1,\mathcal{M}} + u^{n+1}_{2,\sigma} D_{K,\sigma} V^{n+1}_{\mathcal{M}}) + L^n_{12,\sigma} D_{K,\sigma} u^{n+1}_{2,\mathcal{M}}), \\ \mathcal{F}^{n+1}_{2,K,\sigma} &= -\tau_\sigma(L^n_{12,\sigma}(D_{K,\sigma}u^{n+1}_{1,\mathcal{M}} + u^{n+1}_{2,\sigma} D_{K,\sigma} V^{n+1}_{\mathcal{M}}) + L^n_{22,\sigma} D_{K,\sigma} u^{n+1}_{2,\mathcal{M}}). \end{aligned} \tag{34}$$

The coefficients $L^n_{ij,\sigma}$ are approximations of the coefficients of the matrix $\mathbb{L}$ at the interface $\sigma$. The discrete matrix $\mathbb{L}^n_\sigma = (L^n_{ij,\sigma})_{1 \leq i,j \leq 2}$ is defined as

$$\mathbb{L}^n_\sigma = \mathbb{L}\left(\frac{\mathbf{u}^n_K + \mathbf{u}^n_{K,\sigma}}{2}\right), \quad \text{for all } K \in \mathcal{T},\, \sigma \in \mathcal{E}_K.$$

To complete the definition of the scheme (31)–(34), it remains to define $V^{n+1}_\sigma$ involved in (32) and $u^{n+1}_{2,\sigma}$ involved in (34). These definitions will be given later on, our choice being driven by the expected equivalence with a scheme for (14)–(19).

In order to obtain an equivalent scheme for the energy-transport system in the dual entropy variables (14)–(19), we apply the change of variables (13), associated with the definitions of the functions $b_i(\mathbf{w}, V)$ given in (17) and the definitions of the currents $I_i(\mathbf{w}, V)$ given in (19), into the numerical scheme (31)–(34).

For all $K \in \mathcal{T}$ and for all $n \geq 0$, we set

$$w_{1,K}^n = u_{1,K}^n + u_{2,K}^n V_K^n, \qquad w_{2,K}^n = u_{2,K}^n, \tag{35}$$

$$b_{1,K}^n = \rho_{1,K}^n = b_1(\mathbf{w}_K^n, V_K^n), \qquad b_{2,K}^n = \rho_{2,K}^n - \rho_{1,K}^n V_K^n = b_2(\mathbf{w}_K^n, V_K^n). \tag{36}$$

The boundary conditions are similarly defined:

$$w_{1,\sigma}^n = w_{1,\sigma}^D, \quad w_{2,\sigma}^n = w_{2,\sigma}^D, \quad \forall \sigma \in \mathcal{E}_{ext}^D, \quad \forall n \geq 0. \tag{37}$$

From the numerical scheme (31)–(32), we deduce

$$\begin{aligned}
&\mathrm{m}(K)\frac{b_{1,K}^{n+1} - b_{1,K}^n}{\Delta t} + \sum_{\sigma \in \mathcal{E}_K} \mathcal{F}_{1,K,\sigma}^{n+1} = 0,\\
&\mathrm{m}(K)\frac{b_{2,K}^{n+1} - b_{2,K}^n}{\Delta t} + \sum_{\sigma \in \mathcal{E}_K} \left(\mathcal{F}_{2,K,\sigma}^{n+1} - V_\sigma^{n+1}\mathcal{F}_{1,K,\sigma}^{n+1}\right)\\
&\qquad\qquad = \mathrm{m}(K)\widetilde{W}_K^{n+1} - \mathrm{m}(K)\frac{V_K^{n+1} - V_K^n}{\Delta t} b_{1,K}^n.
\end{aligned}$$

It leads to the following scheme for the system written in the dual entropy variables (14)–(16):

$$\mathrm{m}(K)\frac{b_{1,K}^{n+1} - b_{1,K}^n}{\Delta t} + \sum_{\sigma \in \mathcal{E}_K} \mathcal{G}_{1,K,\sigma}^{n+1} = 0, \tag{38}$$

$$\mathrm{m}(K)\frac{b_{2,K}^{n+1} - b_{2,K}^n}{\Delta t} + \sum_{\sigma \in \mathcal{E}_K} \mathcal{G}_{2,K,\sigma}^{n+1} = \mathrm{m}(K)\widetilde{W}_K^{n+1} - \mathrm{m}(K)\frac{V_K^{n+1} - V_K^n}{\Delta t} b_{1,K}^n, \tag{39}$$

$$-\lambda^2 \sum_{\sigma \in \mathcal{E}_K} \tau_\sigma D_{K,\sigma} V_{\mathcal{M}}^{n+1} = \mathrm{m}(K)(C_K - b_{1,K}^{n+1}), \tag{40}$$

where $\widetilde{W}_K^{n+1} = W_K^{n+1} = \widetilde{W}(\mathbf{w}_K^{n+1}, V_K^{n+1})$ and numerical fluxes are defined by:

$$\begin{aligned}
\mathcal{G}_{1,K,\sigma}^{n+1} &= \mathcal{F}_{1,K,\sigma}^{n+1},\\
\mathcal{G}_{2,K,\sigma}^{n+1} &= \mathcal{F}_{2,K,\sigma}^{n+1} - V_\sigma^{n+1}\mathcal{F}_{1,K,\sigma}^{n+1}.
\end{aligned} \tag{41}$$

The crucial point now is to ensure that the new numerical fluxes $\mathcal{G}_{i,K,\sigma}^{n+1}$, $i = 1, 2$, can be seen as approximations of the currents $I_1$ and $I_2$ defined by (19). This means that we want the numerical fluxes to be written as

$$\begin{aligned}
\mathcal{G}_{1,K,\sigma}^{n+1} &= -\tau_\sigma(\mathrm{D}_{11,\sigma}^* D_{K,\sigma} w_{1,\mathcal{M}}^{n+1} + \mathrm{D}_{12,\sigma}^* D_{K,\sigma} w_{2,\mathcal{M}}^{n+1}),\\
\mathcal{G}_{2,K,\sigma}^{n+1} &= -\tau_\sigma(\mathrm{D}_{21,\sigma}^* D_{K,\sigma} w_{1,\mathcal{M}}^{n+1} + \mathrm{D}_{22,\sigma}^* D_{K,\sigma} w_{2,\mathcal{M}}^{n+1}).
\end{aligned} \tag{42}$$

The coefficients $(\mathrm{D}_{ij,\sigma}^*)_{1 \leq i,j \leq 2}$ are approximations of the coefficients of the matrix $\mathbb{D}(\mathbf{w}, V)$ given by (21) at the interface $\sigma$ and on the time interval $(t^n, t^{n+1}]$. They compose the associate matrix $\mathbb{D}_\sigma^*$ which has to be symmetric and uniformly positive definite. This property depend on the definition of the quantities $V_\sigma^{n+1}$ and $u_{2,\sigma}^{n+1}$, respectively involved in the equations (32) and (34), for each $\sigma \in \mathcal{E}$.

2.3. **Equivalence of the schemes in the primal and dual entropy variables.** Observe that the scheme is not a fully backward Euler discretization in time since the last term in the equation (39) involves the quantity $b_{1,K}^n$. This is explained by the proof of the following proposition.

**Proposition 1** (Schemes equivalence)**.** Let us supplement the scheme (31)–(34) with the definitions of the quantities $(V_\sigma^{n+1})_{\sigma\in\mathcal{E},\,n\geq 0}$ and $(u_{2,\sigma}^{n+1})_{\sigma\in\mathcal{E},\,n\geq 0}$. We distinguish two cases:

- Case 1: centered scheme. For all $\sigma\in\mathcal{E}$ and $n\geq 0$, we set:

$$u_{2,\sigma}^{n+1}=\frac{u_{2,K}^{n+1}+u_{2,K,\sigma}^{n+1}}{2}\quad\text{and}\quad V_\sigma^{n+1}=\frac{V_K^{n+1}+V_{K,\sigma}^{n+1}}{2}. \tag{43}$$

- Case 2: upwind scheme. For all $\sigma\in\mathcal{E}$ and $n\geq 0$, we set:

$$u_{2,\sigma}^{n+1}=\begin{cases}u_{2,K,\sigma}^{n+1}, & \text{if } D_{K,\sigma}V_{\mathcal{M}}^{n+1}>0,\\ u_{2,K}^{n+1}, & \text{if } D_{K,\sigma}V_{\mathcal{M}}^{n+1}\leq 0,\end{cases}\quad\text{and}\quad V_\sigma^{n+1}=\min(V_K^{n+1},V_{K,\sigma}^{n+1}). \tag{44}$$

Then, in both cases, the scheme (31)–(34) written in the primal entropy variables is equivalent to the scheme (38)–(42) written in the dual entropy variables, provided that

$$\mathbb{D}_\sigma^*=(\mathbb{P}_\sigma^{n+1})^T\mathbb{L}_\sigma^n\mathbb{P}_\sigma^{n+1}\quad\text{with}\quad\mathbb{P}_\sigma^{n+1}=\begin{pmatrix}1 & -V_\sigma^{n+1}\\ 0 & 1\end{pmatrix}. \tag{45}$$

*Proof.* Starting from the definition (41) of the numerical fluxes $\mathcal{G}_{1,K,\sigma}^{n+1}$ and $\mathcal{G}_{2,K,\sigma}^{n+1}$, we want to establish the relations (42) with $\mathbb{D}_\sigma^*$ defined by (45). Observe that, from (45), one deduces the following relations, which are the discrete counterparts of (21):

$$\mathbb{D}_{11,\sigma}^*=L_{11,\sigma}^n, \tag{46}$$

$$\mathbb{D}_{12,\sigma}^*=\mathbb{D}_{21,\sigma}^*=-V_\sigma^{n+1}L_{11,\sigma}^n+L_{12,\sigma}^n, \tag{47}$$

$$\mathbb{D}_{22,\sigma}^*=\left(V_\sigma^{n+1}\right)^2L_{11,\sigma}^n-2V_\sigma^{n+1}L_{12,\sigma}^n+L_{22,\sigma}^n. \tag{48}$$

Then, due to the change of variables (35), we can rewrite $D_{K,\sigma}u_{1,\mathcal{M}}^{n+1}$ and $D_{K,\sigma}u_{2,\mathcal{M}}^{n+1}$ for all $K\in\mathcal{T}$ and $\sigma\in\mathcal{E}_K$. It is straightforward that $D_{K,\sigma}u_{2,\mathcal{M}}^{n+1}=D_{K,\sigma}w_{2,\mathcal{M}}^{n+1}$. Concerning $D_{K,\sigma}u_{1,\mathcal{M}}^{n+1}$, two formulations are possible, either

$$D_{K,\sigma}u_{1,\mathcal{M}}^{n+1}=D_{K,\sigma}w_{1,\mathcal{M}}^{n+1}-V_K^{n+1}D_{K,\sigma}w_{2,\mathcal{M}}^{n+1}-w_{2,K,\sigma}^{n+1}D_{K,\sigma}V_{\mathcal{M}}^{n+1},$$

or

$$D_{K,\sigma}u_{1,\mathcal{M}}^{n+1}=D_{K,\sigma}w_{1,\mathcal{M}}^{n+1}-V_{K,\sigma}^{n+1}D_{K,\sigma}w_{2,\mathcal{M}}^{n+1}-w_{2,K}^{n+1}D_{K,\sigma}V_{\mathcal{M}}^{n+1}.$$

The definitions of the interfacial quantities given by (43) in the centered case and by (44) in the upwind case lead to the same definition of $D_{K,\sigma}u_{1,\mathcal{M}}^{n+1}$, namely

$$D_{K,\sigma}u_{1,\mathcal{M}}^{n+1}=D_{K,\sigma}w_{1,\mathcal{M}}^{n+1}-V_\sigma^{n+1}D_{K,\sigma}w_{2,\mathcal{M}}^{n+1}-w_{2,\sigma}^{n+1}D_{K,\sigma}V_{\mathcal{M}}^{n+1},$$

where $w_{2,\sigma}^{n+1}=u_{2,\sigma}^{n+1}$. Finally, plugging the expressions (34) of the numerical fluxes $\mathcal{F}_{i,K,\sigma}^{n+1}$ and the latter expressions $D_{K,\sigma}u_{i,\mathcal{M}}^{n+1}$, $i=1,2$, into the relations (41) of the numerical fluxes $\mathcal{G}_{i,K,\sigma}^{n+1}$ gives

$$\begin{aligned}G_{1,K,\sigma}^{n+1}&=-\tau_\sigma\left(L_{11,\sigma}^nD_{K,\sigma}w_{1,\mathcal{M}}^{n+1}+\left(L_{12,\sigma}^n-V_\sigma^{n+1}L_{11,\sigma}^n\right)D_{K,\sigma}w_{2,\mathcal{M}}^{n+1}\right),\\ G_{2,K,\sigma}^{n+1}&=-\tau_\sigma\Big(\left(L_{12,\sigma}^n-V_\sigma^{n+1}L_{11,\sigma}^{n+1}\right)D_{K,\sigma}w_{1,\mathcal{M}}^{n+1}\\ &\quad+\left(L_{22,\sigma}^n-2V_\sigma^{n+1}L_{12,\sigma}^n+(V_\sigma^{n+1})^2L_{11,\sigma}^n\right)D_{K,\sigma}w_{2,\mathcal{M}}^{n+1}\Big).\end{aligned}$$

This corresponds to the expected relations (42) with the coefficients (46)–(48) of the matrix $\mathbb{D}_\sigma^*$. Hence we have shown that the scheme in primal variables (31)–(34), supplemented by either the centered or the upwind definitions of the interfacial quantities $V_\sigma^{n+1}$ and $u_{2,\sigma}^{n+1}$, is equivalent to the numerical scheme in the dual entropy variables (38)–(42). Conversely, starting from the dual scheme (38)–(42) with the matrix $\mathbb{D}_\sigma^*$ given in (45), one deduces the primal scheme (31)–(34).

□

## 3. Discrete entropy inequality and *a priori* estimates

In this section, we establish the discrete counterpart of the entropy inequality (24) and then deduce some *a priori* estimates on the approximate solution.

First of all, recall that the functions $u_1^D$, $u_2^D$, $V^D$ are assumed to be defined on the whole domain $\Omega$. Hence we can set

$$(u_{1,K}^D, u_{2,K}^D, V_K^D) = \frac{1}{\mathrm{m}(K)} \int_K (u_1^D(x), u_2^D(x), V^D(x))dx, \quad \forall K \in \mathcal{T}.$$

Moreover the assumption (10) on the boundary condition implies that

$$u_{K,2}^D = u_2^D < 0, \quad \forall K \in \mathcal{T}. \tag{49}$$

### 3.1. Discrete entropy inequality.

Let $\left(\mathbf{u}_K^n = (u_{1,K}^n, u_{2,K}^n)^T, V_K^n\right)_{K\in\mathcal{T}, n\geq 0}$ be a solution to the scheme in primal variables (31)–(34), supplemented with either the definition (43) or (44) of the interfacial quantities.

For all $n \geq 0$, we define the discrete entropy functional as follows:

$$\begin{aligned} S^n = \sum_{K\in\mathcal{T}} \mathrm{m}(K) \left[\boldsymbol{\rho}_K^n \cdot (\mathbf{u}_K^n - \mathbf{u}_K^D) - (\chi(\mathbf{u}_K^n) - \chi(\mathbf{u}_K^D))\right] \\ - \frac{\lambda^2}{2} u_2^D \sum_{\sigma\in\mathcal{E}} \tau_\sigma (D_\sigma(V_\mathcal{M}^n - V_\mathcal{M}^D))^2, \end{aligned} \tag{50}$$

where $\boldsymbol{\rho}_K^n = \boldsymbol{\rho}(\mathbf{u}_K^n) = (\rho_1(\mathbf{u}_K^n), \rho_2(\mathbf{u}_K^n))^T$. Since $\boldsymbol{\rho}$ is related to the convex potential $\chi$ by Assumptions 1-(iii), one deduces that $S^n$ is nonnegative for all $n \geq 0$.

We now establish the discrete counterpart of the entropy inequality (24).

**Proposition 2** (Discrete entropy dissipation)**.** Let assume that Assumptions 1, (10) and (11) hold. Let us also assume that the boundary conditions are at thermal equilibrium, namely that

$$D_{K,\sigma} w_{1,\mathcal{M}}^D = D_{K,\sigma} w_{2,\mathcal{M}}^D = 0. \tag{51}$$

Let $(\mathbf{u}_K^n, V_K^n)_{K\in\mathcal{T}, n\geq 0}$ be a solution to the scheme in primal variables (31)–(34), supplemented with either (43) or (44). The discrete entropy (50) satisfies the following inequality: for all $n \geq 0$,

$$\begin{aligned} \frac{S^{n+1} - S^n}{\Delta t} \leq - \sum_{\sigma\in\mathcal{E}} \tau_\sigma (D_{K,\sigma} \mathbf{w}_\mathcal{M}^{n+1})^T \mathbb{D}_\sigma^* D_{K,\sigma} \mathbf{w}_\mathcal{M}^{n+1} \\ + \sum_{K\in\mathcal{T}} \mathrm{m}(K) W_K^{n+1} (w_{2,K}^{n+1} - w_{2,K}^D) \leq 0, \end{aligned} \tag{52}$$

where $D_{K,\sigma}\mathbf{w}_\mathcal{M}^{n+1} = (D_{K,\sigma} w_{1,\mathcal{M}}^{n+1}, D_{K,\sigma} w_{2,\mathcal{M}}^{n+1})^T$.

*Proof.* Using the definition (50) of the discrete entropy, one has

$$S^{n+1} - S^n = A + B,$$

where

$$\begin{aligned} A = &\sum_{K\in\mathcal{T}} \mathrm{m}(K) \left[\boldsymbol{\rho}_K^{n+1}\cdot(\mathbf{u}_K^{n+1}-\mathbf{u}_K^D) - (\chi(\mathbf{u}_K^{n+1})-\chi(\mathbf{u}_K^D))\right.\\ &\left.-\boldsymbol{\rho}_K^{n}\cdot(\mathbf{u}_K^{n}-\mathbf{u}_K^D) + \chi(\mathbf{u}_K^{n})-\chi(\mathbf{u}_K^D)\right],\\ B = &-\frac{\lambda^2}{2}u_2^D\sum_{\sigma\in\mathcal{E}}\tau_\sigma\left[(D_\sigma(V_\mathcal{M}^{n+1}-V_\mathcal{M}^D))^2-(D_\sigma(V_\mathcal{M}^n-V_\mathcal{M}^D))^2\right]. \end{aligned}$$

We first consider the term $A$, where the quantities $\chi(\mathbf{u}_K^D)$ cancel each others. Adding and subtracting $\mathbf{u}_K^{n+1}\cdot\boldsymbol{\rho}_K^n$, the term $A$ now reads

$$A = \sum_{K\in\mathcal{T}} \mathrm{m}(K)(\boldsymbol{\rho}_K^{n+1}-\boldsymbol{\rho}_K^n)\cdot(\mathbf{u}_K^{n+1}-\mathbf{u}_K^D) - (\chi(\mathbf{u}_K^{n+1})-\chi(\mathbf{u}_K^n)) - \boldsymbol{\rho}_K^n\cdot(\mathbf{u}_K^n-\mathbf{u}_K^{n+1}).$$

Since $\chi$ is a convex function such that $\boldsymbol{\rho} = \nabla_{\mathbf{u}}\chi$, leading to $\boldsymbol{\rho}_K^n = \nabla_{\mathbf{u}}\chi(\mathbf{u}_K^n)$, we have:

$$\chi(\mathbf{u}_K^{n+1}) - \chi(\mathbf{u}_K^n) - \boldsymbol{\rho}_K^n\cdot(\mathbf{u}_K^{n+1}-\mathbf{u}_K^n) \geq 0.$$

It yields

$$A \leq \sum_{K\in\mathcal{T}} \mathrm{m}(K)(\boldsymbol{\rho}_K^{n+1}-\boldsymbol{\rho}_K^n)\cdot(\mathbf{u}_K^{n+1}-\mathbf{u}_K^D). \tag{53}$$

We now address the term $B$. Since $(a^2-b^2)\leq 2a(a-b)$ for all $a,b\in\mathbb{R}$, one has

$$B \leq -\lambda^2 u_2^D \sum_{\sigma\in\mathcal{E}} \tau_\sigma D_{K,\sigma}(V_\mathcal{M}^{n+1}-V_\mathcal{M}^D) D_{K,\sigma}(V_\mathcal{M}^{n+1}-V_\mathcal{M}^n).$$

A discrete integration by part leads to

$$B \leq \lambda^2 u_2^D \sum_{K\in\mathcal{T}} (V_K^{n+1}-V_K^D)\left(\sum_{\sigma\in\mathcal{E}}\tau_\sigma D_{K,\sigma}(V_\mathcal{M}^{n+1}-V_\mathcal{M}^n)\right).$$

Using the scheme for the Poisson equation (33), one gets

$$B \leq u_2^D \sum_{K\in\mathcal{T}} \mathrm{m}(K)(V_K^{n+1}-V_K^D)(\rho_{1,K}^{n+1}-\rho_{1,K}^n). \tag{54}$$

Combining (53) and (54) gives

$$\begin{aligned} S^{n+1}-S^n \leq &\sum_{K\in\mathcal{T}} \mathrm{m}(K)(\rho_{1,K}^{n+1}-\rho_{1,K}^n)\left[(u_{1,K}^{n+1}-u_{1,K}^D) + u_2^D(V_K^{n+1}-V_K^D)\right]\\ &+\sum_{K\in\mathcal{T}} \mathrm{m}(K)(\rho_{2,K}^{n+1}-\rho_{2,K}^n)(u_{2,K}^{n+1}-u_{2,K}^D). \end{aligned} \tag{55}$$

Using the primal scheme (31)-(32), the inequality (55) becomes

$$\frac{S^{n+1}-S^n}{\Delta t} \leq \sum_K \mathrm{m}(K) W_K^{n+1}(u_{2,K}^{n+1}-u_{2,K}^D) + C + D,$$

with

$$C = -\sum_{K\in\mathcal{T}}\left(\sum_{\sigma\in\mathcal{E}}\mathcal{F}^{n+1}_{1,K,\sigma}\right)\left((u^{n+1}_{1,K}-u^D_{1,K})+V^{n+1}_K(u^{n+1}_{2,K}-u^D_{2,K})\right)$$
$$-u^D_2\sum_{K\in\mathcal{T}}\left(\sum_{\sigma\in\mathcal{E}}\mathcal{F}^{n+1}_{1,K,\sigma}\right)(V^{n+1}_K-V^D_K),$$
$$D = -\sum_{K\in\mathcal{T}}\left(\sum_{\sigma\in\mathcal{E}}\mathcal{F}^{n+1}_{2,K,\sigma}-V^{n+1}_\sigma\mathcal{F}^{n+1}_{1,K,\sigma}\right)(u^{n+1}_{2,K}-u^D_{2,K}).$$

Now using the relation (49), the change of variables (35) and the relations (41) between the primal and dual numerical fluxes, it holds

$$C = -\sum_{K\in\mathcal{T}}\left(\sum_{\sigma\in\mathcal{E}}\mathcal{G}^{n+1}_{1,K,\sigma}\right)(w^{n+1}_{1,K}-w^D_{1,K}),$$
$$D = -\sum_{K\in\mathcal{T}}\left(\sum_{\sigma\in\mathcal{E}}\mathcal{G}^{n+1}_{2,K,\sigma}\right)(w^{n+1}_{2,K}-w^D_{2,K}).$$

Accounting for the boundary conditions at thermal equilibrium (51), we conclude by a discrete integration by parts which gives:

$$\frac{S^{n+1}-S^n}{\Delta t} \le \sum_{\sigma\in\mathcal{E}}\mathcal{G}^{n+1}_{1,K,\sigma}D_{K,\sigma}w^{n+1}_{1,\mathcal{M}}+\sum_{\sigma\in\mathcal{E}}\mathcal{G}^{n+1}_{2,K,\sigma}D_{K,\sigma}w^{n+1}_{2,\mathcal{M}} + \sum_{K\in\mathcal{T}}\mathrm{m}(K)W^{n+1}_K(w^{n+1}_{2,K}-w^D_{2,K}). \tag{56}$$

The formulation (42) of the numerical fluxes $\mathcal{G}^{n+1}_{i,K,\sigma}$ permits to rewrite

$$\sum_{\sigma\in\mathcal{E}}\mathcal{G}^{n+1}_{1,K,\sigma}D_{K,\sigma}w^{n+1}_{1,\mathcal{M}}+\sum_{\sigma\in\mathcal{E}}\mathcal{G}^{n+1}_{2,K,\sigma}D_{K,\sigma}w^{n+1}_{2,\mathcal{M}} = -\sum_{\sigma\in\mathcal{E}}\tau_\sigma(D_{K,\sigma}\mathbf{w}^{n+1}_{\mathcal{M}})^T\mathbb{D}^*_\sigma D_{K,\sigma}\mathbf{w}^{n+1}_{\mathcal{M}} \tag{57}$$

From equations (56) and (57), one deduces the entropy inequality (52). The assumption (11) on the energy relaxation term and the positive definiteness of the matrices $\mathbb{D}^*_\sigma$ ensure the nonpositivity of the right-hand-side of (52) and the decay of the discrete entropy. □

3.2. ***A priori* estimates on the approximate solutions.** We start with a uniform $L^\infty$ bound on $(V^n_K)_{K\in\mathcal{T},\,n\ge0}$, which will be useful to derive other estimates.

**Lemma 1** ($L^\infty$ bound on $V_\mathcal{T}$)**.** Let $(\mathbf{u}^n_K, V^n_K)_{K\in\mathcal{T},n\ge0}$ be a solution to the scheme (31)–(34). There exists a constant $C_V>0$ only depending on $\Omega$, $\|C\|_\infty$, $C_\rho$ and $\|V^D\|_\infty$ such that

$$|V^n_K|\le C_V,\quad \forall K\in\mathcal{T},\,\forall n\ge0. \tag{58}$$

*Proof.* The assumptions stated in Section 1.1 ensure that the doping profile $C$, the electron density $\rho_1(\mathbf{u})$ and the potential boundary condition $V^D$ are in $L^\infty(\Omega)$. Hence, one can apply [8, Proposition A.1] to get the expected bound. □

From the discrete entropy estimate (52), one may now deduce the following uniform discrete bounds.

**Proposition 3** (Discrete *a priori* estimates)**.** Assume Assumptions 1 hold. Let $\mathcal{T}$ be an admissible mesh of $\Omega$ and $\Delta t > 0$. Assume the initial and boundary conditions satisfy the hypotheses (10) and (11). Let $(\mathbf{u}_K^n, V_K^n)_{K\in\mathcal{T}, n\geq 0}$ be a solution to the scheme in primal variables (31)–(34), supplemented with either (43) or (44). Then there exist a constant $C_1 > 0$, depending only on the initial and boundary conditions and on the Debye length $\lambda^2$, and a constant $C_2 > 0$, depending only on the initial and boundary conditions, on $\alpha$ defined in (6) and on $C_V$ defined in (58) such that

$$\sup_{n=0,\dots,N_T} \left( \|u_{1,\mathcal{T}}^n - u_{1,\mathcal{T}}^D\|_{L^2} + \|u_{2,\mathcal{T}}^n - u_{2,\mathcal{T}}^D\|_{L^2} + |V_{\mathcal{M}}^{n+1} - V_{\mathcal{M}}^D|_{1,\mathcal{M}}^2 \right) \leq C_1, \tag{59}$$

$$\sum_{n=0}^{N_T-1} \Delta t \left( \left| w_{1,\mathcal{M}}^{n+1} \right|_{1,\mathcal{M}}^2 + \left| w_{2,\mathcal{M}}^{n+1} \right|_{1,\mathcal{M}}^2 \right) \leq C_2. \tag{60}$$

*Proof.* Proposition 2 ensures the dissipation of the discrete entropy. Since $\chi \in C^1(\mathbb{R}^2, \mathbb{R})$ satisfies $\nabla\chi = \boldsymbol{\rho}$, the Taylor's theorem gives, for all $K \in \mathcal{T}$

$$\chi(\mathbf{u}_K^n) - \chi(\mathbf{u}_K^D) = \int_0^1 \boldsymbol{\rho}(\mathbf{u}_K^D + s(\mathbf{u}_K^n - \mathbf{u}_K^D)) \cdot (\mathbf{u}_K^n - \mathbf{u}_K^D) ds.$$

Then one can rewrite the first term of the discrete entropy $S^n$, defined in equation (50), as follow:

$$\sum_{K\in\mathcal{T}} \mathrm{m}(K) \int_0^1 \left( \boldsymbol{\rho}\left(\mathbf{u}_K^n\right) - \boldsymbol{\rho}\left(\mathbf{u}_K^D + s\left(\mathbf{u}_K^n - \mathbf{u}_K^D\right)\right)\right) \cdot \left(\mathbf{u}_K^n - \mathbf{u}_K^D\right) ds =$$

$$\sum_{K\in\mathcal{T}} \mathrm{m}(K) \int_0^1 \left( \boldsymbol{\rho}(\mathbf{u}_K^n) - \boldsymbol{\rho}\left(\mathbf{u}_K^D + s(\mathbf{u}_K^n - \mathbf{u}_K^D)\right)\right) \cdot \left(\mathbf{u}_K^n - (s\mathbf{u}_K^n + (1-s)\mathbf{u}_K^D)\right) \frac{ds}{1-s}.$$

Using the monotony assumption (12) of $\boldsymbol{\rho}$, we finally get

$$S^0 \geq S^n \geq \frac{C_0}{2} \sum_{K\in\mathcal{T}} \mathrm{m}(K) |\mathbf{u}_K^n - \mathbf{u}_K^D|^2 - \frac{\lambda^2}{2} u_2^D \sum_{\sigma\in\mathcal{E}} \tau_\sigma D_\sigma (V_{\mathcal{M}}^n - V_{\mathcal{M}}^D)^2,$$

which yields the *a priori* estimate (59). The estimate (60) is obtained by summing the discrete entropy inequality (52) for $n = 0, \dots, N_T - 1$. Since the energy relaxation satisfies (11), the control of the entropy dissipation writes

$$\sum_{n=0}^{N_T-1} \Delta t \sum_{\sigma\in\mathcal{E}} \tau_\sigma (D_{K,\sigma} \mathbf{w}_{\mathcal{M}}^{n+1})^T \mathbb{D}_\sigma^* D_{K,\sigma} \mathbf{w}_{\mathcal{M}}^{n+1} \leq S^0. \tag{61}$$

Using the definition (45) of $\mathbb{D}_\sigma^*$ and assumption (6) on $\mathbb{L}_\sigma^n$, we have, for all $\sigma \in \mathcal{E}$ and all $\mathbf{v} \in \mathbb{R}^2$

$$\mathbf{v}^T \mathbb{D}_\sigma^* \mathbf{v} \geq \alpha |\mathbb{P}_\sigma^{n+1} \mathbf{v}|^2 \geq \frac{\alpha}{\|(\mathbb{P}_\sigma^{n+1})^{-1}\|^2} |\mathbf{v}|^2.$$

Now, using the definition of $\mathbb{P}_\sigma^{n+1}$, we observe that $\|(\mathbb{P}_\sigma^{n+1})^{-1}\|$ is a nondecreasing function of $|V_\sigma^{n+1}|$, where $V_\sigma^{n+1}$ is defined either by (43) or by (44). Using the $L^\infty$

bound (58), we finally obtain that there exists a constant $\beta > 0$, depending on $C_V$ and $\alpha$, such that for all $\sigma \in \mathcal{E}$, for all $\mathbf{v} \in \mathbb{R}^2$

$$\mathbf{v}^T \, \mathbb{D}_\sigma^* \, \mathbf{v} \geq \beta |\mathbf{v}|^2.$$

Applying this estimate in (61) yields the *a priori* estimate (60). □

Since the mesh satisfies the regularity constraint (25), and as a consequence of a discrete Poincaré inequality [4], it holds

$$\sum_{n=0}^{N_T-1} \Delta t \left( \|w_{1,\mathcal{M}}^{n+1}\|_{1,\mathcal{M}}^2 + \|w_{2,\mathcal{M}}^{n+1}\|_{1,\mathcal{M}}^2 \right) \leq c, \tag{62}$$

$$\sup_{n=0,\dots,N_T} \left( \|V_{\mathcal{M}}^{n+1} - V_{\mathcal{M}}^D\|_{1,\mathcal{M}} \right) \leq c. \tag{63}$$

## 4. Existence of a solution to the schemes

In this section, we state the existence of a solution to the schemes (both primal and dual, due to the equivalence property, see Proposition 1). The proof is based on a Leray-Schauder fixed point theorem, and follows the same guidelines as the existence proof given in [14, Lemma 3.2].

**Theorem 1** (Existence of a solution)**.** Let $(\mathcal{T}, \mathcal{E}, \mathcal{P})$ be an admissible mesh of $\Omega$ satisfying the regularity constraint (25). Assume Assumptions 1, (10) and (11) hold. Let $(w_{1,\mathcal{T}}^0, w_{2,\mathcal{T}}^0)$ be defined by (35) and (28). Then, for all $0 \leq n \leq N_T$, the primal scheme (31)–(34) has a solution $(u_{1,\mathcal{T}}^n, u_{2,\mathcal{T}}^n, V_{\mathcal{T}}^n)$. As a consequence, by equivalence, the dual scheme (38)–(42) has a solution $(w_{1,\mathcal{T}}^n, w_{2,\mathcal{T}}^n, V_{\mathcal{T}}^n)$ for all $0 \leq n \leq N_T$.

*Proof.* The result is proved by induction on $0 \leq n \leq N_T$. The principle of Leray-Schauder fixed point theorem is to transform continuously the nonlinear system into a linear one while ensuring that *a priori* estimates controlling the solution remain all along the homotopy. We sketch the main idea of the proof, following [14, Lemma 3.2], parametrizing the homotopy by $\kappa \in [0, 1]$.

Let $n \geq 0$ and $w_{1,\mathcal{T}}^n$, $w_{2,\mathcal{T}}^n$, $V_{\mathcal{T}}^n$ be given, as well as the Dirichlet boundary conditions $w_{1,\sigma}^D$, $w_{2,\sigma}^D$ and $V_\sigma^D$, $\forall \sigma \in \mathcal{E}_{ext}^D$. We apply the Leray-Schauder theorem to the application

$$\begin{aligned} \mathcal{L}^n : \mathbb{R}^{2\theta} \times [0,1] &\to \mathbb{R}^{2\theta} \\ (\widetilde{\mathbf{u}}_{\mathcal{T}}, \kappa) &\mapsto \mathbf{u}_{\mathcal{T}} \end{aligned}$$

defined in four steps as follows. Here $\theta$ denotes the cardinal of $\mathcal{T}$, see Section 2.1.

*Step 1.* Let $V_{\mathcal{T}} \in \mathbb{R}^\theta$ be the solution to the following linear system

$$\begin{cases} -\lambda^2 \displaystyle\sum_{\sigma \in \mathcal{E}} \tau_\sigma D_{K,\sigma} V_{\mathcal{M}} = \mathrm{m}(K) \left( C_K - \rho_1(\widetilde{\mathbf{u}}_K) \right), & \forall K \in \mathcal{T}, \\ V_\sigma = V_\sigma^D, & \forall \sigma \in \mathcal{E}_{ext}^D. \end{cases} \tag{64}$$

*Step 2.* We make use of the change of variables (13) to define $\widetilde{\mathbf{w}}_{\mathcal{T}} = (\widetilde{w}_{1,\mathcal{T}}, \widetilde{w}_{2,\mathcal{T}}) \in \mathbb{R}^{2\theta}$ as

$$\widetilde{w}_{1,K} = \widetilde{u}_{1,K} + V_K \widetilde{u}_{2,K}, \quad \widetilde{w}_{2,K} = \widetilde{u}_{2,K}, \quad \forall K \in \mathcal{T}.$$

*Step 3.* We define $\mathbf{w}_\mathcal{T} = (w_{1,\mathcal{T}}, w_{2,\mathcal{T}}) \in \mathbb{R}^{2\theta}$ as the solution to the following linear problem: for all $K \in \mathcal{T}$,

(65)
$$
\begin{aligned}
-\sum_{\sigma\in\mathcal{E}} \tau_\sigma \left(D_{11,\sigma}(\widetilde{\mathbf{w}}, V) D_{K,\sigma} w_{1,\mathcal{M}} \right. &\left. + D_{12,\sigma}(\widetilde{\mathbf{w}}, V) D_{K,\sigma} w_{2,\mathcal{M}}\right) \\
&= -\kappa\, \mathrm{m}(K) \frac{b_1(\widetilde{\mathbf{w}}_K, V_K) - b_1(\mathbf{w}_K^n, V_K^n)}{\Delta t}, \\
-\sum_{\sigma\in\mathcal{E}} \tau_\sigma \left(D_{21,\sigma}(\widetilde{\mathbf{w}}, V) D_{K,\sigma} w_{1,\mathcal{M}} \right. &\left. + D_{22,\sigma}(\widetilde{\mathbf{w}}, V) D_{K,\sigma} w_{2,\mathcal{M}}\right) \\
&= \kappa \left[ -\mathrm{m}(K) \frac{b_2(\widetilde{\mathbf{w}}_K, V_K) - b_2(\mathbf{w}_K^n, V_K^n)}{\Delta t} \right. \\
&\left. + \mathrm{m}(K) \widetilde{W}(\widetilde{\mathbf{w}}_K, V_K) - \mathrm{m}(K) \frac{V_K - V_K^n}{\Delta t} b_1(\mathbf{w}_K^n, V_K^n) \right],
\end{aligned}
$$

associated with the following boundary conditions

$$
\begin{aligned}
&w_{1,\sigma} = \kappa w_{1,\sigma}^D, \quad w_{2,\sigma} = \kappa w_{2,\sigma}^D, \quad \forall \sigma \in \mathcal{E}_{ext}^D, \\
&D_{11,\sigma}(\widetilde{\mathbf{w}}, V) D_{K,\sigma} w_{1,\mathcal{M}} + D_{12,\sigma}(\widetilde{\mathbf{w}}, V) D_{K,\sigma} w_{2,\mathcal{M}} = 0, \quad \sigma \in \mathcal{E}_{ext}^N, \\
&D_{21,\sigma}(\widetilde{\mathbf{w}}, V) D_{K,\sigma} w_{1,\mathcal{M}} + D_{22,\sigma}(\widetilde{\mathbf{w}}, V) D_{K,\sigma} w_{2,\mathcal{M}} = 0, \quad \sigma \in \mathcal{E}_{ext}^N.
\end{aligned} \tag{66}
$$

*Step 4.* Making use of the entropy-primal change of variables (13), we define $\mathbf{u}_\mathcal{T} = (u_{1,\mathcal{T}}, u_{2,\mathcal{T}}) \in \mathbb{R}^{2\theta}$ as

$$
u_{1,K} = w_{1,K} - w_{2,K} V_K, \quad u_{2,K} = w_{2,K}, \quad \forall K \in \mathcal{T}.
$$

We remark that a fixed point $\mathbf{u}_\mathcal{T}$ of $\mathcal{L}^n(\cdot, \kappa = 1)$, together with the corresponding $V_\mathcal{T}$, gives a solution $(u_{1,\mathcal{T}}^{n+1}, u_{2,\mathcal{T}}^{n+1}, V_\mathcal{T}^{n+1})$ to the scheme (31)–(34). First of all, we have to check that the application $\mathcal{L}^n$ is well-defined. Existence and uniqueness of $V_\mathcal{T}$ solution to the linear system (64) is obvious. Moreover, Lemma 1 ensures an $L^\infty$ bound on $V_\mathcal{T}$. Then, we prove the existence and uniqueness of $\mathbf{w}_\mathcal{T} = (w_{1,\mathcal{T}}, w_{2,\mathcal{T}})$, solution to (65)–(66). Since it is a finite dimensional linear system, it is sufficient to prove uniqueness of $\mathbf{w}_\mathcal{T}$, and next that if $\kappa = 0$, then $\mathbf{w}_\mathcal{T} = \mathbf{0}$. Let us multiply the first equation of (65) by $w_{1,K}$ and the second by $w_{2,K}$. Adding both equations, summing over all the control volumes $K \in \mathcal{T}$ and performing a discrete integration by parts leads to

$$
\sum_{\sigma\in\mathcal{E}} \tau_\sigma (D_{K,\sigma} \mathbf{w}_\mathcal{M})^T \widetilde{\mathbb{D}}_\sigma D_{K,\sigma} \mathbf{w}_\mathcal{M} = 0,
$$

where $(\widetilde{\mathbb{D}}_\sigma)_{ij} = D_{ij,\sigma}(\widetilde{\mathbf{w}}, V)$, $i = 1, 2$. Since the matrices $\widetilde{\mathbb{D}}_\sigma$ are symmetric positive definite, $D_{K,\sigma} w_{i,\mathcal{M}} = 0$ for all $\sigma \in \mathcal{E}$, $i = 1, 2$. Due to the homogeneous Dirichlet boundary conditions (66) (with $\kappa = 0$), the discrete Poincaré inequality [4] ensures that $w_{1,K} = w_{2,K} = 0$, for all $K \in \mathcal{T}$. Hence the system is invertible.

We now have to verify the assumptions of the Leray-Schauder theorem. First observe that every step of construction of $\mathcal{L}^n$ is continuous on $\mathbb{R}^{2\theta} \times [0, 1]$. Since the problem is set in finite dimension, it ensures the required compactness. We have already established that $\mathcal{L}^n(\widetilde{\mathbf{u}}_\mathcal{T}, 0) = \mathbf{0}$. Now, it remains to verify that there exists a constant $M > 0$ such that, for all $\kappa \in [0, 1]$, for all $\mathbf{u}_\mathcal{T}$ satisfying $\mathcal{L}^n(\mathbf{u}_\mathcal{T}, \kappa) = \mathbf{u}_\mathcal{T}$, it holds $\|\mathbf{u}_\mathcal{T}\|_{L^2} \leq M$. Let us remark that if $\mathcal{L}^n(\mathbf{u}_\mathcal{T}, \kappa) = \mathbf{u}_\mathcal{T}$ then $\widetilde{\mathbf{w}}_\mathcal{T} = \mathbf{w}_\mathcal{T}$.

Using the proofs of Propositions 2 and 3, we get

$$\|\mathbf{u}_\mathcal{T} - \mathbf{u}_\mathcal{T}^D\|_{L^2}^2 \leq \frac{2S^n}{C_0},$$

where $C_0$ is the constant in (12) and $S^n$ is the discrete entropy defined from the known quantities $w_{1,\mathcal{T}}^n$, $w_{2,\mathcal{T}}^n$ and $V_\mathcal{T}^n$, together with the change of variable (13). Thus the Leray-Schauder theorem applies, which concludes the proof. □

## 5. Long time behavior

In this section, we prove the exponential convergence of the approximate solution of the energy-transport system towards an approximation of the thermal equilibrium as time tends to infinity.

5.1. **Definition of the thermal equilibrium.** According to [14], the thermal equilibrium is defined as a particular steady state of the energy-transport system (2)–(5) for which $J_1 = 0 = J_2$, which is equivalent to $\nabla \mathbf{w} = 0$. Hence the temperature $T = -1/w_2$ is constant in all the domain $\Omega$. As in [14], we need some additional compatibility assumptions on the Dirichlet boundary conditions in this section.

**Assumptions 2.** The functions $\mathbf{u}^D \in H^1(\Omega)$, $V^D \in L^\infty(\Omega) \cap H^1(\Omega)$ satisfy

(i) $W(\mathbf{u}^D) = 0$,
(ii) $\nabla \mathbf{w}^D = 0$,
(iii) $-\lambda^2 \Delta V^D = C - \rho_1(\mathbf{u}^D)$.

We also assume that

$$u_1 \mapsto \rho_1(\mathbf{u}) \text{ is increasing.} \tag{67}$$

This assumption is fulfilled by standard examples, proposed for instance in [31, 14], which we consider in Section 6.

Under these assumptions, the thermal equilibrium $(\mathbf{w}^{eq}, V^{eq})$ is defined as

$$\mathbf{w}^{eq} = \mathbf{w}^D \quad \text{on } \Omega, \tag{68}$$

$$-\lambda^2 \Delta V^{eq} = C - b_1(\mathbf{w}^{eq}, V^{eq}), \quad \text{on } \Omega, \tag{69}$$

with mixed boundary conditions $V^{eq} = V^D$ on $\Gamma^D$ and $\nabla V^{eq} \cdot \mathbf{n} = 0$ on $\Gamma^N$. Under the previous hypotheses, there exists a unique solution $V^{eq} \in H^1(\Omega)$ to this nonlinear elliptic equation.

Now consider a finite volume approximation $(\mathbf{w}_\mathcal{T}^{eq}, V_\mathcal{T}^{eq})$ of the thermal equilibrium. We set, for all $K \in \mathcal{T}$,

$$\mathbf{w}_K^{eq} = \mathbf{w}_K^D, \tag{70}$$

$$-\lambda^2 \sum_{\sigma \in \mathcal{E}_K} \tau_\sigma D_{K,\sigma} V_\mathcal{M}^{eq} = \mathrm{m}(K) \left( C_K - b_1(\mathbf{w}_K^{eq}, V_K^{eq}) \right), \tag{71}$$

with $V_\sigma^{eq} = V_\sigma^D$ for all $\sigma \in \mathcal{E}_{ext}^D$. Existence and uniqueness of a solution to the numerical scheme (70)–(71) is established in [10, Proposition 3.1]. Under Assumptions 2 on the compatible boundary conditions and the monotony assumption (67) on $\rho_1$, an application of the Leray-Schauder fixed point theorem states that there exists a unique solution $V_\mathcal{T}^{eq}$ to the scheme (70)–(71) satisfying a $L^\infty$ estimate: $|V_K^{eq}| \leq C_{V^{eq}}$ for all $K \in \mathcal{T}$.

5.2. **Exponential decay towards equilibrium.** In order to study the decay towards the thermal equilibrium, we introduce the discrete relative entropy with respect to the thermal equilibrium: for $n \geq 0$,

$$\widetilde{S}^n = \sum_{K\in\mathcal{T}} \mathrm{m}(K)\left[\boldsymbol{\rho}_K^n \cdot (\mathbf{u}_K^n - \mathbf{u}_K^{eq}) - (\chi(\mathbf{u}_K^n) - \chi(\mathbf{u}_K^{eq}))\right] - \frac{\lambda^2}{2} u_2^D \sum_{\sigma\in\mathcal{E}} \tau_\sigma (D_\sigma(V_\mathcal{M}^n - V_\mathcal{M}^{eq}))^2, \tag{72}$$

where $\mathbf{u}_\mathcal{T}^{eq}$ is defined with $\mathbf{w}_\mathcal{T}^{eq}$ and $V_\mathcal{T}^{eq}$ using the change of variables (35). We establish the exponential decay of $\widetilde{S}^n$.

**Theorem 2** (Exponential decay towards thermal equilibrium)**.** Let $(\mathcal{T},\mathcal{E},\mathcal{P})$ be an admissible mesh of $\Omega$ satisfying the regularity constraint (25). Under Assumptions 1 and 2, there exists a constant $C_{SI} > 0$ only depending on $\|\boldsymbol{\rho}\|_{W^{1,\infty}}$, $\lambda^2$, $u_2^D$, $C_0$ given in (12), $C_V$ defined in (58) and the regularity parameter of the mesh $\xi$ such that for all $n \geq 0$,

$$\widetilde{S}^n \leq \left(1 + \frac{\Delta t}{C_{SI}}\right)^{-n} \widetilde{S}^0. \tag{73}$$

Then for all $k > 0$, if $\Delta t \leq k$, one has

$$\widetilde{S}^n \leq \widetilde{S}^0 e^{-\gamma t^n}, \tag{74}$$

with $\gamma = \log(1+k/C_{SI})/k$. Moreover, there exists a constant $c > 0$ only depending on the data and on $\xi$ such that for all $k > 0$, if $\Delta t \leq k$,

$$\|\mathbf{u}_\mathcal{T}^n - \mathbf{u}_\mathcal{T}^{eq}\|_{L^2}^2 + \|V_\mathcal{M}^n - V_\mathcal{M}^{eq}\|_{1,\mathcal{M}}^2 \leq c\, \widetilde{S}^0 e^{-\gamma t^n} \quad \forall n \geq 0. \tag{75}$$

*Proof.* Throughout this proof, $c_i > 0$ denotes constants only depending on the data. We proceed as in [3, Theorem 3.1]. First of all, following the same computations as in the proofs of propositions 2 and 3, we obtain that for all $n \geq 0$,

$$\widetilde{S}^{n+1} - \widetilde{S}^n \leq -\Delta t \left( \left| w_{1,\mathcal{M}}^{n+1} \right|_{1,\mathcal{M}}^2 + \left| w_{2,\mathcal{M}}^{n+1} \right|_{1,\mathcal{M}}^2 \right) =: -\Delta t\, \mathbb{I}^{n+1}. \tag{76}$$

Then, we now have to establish that there exists a constant $C_{SI} > 0$ such that for all $n \geq 0$,

$$\widetilde{S}^n \leq C_{SI} \mathbb{I}^n. \tag{77}$$

To this end, we introduce

$$\mathbb{F}^n := \sum_{K\in\mathcal{T}} \mathrm{m}(K) |\mathbf{u}_K^n - \mathbf{u}_K^{eq}|^2 - \frac{\lambda^2}{2} u_2^D \sum_{\sigma\in\mathcal{E}} \tau_\sigma |D_\sigma(V_\mathcal{M}^n - V_\mathcal{M}^{eq})|^2, \tag{78}$$

and we prove that

$$\widetilde{S}^n \leq C_{SF} \mathbb{F}^n \leq C_{FI} \mathbb{I}^n.$$

Using that $\boldsymbol{\rho} \in W^{1,\infty}(\mathbb{R}^2)$ and that $\boldsymbol{\rho} = \nabla\chi$ with $\chi$ convex, we have

$$\begin{aligned} \boldsymbol{\rho}(\mathbf{u}^n) \cdot (\mathbf{u}^n - \mathbf{u}^{eq}) - (\chi(\mathbf{u}^n) - \chi(\mathbf{u}^{eq})) &\leq (\boldsymbol{\rho}(\mathbf{u}^n) - \boldsymbol{\rho}(\mathbf{u}^{eq})) \cdot (\mathbf{u}^n - \mathbf{u}^{eq}) \\ &\leq \|\boldsymbol{\rho}\|_{W^{1,\infty}} |\mathbf{u}^n - \mathbf{u}^{eq}|^2. \end{aligned}$$

Then, using the definition (72) of $\widetilde{S}^n$, we immediately obtain that there exists a constant $C_{SF} > 0$ only depending on $\|\boldsymbol{\rho}\|_{W^{1,\infty}}$, $\lambda^2$ and $u_2^D$ such that

$$\widetilde{S}^n \leq C_{SF} \mathbb{F}^n.$$

It remains now to show that there exists a constant $C'_{FI} > 0$ such that $\mathbb{F}^n \leq C'_{FI}\mathbb{I}^n$. Using the $W^{1,\infty}$ bound on $\boldsymbol{\rho}$ together with Cauchy-Schwarz et Young inequalities, we have

$$\begin{aligned}\frac{C_0}{4}\sum_{K\in\mathcal{T}}\mathrm{m}(K)|\mathbf{u}_K^n-\mathbf{u}_K^{eq}|^2+c_1\sum_{K\in\mathcal{T}}\mathrm{m}(K)|\mathbf{w}_K^n-\mathbf{w}_K^{eq}|^2\\ \geq\sum_{K\in\mathcal{T}}\mathrm{m}(K)\left(\boldsymbol{\rho}_K^n-\boldsymbol{\rho}_K^{eq}\right)\cdot(\mathbf{w}_K^n-\mathbf{w}_K^{eq}).\end{aligned}$$

Using the change of variables (35), we have

$$\begin{aligned}&\frac{C_0}{4}\sum_{K\in\mathcal{T}}\mathrm{m}(K)|\mathbf{u}_K^n-\mathbf{u}_K^{eq}|^2+c_1\sum_{K\in\mathcal{T}}\mathrm{m}(K)|\mathbf{w}_K^n-\mathbf{w}_K^{eq}|^2\\ &\geq\sum_{K\in\mathcal{T}}\mathrm{m}(K)\Big((\rho_{1,K}^n-\rho_{1,K}^{eq})(u_{1,K}^n+u_{2,K}^nV_K^n-u_{1,K}^{eq}-u_{2,K}^{eq}V_K^{eq})\\ &\qquad+(\rho_{2,K}^n-\rho_{2,K}^{eq})(u_{2,K}^n-u_{2,K}^{eq})\Big)\\ &\geq\sum_{K\in\mathcal{T}}\mathrm{m}(K)(\boldsymbol{\rho}_K^n-\boldsymbol{\rho}_K^{eq})\cdot(\mathbf{u}_K^n-\mathbf{u}_K^{eq})\\ &\quad+\sum_{K\in\mathcal{T}}\mathrm{m}(K)(\rho_{1,K}^n-\rho_{1,K}^{eq})(u_{2,K}^nV_K^n-u_{2,K}^{eq}V_K^{eq}).\end{aligned}$$

Next using the monotony of $\boldsymbol{\rho}$, see (12), we obtain

$$\begin{aligned}&\frac{C_0}{4}\sum_{K\in\mathcal{T}}\mathrm{m}(K)|\mathbf{u}_K^n-\mathbf{u}_K^{eq}|^2+c_1\sum_{K\in\mathcal{T}}\mathrm{m}(K)|\mathbf{w}_K^n-\mathbf{w}_K^{eq}|^2\\ &\qquad\geq C_0\sum_{K\in\mathcal{T}}\mathrm{m}(K)|\mathbf{u}_K^n-\mathbf{u}_K^{eq}|^2+T_1+T_2,\end{aligned}\tag{79}$$

with

$$\begin{aligned}T_1&=\sum_{K\in\mathcal{T}}\mathrm{m}(K)(\rho_{1,K}^n-\rho_{1,K}^{eq})(u_{2,K}^n-u_{2,K}^{eq})V_K^n,\\ T_2&=u_2^D\sum_{K\in\mathcal{T}}\mathrm{m}(K)(\rho_{1,K}^n-\rho_{1,K}^{eq})(V_K^n-V_K^{eq}).\end{aligned}$$

Using the $L^\infty$ estimate (58) on $V_{\mathcal{T}}^n$ and the Cauchy-Schwarz inequality, we get

$$T_1\geq -C_V\left(\sum_{K\in\mathcal{T}}\mathrm{m}(K)(\rho_{1,K}^n-\rho_{1,K}^{eq})^2\right)^{1/2}\left(\sum_{K\in\mathcal{T}}\mathrm{m}(K)(u_{2,K}^n-u_{2,K}^{eq})^2\right)^{1/2}.$$

Then, since $\boldsymbol{\rho}\in W^{1,\infty}(\mathbb{R}^2)$ and thanks to Young inequality, it yields

$$T_1\geq-\frac{C_0}{2}\sum_{K\in\mathcal{T}}\mathrm{m}(K)|\mathbf{u}_K^n-\mathbf{u}_K^{eq}|^2-c_2\sum_{K\in\mathcal{T}}\mathrm{m}(K)|\mathbf{w}_K^n-\mathbf{w}_K^{eq}|^2.\tag{80}$$

Concerning $T_2$, we have using schemes (33) and (69), and a discrete integration by parts:

$$\begin{aligned} T_2 &= u_2^D \lambda^2 \sum_{K\in\mathcal{T}} \sum_{\sigma\in\mathcal{E}_K} \tau_\sigma D_{K,\sigma}(V_\mathcal{M}^n - V_\mathcal{M}^{eq})(V_K^n - V_K^{eq}) \\ &= -u_2^D \lambda^2 \sum_{\substack{\sigma\in\mathcal{E} \\ (K=K_\sigma)}} \tau_\sigma D_\sigma (V_\mathcal{M}^n - V_\mathcal{M}^{eq})^2. \end{aligned} \tag{81}$$

Finally, plugging equations (80) and (81) in (79), we have

$$-u_2^D\lambda^2 \, |V_\mathcal{M}^n - V_\mathcal{M}^{eq}|_{1,\mathcal{M}}^2 + \frac{C_0}{4} \sum_{K\in\mathcal{T}} \mathrm{m}(K)|\mathbf{u}_K^n - \mathbf{u}_K^{eq}|^2 \le (c_1+c_2) \sum_{K\in\mathcal{T}} \mathrm{m}(K)|\mathbf{w}_K^n - \mathbf{w}_K^{eq}|^2,$$

from which we deduce $\mathbb{F}^n \le C'_{FI}\mathbb{I}^n$ thanks to the discrete Poincaré inequality.

We can now deduce (73) from

$$\widetilde{S}^{n+1} - \widetilde{S}^n \le -\Delta t \mathbb{I}^{n+1} \le -\frac{\Delta t}{C_{SI}} \widetilde{S}^{n+1}.$$

Then, using the fact that $t^n = n\Delta t$ together with the monotony of $x \mapsto \log(1 + x/C_{SI})/x$ on $\mathbb{R}_+^*$, one gets (74) and then (75). □

## 6. Numerical experiments

6.1. **Physical models.** We consider the unipolar energy-transport model under Boltzmann statistics, as in [31, 14], with the following densities $\rho_i(\mathbf{u})$, $i = 1, 2$:

$$\begin{cases} \rho_1(\mathbf{u}) = \left(-\dfrac{1}{u_2}\right)^{3/2} \exp(u_1), \\ \rho_2(\mathbf{u}) = \dfrac{3}{2}\left(-\dfrac{1}{u_2}\right)^{5/2} \exp(u_1). \end{cases} \tag{82}$$

The diffusion matrix $\mathbb{L}(\mathbf{u}) = (L_{ij}(\mathbf{u}))_{1\le i,j\le 2}$ depends on $\mathbf{u}$ with

$$\mathbb{L}(\mathbf{u}) = c_0\rho_1(\mathbf{u})T^{1/2-\beta} \begin{pmatrix} 1 & (2-\beta)T \\ (2-\beta)T & (3-\beta)(2-\beta)T^2 \end{pmatrix}, \tag{83}$$

and the relaxation term $W$ reads:

$$W(\mathbf{u}) = c_1\rho_1(\mathbf{u}) \frac{l_x^2}{\tau_0\mu_0 U_T} T^{\beta-1/2}(1-T), \tag{84}$$

where $T = -1/u_2$ is the temperature. The usual values are $c_0 = 1$, $c_1 = 3/2$, $\beta = 1/2$ for the Chen model, and $c_0 = c_1 = 2/\sqrt{\pi}$, $\beta = 0$ for the Lyumkis model [31]. To sum up, we take for the Chen model

$$\begin{aligned} \mathbb{L}(\mathbf{u}) &= \rho_1(\mathbf{u}) \begin{pmatrix} 1 & 3T/2 \\ 3T/2 & 15T^2/4 \end{pmatrix}, \\ W(\mathbf{u}) &= \left(\frac{3}{2}\frac{l_x^2}{\tau_0\mu_0 U_T}\right)\rho_1(\mathbf{u}) - \left(\frac{l_x^2}{\tau_0\mu_0 U_T}\right)\rho_2(\mathbf{u}), \end{aligned}$$

and for the Lyumkis model

$$\begin{aligned} \mathbb{L}(\mathbf{u}) &= \frac{2}{\sqrt{\pi}}\sqrt{T}\rho_1(\mathbf{u}) \begin{pmatrix} 1 & 2T \\ 2T & 6T^2 \end{pmatrix}, \\ W(\mathbf{u}) &= \left(\frac{2}{\sqrt{\pi}}\frac{1}{\sqrt{T}}\frac{l_x^2}{\tau_0\mu_0 U_T}\right)\rho_1(\mathbf{u}) - \left(\frac{4}{3\sqrt{\pi}}\frac{1}{\sqrt{T}}\frac{l_x^2}{\tau_0\mu_0 U_T}\right)\rho_2(\mathbf{u}). \end{aligned}$$

TABLE 1. Physical parameters.

| Parameter | Physical meaning | Numerical value |
|---|---|---|
| $q$ | elementary charge | $1.6 \times 10^{-19}$ As |
| $n_i$ | intrinsic density | $1.4 \times 10^{10}\text{cm}^{-3}$ |
| $T_0$ | ambient temperature | 300K |
| $\varepsilon$ | permittivity constant | $10^{-12}\text{AsV}^{-1}\text{cm}^{-1}$ |
| $\mu_0$ | low field mobility | $1.5 \times 10^{3}\text{cm}^2\text{V}^{-1}\text{s}^{-1}$ |
| $U_T$ | thermal voltage at $T_0 = 300$K | 0.0259 V |
| $\tau_0$ | energy relaxation time | $0.4 \times 10^{-12}$ s |

For these models, one can check that the matrix $\mathbb{L}(\mathbf{u})$ is symmetric positive definite, as soon as the temperature $T$ is positive. In all this section, we consider the numerical values of the physical parameters for silicon given in Table 1. Finally, the scaling ensures that the Debye length is

$$\lambda^2 = \frac{\varepsilon U_T}{q l_x^2 C_m},$$

where $l_x$ is the characteristic length of the device.

6.2. **Implementation of the whole method.** Since the numerical schemes (31)–(34) and (38)–(42) are fully implicit, they require to solve a nonlinear system at each time iteration. We use a Newton method which stops when either the difference between two successive computed solutions is lower than $10^{-9}$, or the number of iterations of the method is larger than 30.

Although the discretization is fully implicit, it is necessary to use an adaptative time step: for the first iterations, small time steps are needed in order to allow the convergence of the Newton method, while larger time steps are sufficient when moving to the stationary regime. In practice, the minimal time step is $\Delta t_{min} = 10^{-8}$, the maximal one is $\Delta t_{max} = 10^{-3}$. The method starts with $\Delta t = \Delta t_{max}$. At each time iteration, we consider that the whole method has converged if the tolerance of $10^{-9}$ is reached and the quantity $u_{2,K}$ is negative for all $K \in \mathcal{T}$. Otherwise, the current time step is set to $\max(\Delta t/2, \Delta t_{min})$ and Newton method is relaunched until the convergence is achieved. At the next time iteration, the new current time step is set to $\min(2\Delta t, \Delta t_{max})$.

In the sequel, we consider realistic test cases presented in the literature. Let us emphasize that the assumptions, which are essential for the theoretical results, namely the $W^{1,\infty}$ bound on $\boldsymbol{\rho}$ and the boundary conditions at thermal equilibrium, are not satisfied by the considered test cases.

In particular, we investigate the large-time convergence of the numerical solution toward an approximate stationary state, which is not necessary the thermal equilibrium. In what follows, the stationary state is computed setting a final time $T_{max} = 5$. The relative entropy $\tilde{S}^n$ is defined by (72), using this computed steady state. For all the two-dimensional test cases, we illustrate the exponential decay of the relative entropy in time.

6.3. **Test cases.** Subsection 6.3.1 presents a one-dimensional validation test-case of a $n^+nn^+$ silicon diode. It allows to compare the primal and dual schemes (both with the centered (43) and the upwind (44) definitions of the interfacial quantities) in terms of Newton iterations and adaptative time steps. Subsection 6.3.2 is devoted

to the same test case in a two-dimensional domain. Next Subsections 6.3.3 and 6.3.4 present more realistic two-dimensional test cases, namely the MESFET and double-gate MESFET device test cases. All the two-dimensional test cases are performed on a mesh made of 3584 triangles, with final time $T_{max} = 1$, as the steady state is reached.

6.3.1. *Validation test case.* Performances in terms of total numbers of Newton iterations and adaptative time stepping is compared for both dual and primal schemes. We focus on a $n^+nn^+$ silicon diode test case, proposed in [15, 25, 9]. It corresponds to a model for the channel of a MOS transistor.

The domain is $\Omega = (0, l_x)$ with $l_x = 0.6\mu m$. The $n$ region is $(0.1, 0.5)$ and the $n^+$ region is $(0, 0.1) \cup (0.5, 0.6)$.

The doping profile is

$$C = C_m = 5 \times 10^{17} cm^{-3} \quad \text{in the } n^+ \text{ region},$$
$$C = C_n = 2 \times 10^{15} cm^{-3} \quad \text{in the } n \text{ region}.$$

The boundary conditions are

$$V(0) = 1.5\text{V and } V(0.6) = 0,$$
$$u_2(0) = -1/T_0 = u_2(0.6),$$
$$\rho_1(\mathbf{u})(0) = C_m = \rho_1(\mathbf{u})(0.6),$$

the latest giving the boundary condition for $u_1$ according to (82). The initial conditions for $u_1$ and $u_2$ are constant and equal to the Dirichlet boundary conditions.

The comparison between the primal and dual schemes is performed for the Chen model, on an uniform mesh of $(0, l_x)$ made of 200 cells with final time $T_{max} = 1$, as the stationary state is reached. The initial time step is $\Delta t = 10^{-3}$. In order to ensure the Newton method convergence, the time step may be adapted, as explained in Section 6.2.

Table 2 records the number of Newton iterations (including iterations for which the Newton method did not converge), the number of time steps and the number of adapted time steps for the primal and dual schemes, for the Chen and the Lyumkis models. The total amounts of iterations is much more important for the dual scheme, that's why the further simulations are performed with the primal scheme. This phenomenon may be explained by a lack of smoothness of the functions involved in the Newton method for the dual scheme. Obviously the Newton method convergence is not quadratic in this case. It turns out that the choice of variables involved in the Newton method has a strong impact on the performance of the algorithm, as underlined for exemple by [6]. Concerning the impact of the centered and upwind choices, one observes that both provide comparable results for the primal scheme, whereas the centered discretization is slightly more efficient than the upwind one for the dual scheme.

We present in Figures 1 and 2 the distribution of the Newton method iterations in time, for the Chen and Lyumkis models respectively, both for the primal and dual schemes in the centered and upwind configurations. Because the Newton method requires more iterations at the beginning of the computation, the graphs are plotted in semi-log scale in time. The distribution of the iterations is different between the Chen and the Lyumkis models, since a burst in the amounts of Newton iteration appears in the Lyumkis model around $t = 0.1$. Figures 1 and 2 confirm that the primal scheme is more efficient than the dual one.

TABLE 2. 1D-ballistic diode test case: Newton iterations for the primal and dual schemes, number of time steps and adaptative time steps for the Chen model. The performance for the Lyumkis model are indicated within brackets.

| Schemes | | Newton's iterations | Number of time steps | Adaptative time steps |
|---|---|---|---|---|
| Primal | centered | 1630 (1791) | 1011 (1011) | 21 (21) |
| | upwind | 1623 (1868) | 1009 (1010) | 19 (21) |
| Dual | centered | 4481 (4324) | 1106 (1096) | 214 (193) |
| | upwind | 5090 (5256) | 1132 (1122) | 266 (245) |

6.3.2. *Two-dimensional ballistic diode.* We now present the results obtained for the two-dimensional simulation performed with the primal scheme (31)–(34) with the centered definition (43) of the interfacial quantities. The numerical results can be compared with those obtained in [25, 9], and also with the one-dimensional results provided in [19, 11, 15].

The domain is $\Omega = (0, l_x) \times (0, l_y)$ with $l_x = 0.6\mu m$ and $l_y = 0.2\mu m$. The channel length is $0.4\mu m$, see Figure 3.

The doping profile is

$$C = C_m = 5 \times 10^{17} cm^{-3} \quad \text{in the } n^+ \text{ region},$$
$$C = C_n = 2 \times 10^{15} cm^{-3} \quad \text{in the } n \text{ region}.$$

The boundary conditions are

$$V = 1.5\text{V on } \Gamma_1^D \text{ and } V = 0 \text{ on } \Gamma_2^D,$$
$$u_2 = -1/T_0, \text{ with } T_0, \text{ on } \Gamma_1^D \cup \Gamma_2^D,$$
$$\rho_1(\mathbf{u}) = C_m \text{ on } \Gamma_1^D \cup \Gamma_2^D,$$

the latest giving the boundary condition for $u_1$ according to (82). On the other boundary segments, we consider homogeneous Neumann boundary conditions for $J_1$, $J_2$ and $V$. The initial conditions for $u_1$ and $u_2$ are constant and equal to the Dirichlet boundary conditions.

As expected, the computed quantities are almost uniform in one space direction. One observes in Figure 4 the expected hot electron effect in the channel, which compares with the results given in [25, 9]. Moreover, one observes in Figure 5 the exponential decay of the relative entropy and of the error in $L^2$-norm $\|\mathbf{u}_{\mathcal{T}}^n - \mathbf{u}_{\mathcal{T}}^{eq}\|_{L^2} + \|V_{\mathcal{M}}^n - V_{\mathcal{M}}^{eq}\|_{1,\mathcal{M}}$. Considering the estimate (75), it appears that the decay rate of the error in $L^2$-norm is almost the half of the one of the relative entropy $\tilde{S}^n$.

6.3.3. *The MESFET device.* We now consider a MESFET device test case, proposed in [27, 25]. The MESFET device, namely metal-semiconductor field-effect transistor, is used as a switch or amplifier. It consists in two highly-doped $n^+$ regions near the Ohmic contacts, called source and drain, and an $n$ region with a Schottky contact, called gate. The domain is still $\Omega = (0, l_x) \times (0, l_y)$, with $l_x = 0.6\mu m$ and $l_y = 0.2\mu m$. The source and drain contact lengths are $0.1\mu m$ and the gate contact is $0.2\mu m$. The complete geometry is presented in Figure 6.

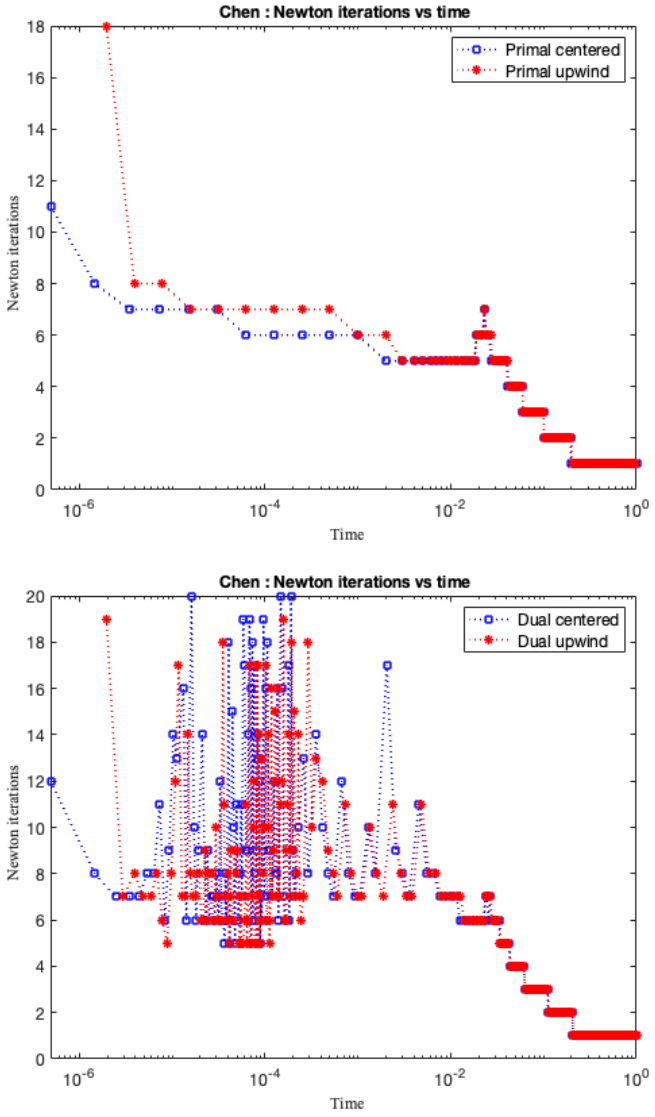



FIGURE 1. Distribution of the Newton iterations in time for the 1-D $n^+nn^+$ diode with the Chen model: primal scheme (above) and dual scheme (below).

The doping profile is

$$C = C_m = 3 \times 10^{17} cm^{-3} \quad \text{in the } n^+ \text{ region},$$
$$C = C_n = 1 \times 10^{15} cm^{-3} \quad \text{in the } n \text{ region}.$$

The boundary conditions are defined as follows:

- at the source S: $\rho_1(\mathbf{u}) = C_m$, $u_2 = -1/T_0$, $V = \Phi_0$,
- at the drain D: $\rho_1(\mathbf{u}) = C_m$, $u_2 = -1/T_0$, $V = \Phi_0 + 2\text{V}$,

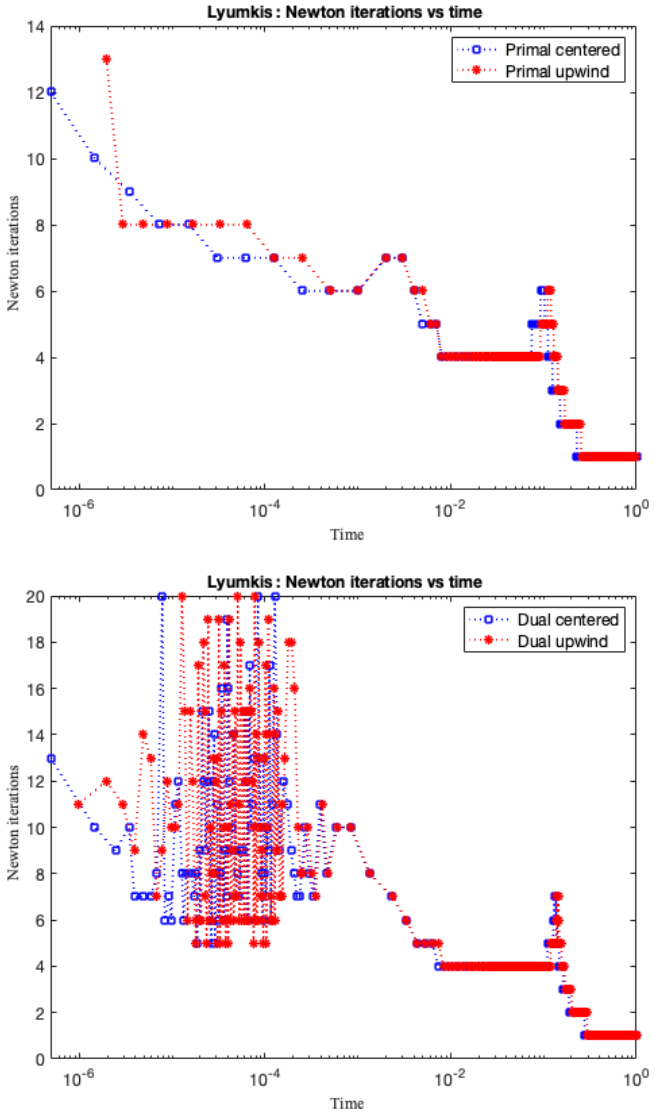



FIGURE 2. Distribution of the Newton iterations in time for the 1-D $n^+nn^+$ diode with the Lyumkis model: primal scheme (above) and dual scheme (below).

- at the gate G: $\rho_1(\mathbf{u}) = 3.9 \times 10^5 \text{cm}^{-3}$, $u_2 = -1/T_0$, $V = \Phi_0 - 0.8\text{V}$,

where $\Phi_0$ is the built-in potential defined as

$$\Phi_0 = U_T \log\left(\frac{\rho_1(\mathbf{u})}{n_i}\right). \tag{85}$$

For the remaining boundary segments, homogeneous boundary conditions for $J_1$, $J_2$ and $V$ are prescribed.

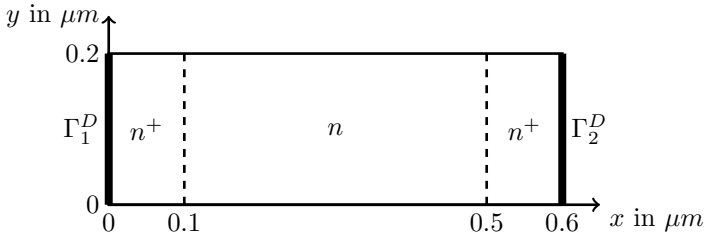



FIGURE 3. Geometry of the $n^+nn^+$ ballistic diode.

One observes in Figure 7 the depletion of the electron density around the gate contact. The temperature is large near the drain contact, with a strong gradient near the gate. As expected, one observes on the electrical potential profile that the electrons are moving from the source to the drain. Figure 8 illustrates the exponential decay towards the steady state, in relative entropy and in $L^2$-norm, in agreement with (74) and (75).

6.3.4. *The double-gate MESFET device.* We finally consider a double-gate MESFET device as in [26]. It consists of two highly doped $n^+$ regions near the Ohmic contacts (source and drain) and an $n$ region with an upper and lower Schottky contact (gate) in a "sandwich" configuration. The domain is $\Omega = (0, l_x) \times (0, l_y)$, with $l_x = 0.6\mu m$ and $l_y = 0.24\mu m$. The source and drain contact lengths are $0.24\mu m$ and the gate contact length is $0.2\mu m$. The length of the low doped $n$ region (channel) is $0.36\mu m$. The full geometry of the double-gate MESFET device is shown in Figure 9.

The doping profile is defined as

$$C = C_m = 3 \times 10^{17} cm^{-3} \quad \text{in the } n^+ \text{ region},$$
$$C = C_n = 1 \times 10^{17} cm^{-3} \quad \text{in the } n \text{ region}.$$

The boundary conditions are given by

- at the source (S): $\rho_1(\mathbf{u}) = C_m$, $u_2 = -1/T_0$, $V = \Phi_0$,
- at the drain (D): $\rho_1(\mathbf{u}) = C_m$, $u_2 = -1/T_0$, $V = \Phi_0 + 2\text{V}$.

Two possible boundary conditions are considered for the gates, which correspond respectively to the open and the closed configurations. The open state corresponds to a large gate voltage with a less significant decrease of the electron density. The closed configuration corresponds to a decreased gate voltage which induces a reduced current. The boundary conditions at the gates (G) are the following:

- open state: $\rho_1(\mathbf{u}) = 3.9 \times 10^5 \text{cm}^{-3}$, $u_2 = -1/T_0$, $V = \Phi_0 - 0.8\text{V}$,
- close state: $\rho_1(\mathbf{u}) = 2.4 \times 10^5 \text{cm}^{-3}$, $u_2 = -1/T_0$, $V = \Phi_0 - 2\text{V}$.

For the remaining boundary segments, homogeneous Neumann boundary conditions for $J_1$, $J_2$ and $V$ are prescribed.

Numerical results for the open configuration are plotted in Figures 10 and 11, the results for the closed state being given in Figures 12 and 13. The numerical results are comparable to those given in [26]. The energy density is larger in the open configuration than in the closed state. One observes also the channel between the gate contacts. Again the exponential convergence towards the steady state is illustrated in Figures 11 and 13 for the open and closed configuration respectively.

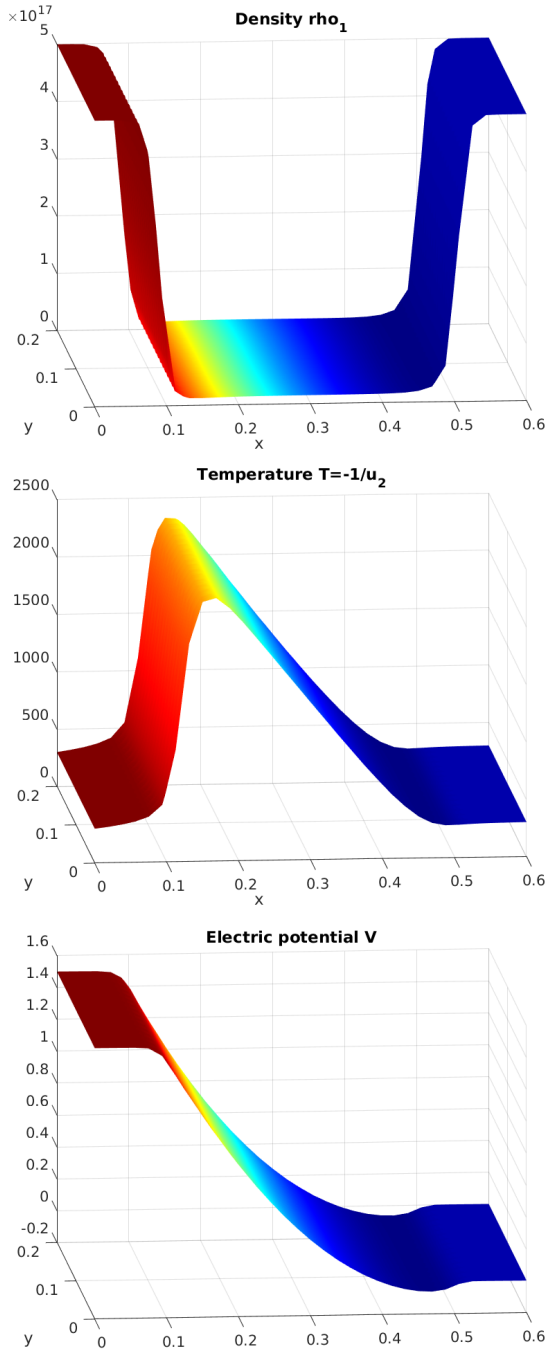



FIGURE 4. 2D-ballistic diode test case with the Chen model. From top to bottom: electron density $\rho_1$, temperature $T = -1/u_2$, electrical potential $V$.

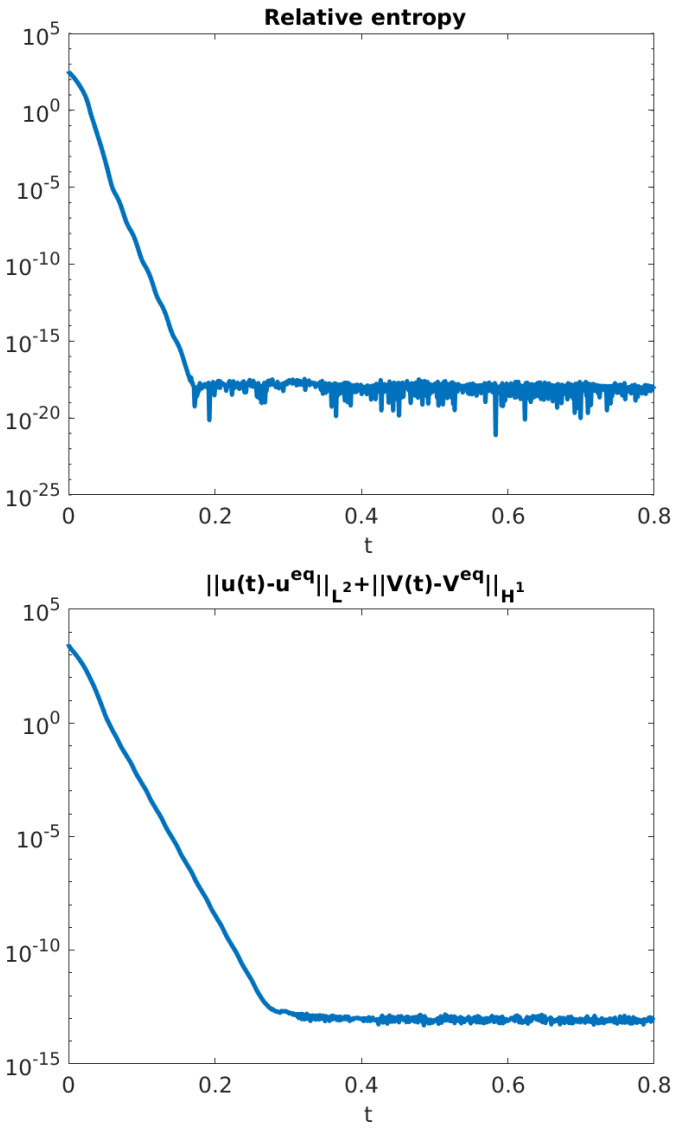



FIGURE 5. 2D-ballistic diode test case: evolution of the relative entropy (top) and of the norm $\|\mathbf{u}_{\mathcal{T}}^n - \mathbf{u}_{\mathcal{T}}^{eq}\|_{L^2} + \|V_{\mathcal{M}}^n - V_{\mathcal{M}}^{eq}\|_{1,\mathcal{M}}$ (bottom) in time.

## 7. Conclusion

We propose in this paper two equivalent TPFA schemes for the energy-transport model. Their definitions strongly rely on the dual entropy variables. Following the approach of [14], we prove a discrete entropy inequality which leads to *a priori* estimates, given in Proposition 3 and in (62)-(63). By means of a Leray-Schauder theorem, one deduces the existence of a solution to the scheme (actually both primal and dual schemes). The convergence of the schemes is not established in

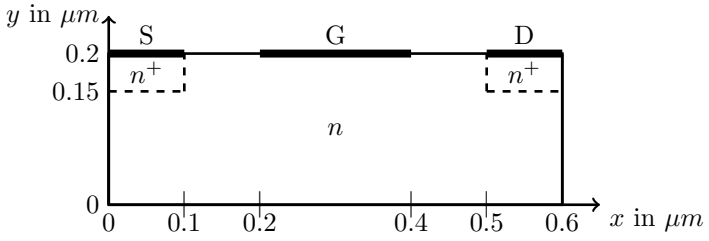



FIGURE 6. Geometry of the MESFET device.

this paper. However, let us mention that the estimates (62)-(63) account for a first step in the proof of compactness of the sequences of approximate solutions. Due to the nonlinearity of the densities, one may use the general result given in [1] to deduce the space-time compactness from the space estimates. Then, up to the identification of the obtained limit as a weak solution of the continuous problem, the proof of convergence of the schemes could be achieved. Similarly to the continuous framework, we prove the exponential decay of the discrete relative entropy towards the thermal equilibrium. Numerical results assess the good behaviour of the whole numerical scheme, although one notices a better efficiency of the primal scheme.

## REFERENCES

[1] B. Andreianov, C. Cancès, and A. Moussa. A nonlinear time compactness result and applications to discretization of degenerate parabolic-elliptic PDEs. *J. Funct. Anal.*, 273(12):3633–3670, 2017.

[2] N. Ben Abdallah and P. Degond. On a hierarchy of macroscopic models for semiconductors. *J. Math. Phys.*, 37(7):3306–3333, 1996.

[3] M. Bessemoulin-Chatard and C. Chainais-Hillairet. Exponential decay of a finite volume scheme to the thermal equilibrium for drift-diffusion systems. *J. Numer. Math.*, 25(3):147–168, 2017.

[4] M. Bessemoulin-Chatard, C. Chainais-Hillairet, and F. Filbet. On discrete functional inequalities for some finite volume schemes. *IMA J. Numer. Anal.*, 35(3):1125–1149, 2015.

[5] A. Bradji and R. Herbin. Discretization of coupled heat and electrical diffusion problems by finite-element and finite-volume methods. *IMA J. Numer. Anal.*, 28(3):469–495, 2008.

[6] K. Brenner and C. Cancès. Improving Newton's method performance by parametrization: the case of the Richards equation. *SIAM J. Numer. Anal.*, 55(4):1760–1785, 2017.

[7] C. Calgaro, C. Colin, and E. Creusé. A combined finite volume - finite element scheme for a low-Mach system involving a Joule term. *AIMS Mathematics*, 5(1):311–331, 2019.

[8] C. Cancès, C. Chainais-Hillairet, J. Fuhrmann, and B. Gaudel. A numerical analysis focused comparison of several finite volume schemes for a unipolar degenerate drift-diffusion model. Preprint, https://hal.archives-ouvertes.fr/hal-02194604v3/document, 2020.

[9] C. Chainais-Hillairet. Discrete duality finite volume schemes for two-dimensional drift-diffusion and energy-transport models. *Internat. J. Numer. Methods Fluids*, 59(3):239–257, 2009.

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

[11] Claire Chainais-Hillairet and Yue-Jun Peng. Finite volume scheme for semiconductor energy-transport model. In *Elliptic and parabolic problems*, volume 63 of *Progr. Nonlinear Differential Equations Appl.*, pages 139–146. Birkhäuser, Basel, 2005.

[12] L. Chen and L. Hsiao. The solution of Lyumkis energy transport model in semiconductor science. *Math. Methods Appl. Sci.*, 26(16):1421–1433, 2003.

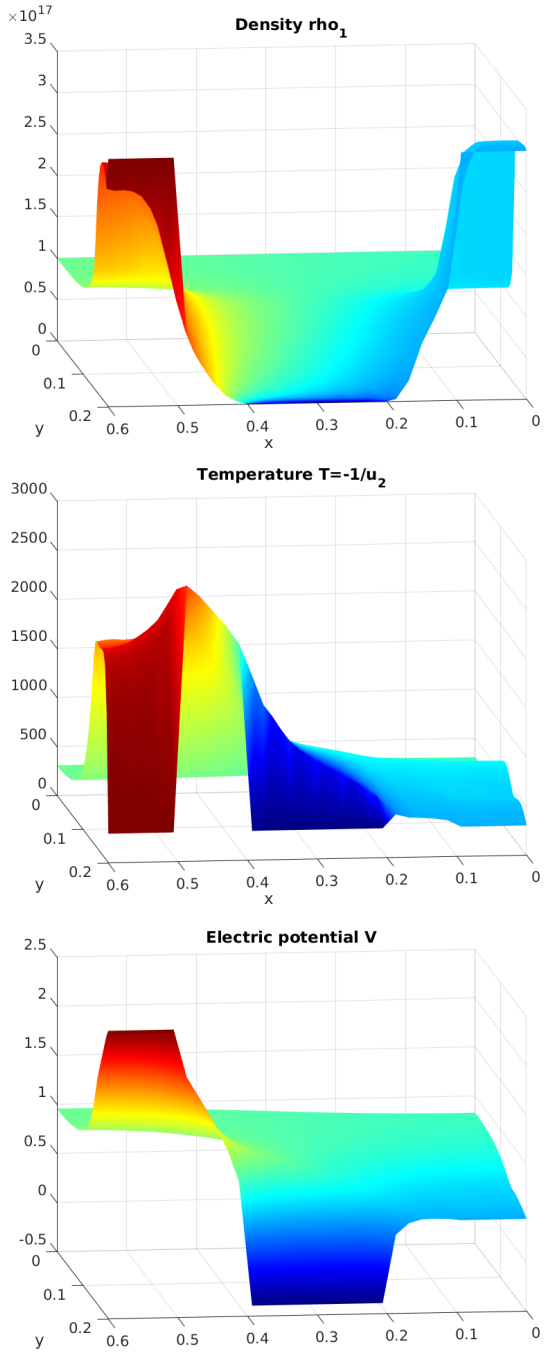



FIGURE 7. MESFET test case (from top to bottom): electron density $\rho_1$, temperature $T = -1/u_2$, electrical potential $V$.

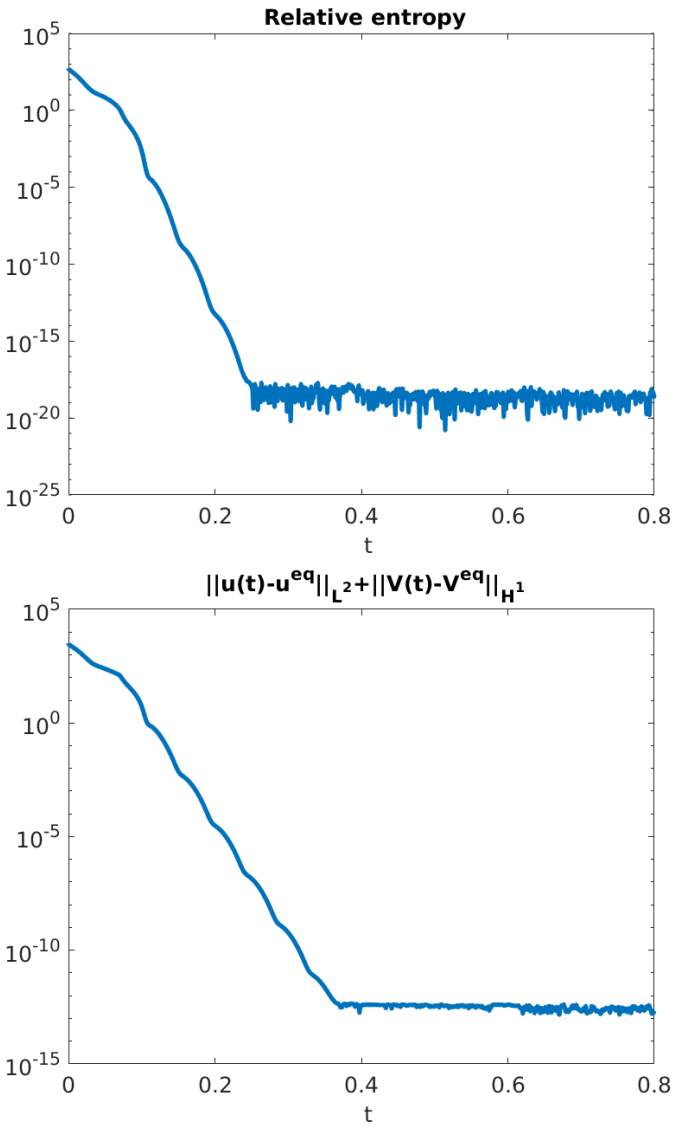



FIGURE 8. MESFET test case: evolution of the relative entropy (top) and of the norm $\|\mathbf{u}_{\mathcal{T}}^n - \mathbf{u}_{\mathcal{T}}^{eq}\|_{L^2} + \|V_{\mathcal{M}}^n - V_{\mathcal{M}}^{eq}\|_{1,\mathcal{M}}$ (bottom) in time.

[13] L. Chen, L. Hsiao, and Y. Li. Large time behavior and energy relaxation time limit of the solutions to an energy transport model in semiconductors. *J. Math. Anal. Appl.*, 312(2):596–619, 2005.

[14] P. Degond, S. Génieys, and A. Jüngel. A system of parabolic equations in nonequilibrium thermodynamics including thermal and electrical effects. *J. Math. Pures Appl. (9)*, 76(10):991–1015, 1997.

[15] P. Degond, A. Jüngel, and P. Pietra. Numerical discretization of energy-transport models for semiconductors with nonparabolic band structure. *SIAM J. Sci. Comput.*, 22(3):986–1007, 2000.

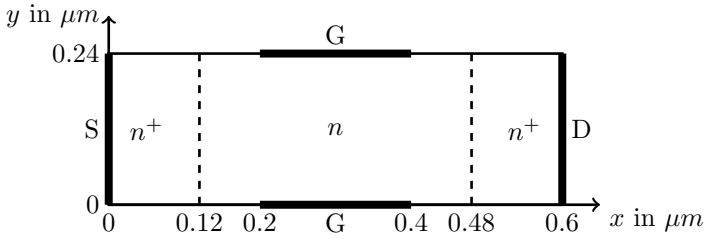



FIGURE 9. Geometry of the double-gate MESFET device.


[16] D. H. Doan, A. Fischer, J. Fuhrmann, A. Glitzky, and M. Liero. Drift-diffusion simulation of s-shaped current-voltage relations for organic semiconductor devices. Working paper or preprint, http://www.wias-berlin.de/preprint/2630/wias_preprints_2630.pdf, 2019.

[17] R. Eymard, T. Gallouët, and R. Herbin. Finite volume methods. In *Handbook of numerical analysis*, volume VII, pages 713–1020. North-Holland, Amsterdam, 2000.

[18] W. Fang and K. Ito. Existence of stationary solutions to an energy drift-diffusion model for semiconductor devices. *Math. Models Methods Appl. Sci.*, 11(5):827–840, 2001.

[19] M. Fournié. Numerical discretization of energy-transport model for semiconductors using high-order compact schemes. *Appl. Math. Lett.*, 15(6):721–726, 2002.

[20] S. Gadau and A. Jüngel. A three-dimensional mixed finite-element approximation of the semiconductor energy-transport equations. *SIAM J. Sci. Comput.*, 31(2):1120–1140, 2008/09.

[21] H. Gajewski and K. Gröger. On the basic equations for carrier transport in semiconductors. *J. Math. Anal. Appl.*, 113(1):12–35, 1986.

[22] H. Gajewski and K. Gröger. Semiconductor equations for variable mobilities based on Boltzmann statistics or Fermi-Dirac statistics. *Math. Nachr.*, 140:7–36, 1989.

[23] Ingenuin Gasser and Roberto Natalini. The energy transport and the drift diffusion equations as relaxation limits of the hydrodynamic model for semiconductors. *Quart. Appl. Math.*, 57(2):269–282, 1999.

[24] J. A. Griepentrog. An application of the implicit function theorem to an energy model of the semiconductor theory. *ZAMM Z. Angew. Math. Mech.*, 79(1):43–51, 1999.

[25] S. Holst, A. Jüngel, and P. Pietra. A mixed finite-element discretization of the energy-transport model for semiconductors. *SIAM J. Sci. Comput.*, 24(6):2058–2075, 2003.

[26] S. Holst, A. Jüngel, and P. Pietra. An adaptive mixed scheme for energy-transport simulations of field-effect transistors. *SIAM J. Sci. Comput.*, 25(5):1698–1716, 2004.

[27] J. W. Jerome and C.-W. Shu. Energy models for one-carrier transport in semiconductor devices. In *Semiconductors*, pages 185–207. Springer, 1994.

[28] A. Jüngel. Regularity and uniqueness of solutions to a parabolic system in nonequilibrium thermodynamics. *Nonlinear Anal.*, 41(5-6, Ser. A: Theory Methods):669–688, 2000.

[29] A. Jüngel. *Transport equations for semiconductors*, volume 773 of *Lecture Notes in Physics*. Springer-Verlag, Berlin, 2009.

[30] A. Jüngel, R. Pinnau, and E. Röhrig. Existence analysis for a simplified transient energy-transport model for semiconductors. *Math. Methods Appl. Sci.*, 36(13):1701–1712, 2013.

[31] Ansgar Jüngel. *Quasi-hydrodynamic semiconductor equations*, volume 41 of *Progress in Nonlinear Differential Equations and their Applications*. Birkhäuser Verlag, Basel, 2001.

[32] M. Kantner. Generalized Scharfetter–Gummel schemes for electro-thermal transport in degenerate semiconductors using the Kelvin formula for the Seebeck coefficient. *J. Comput. Phys.*, 402:109091, 2020.

[33] V. Romano. 2D numerical simulation of the MEP energy-transport model with a finite difference scheme. *J. Comput. Phys.*, 221(2):439–468, 2007.

[34] W. Van Roosbroeck. Theory of the flow of electrons and holes in germanium and other semiconductors. *Bell System Tech. J.*, 29:560–607, 1950.

[35] N. Zamponi and A. Jüngel. Global existence analysis for degenerate energy-transport models for semiconductors. *J. Differential Equations*, 258(7):2339–2363, 2015.

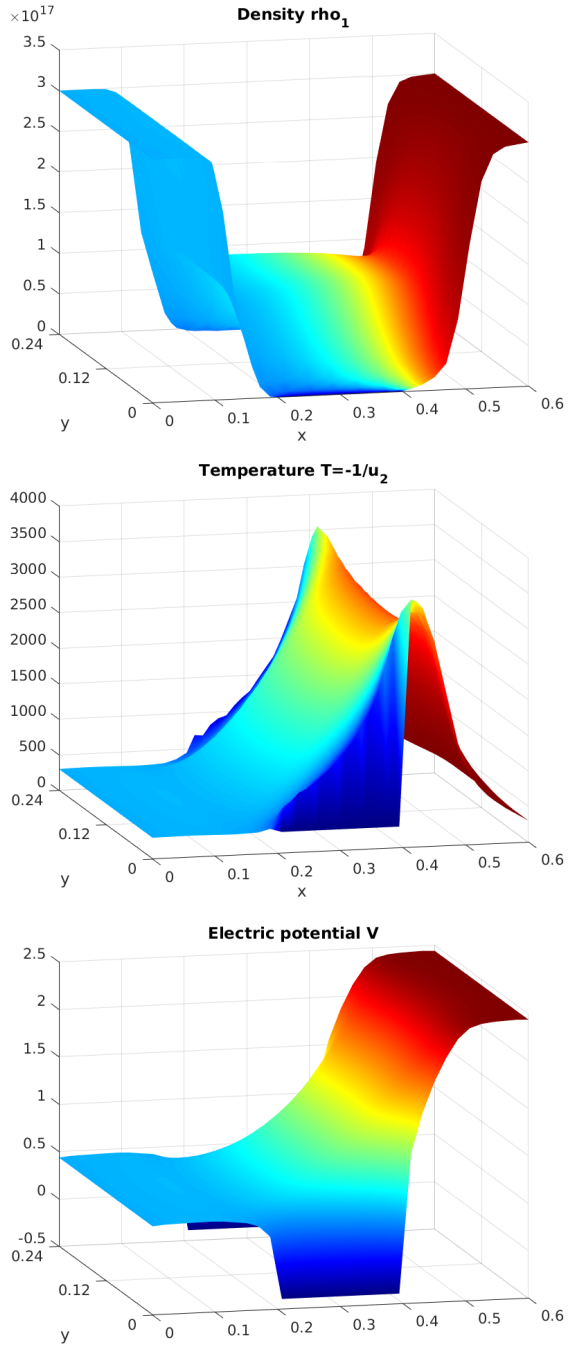



FIGURE 10. Double MESFET test case in the open configuration (from top to bottom): electron density $\rho_1$, temperature $T = -1/u_2$, electrical potential $V$.

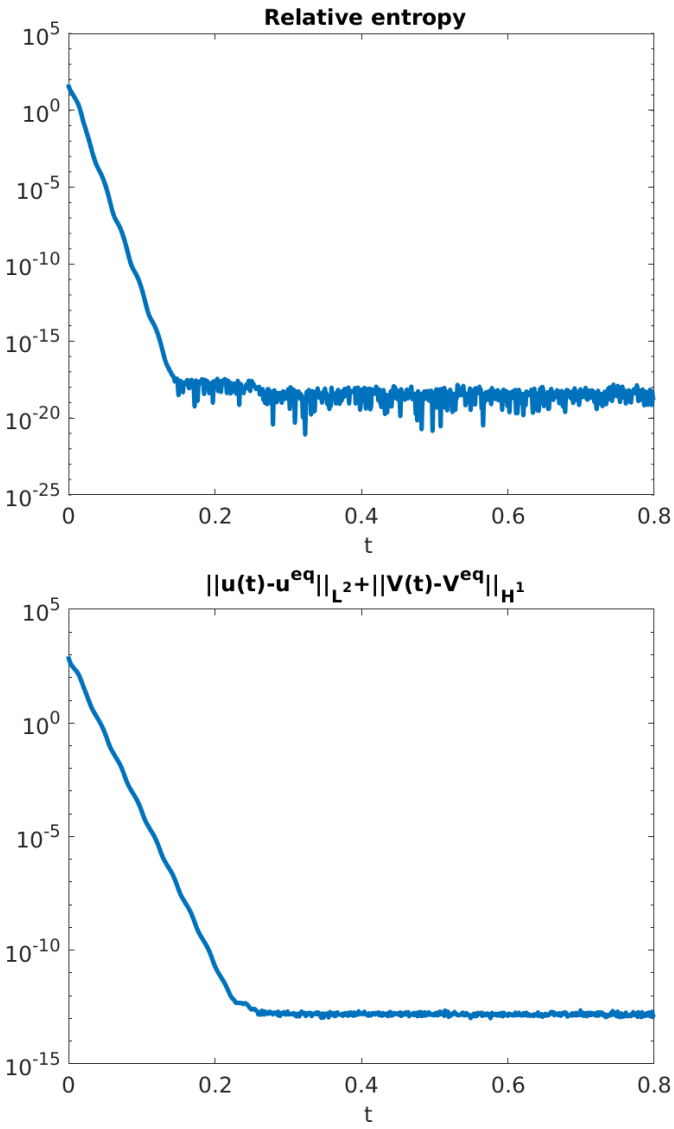



FIGURE 11. Double MESFET test case in the open configuration: evolution of the relative entropy (top) and of the norm $\|\mathbf{u}_{\mathcal{T}}^n - \mathbf{u}_{\mathcal{T}}^{eq}\|_{L^2} + \|V_{\mathcal{M}}^n - V_{\mathcal{M}}^{eq}\|_{1,\mathcal{M}}$ (bottom) in time.

LABORATOIRE DE MATHÉMATIQUES JEAN LERAY, UNIVERSITÉ DE NANTES & CNRS UMR 6629, BP 92208, F-44322 NANTES CEDEX 3, FRANCE

UNIV. LILLE, CNRS, UMR 8524, INRIA - LABORATOIRE PAUL PAINLEVÉ, F-59000 LILLE, FRANCE.

LABORATOIRE DE MATHÉMATIQUES JEAN LERAY, UNIVERSITÉ DE NANTES & CNRS UMR 6629, BP 92208, F-44322 NANTES CEDEX 3, FRANCE

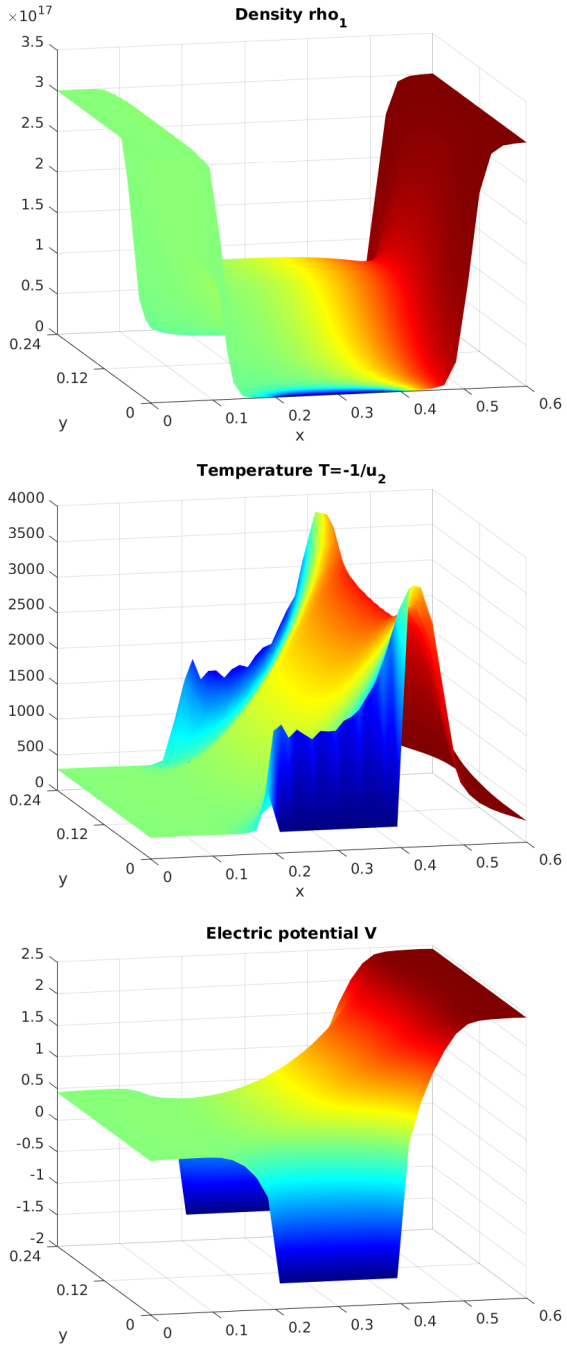



FIGURE 12. Double MESFET test case in the closed configuration (from top to bottom): electron density $\rho_1$, temperature $T = -1/u_2$, electrical potential $V$.

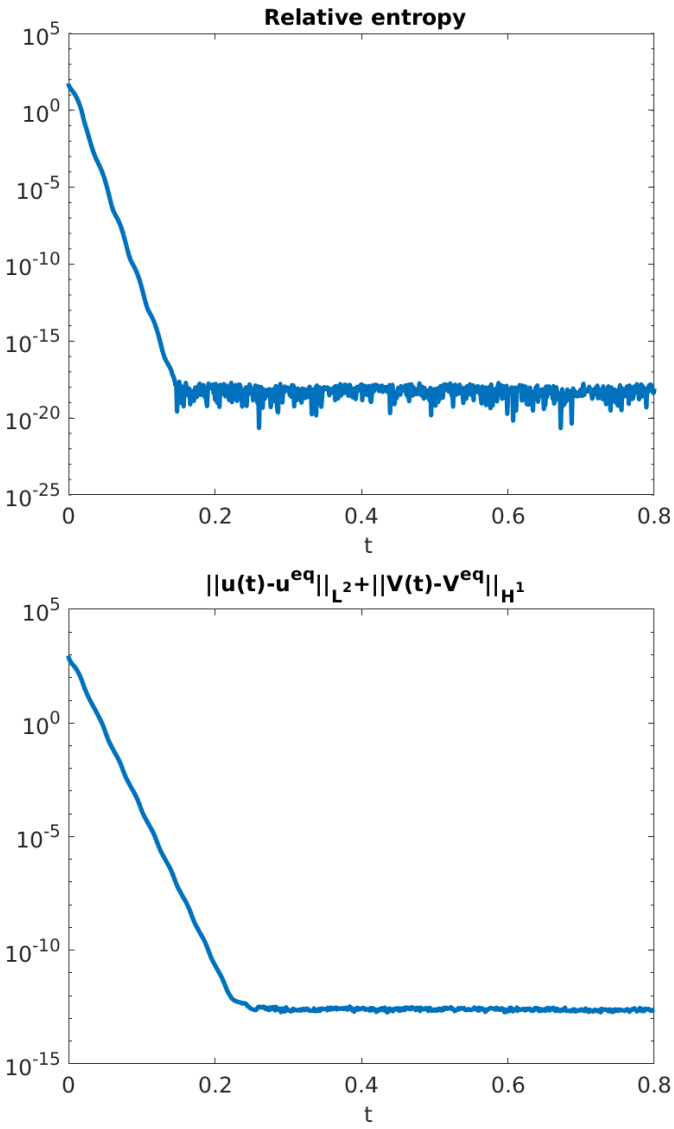



FIGURE 13. Double MESFET test case in the closed configuration: evolution of the relative entropy (top) and of the norm $\|\mathbf{u}_{\mathcal{T}}^n - \mathbf{u}_{\mathcal{T}}^{eq}\|_{L^2} + \|V_{\mathcal{M}}^n - V_{\mathcal{M}}^{eq}\|_{1,\mathcal{M}}$ (bottom) in time.