ESAIM: M2AN 48 (2014) 1701–1724
DOI: 10.1051/m2an/2014016

ESAIM: Mathematical Modelling and Numerical Analysis
www.esaim-m2an.org

HIGHLY ANISOTROPIC NONLINEAR TEMPERATURE BALANCE EQUATION AND ITS NUMERICAL SOLUTION USING ASYMPTOTIC-PRESERVING SCHEMES OF SECOND ORDER IN TIME *

Alexei Lozinski[1,2], Jacek Narski[1] and Claudia Negulescu[1]

Abstract. This paper deals with the numerical study of a nonlinear, strongly anisotropic heat equation. The use of standard schemes in this situation leads to poor results, due to the high anisotropy. An Asymptotic-Preserving method is introduced in this paper, which is second-order accurate in both, temporal and spacial variables. The discretization in time is done using an L-stable Runge−Kutta scheme. The convergence of the method is shown to be independent of the anisotropy parameter $0 < \varepsilon < 1$, and this for fixed coarse Cartesian grids and for variable anisotropy directions. The context of this work are magnetically confined fusion plasmas.

Mathematics Subject Classification. 65N30, 65Z05, 35K20.

Received April 20, 2012. Revised February 14, 2014.
Published online September 26, 2014.

1. Introduction

Magnetically confined plasmas are characterized by highly anisotropic properties induced by the applied strong magnetic field. Indeed, the charged particles constituting the plasma move rapidly around the magnetic field lines, their transverse motion away from the field lines being constrained by the Lorentz force. In contrast, their motion along the field lines is relatively unconstrained, so that rather rapid dynamics along the magnetic fields occurs. This results in an extremely large ratio of the parallel to the transverse thermal conductivities, as well as of other parameters characterizing the plasma evolution.

A prototype simplified model for the heat diffusion in a magnetically confined plasma can be expressed by the following nonlinear, degenerate parabolic equation

$$\partial_t u - \nabla_{||} \cdot (\kappa_{||}(u)\nabla_{||}u) - \nabla_\perp \cdot (\kappa_\perp \nabla_\perp u) = 0, \tag{1.1}$$

Keywords and phrases. Anisotropic parabolic equation, Ill-conditioned problem, singular perturbation model, limit model, asymptotic preserving scheme.

* *This work has been partially supported by the ANR project ESPOIR (Edge Simulation of the Physics Of Iter Relevant turbulent transport, 2009-2013) and the ANR project BOOST (Building the future Of numerical methOdS for iTer, 2010-2014).*

[1] Université de Toulouse, UPS, INSA, UT1, UTM, Institut de Mathématiques de Toulouse, 118 route de Narbonne, 31062 Toulouse, France. jacek.narski@math.univ-toulouse.fr; claudia.negulescu@math.univ-toulouse.fr

[2] Laboratoire de Mathematiques CNRS UMR 6623, Université de Franche-Comté, 16 route de Gray, 25030 Besançon cedex, France. alexei.lozinski@univ-fcomte.fr

Article published by EDP Sciences

where the subscripts $||$ (resp. $\perp$) refer to the direction parallel (resp. perpendicular) to the magnetic field lines and u designates the temperature. In writing out the equation above we have ignored some important physical phenomena coming from convection and turbulence. Nevertheless, our equation contains some important features inherited from the full model that lead to substantial difficulties in the numerical treatment of both the full model and our simplified one. The diffusion in the direction perpendicular to the magnetic field lines is usually slow since the charged particles move mostly along the field lines. The corresponding diffusion coefficient $\kappa_\perp$ can be taken temperature independent. On the other hand, the coefficient describing the diffusion in the direction parallel to the magnetic field lines, $\kappa_{||}$, is normally much larger and strongly temperature dependent. It can be described by the Spitzer-Härm law $\kappa_{||}(u) = \kappa_0 u^{5/2}$ [28]. Moreover, plasma temperatures are extremely high, so that this diffusion coefficient can become very big. Passing to non-dimensional variables, we shall write therefore the law for $\kappa_{||}$ as

$$\kappa_{||}(u) = \frac{1}{\varepsilon} u^{5/2},$$

where ε is a small parameter, $0 < \varepsilon \ll 1$. An accurate resolution of the parallel and perpendicular diffusion processes plays a crucial role in the understanding of the plasma dynamics and the energy transport phenomena. It is therefore very important to develop and to study efficient numerical schemes to solve problem (1.1). It is also desirable to have a scheme that works robustly for all values of ε from $\varepsilon \ll 1$ to $\varepsilon \sim O(1)$ since this parameter enters the equation in combination with a non-linear term so that the effective value of the diffusion coefficient can vary strongly over the computation domain following the variations in u. This is the primary motivation of the present work.

Anisotropic, nonlinear diffusion equations of the type (1.1) arise in several other fields of application and a lot of efforts were made to construct efficient numerical methods for this challenging problem. To mention some examples, such non-linear evolution equations of parabolic type occur in the description of isentropic gas flows through a porous media [2] or in the description of transport phenomena in heterogeneous geologic formations, such as fractured rock systems [5], which are of fundamental interest for petroleum or groundwater engineering. In addition, these equations appear also in image processing, related to the elimination of noise and small-scale details from an image [3, 20, 27] or in the description of the anisotropic water diffusion in tissues of the nervous system [4].

From a numerical point of view, problems of the type (1.1) are very challenging, as one deals with singularly perturbed problems, the model changing its type in the limit $\varepsilon \to 0$. Standard schemes suffer from the presence of very ill conditionned matrices (typically with a condition number of order $1/(\varepsilon h^2)$ where h is the discretization step in space). Solving an equation with such a matrix on a computer accumulates the rounding errors and may lead to completely wrong results. Note that this drawback cannot be overcome by a mesh refinement since it results only in worsening the condition numbers of the matrices in the discretized problem.

Several methods were investigated in literature to cope with this type of anisotropic problems, using for example high order finite element schemes [11], preconditioned conjugate gradient methods in a mixed spectral/finite difference scheme [17] or introducing an artificial "sound" method, to represent the fast thermal equilibrium along the field lines [19]. All these methods however are rather involved and moreover their range of application is limited, as they are efficient only until a threshold value for ε, and cannot thus recover the limit regime $\varepsilon \to 0$. Another class of employed numerical methods are hybrid strategies, which consist in coupling different numerical schemes valid in different regions of the domain. For example in this case, one can couple the resolution of the singular perturbation problem there where $\varepsilon \sim \mathcal{O}(1)$ with the resolution of a limit problem for $\varepsilon \ll 1$. These methods suffer however from the fact that the coupling conditions between the two models are hard to establish and the interface between the two regions difficult to localize.

The objective of the present paper is to introduce an efficient numerical scheme based on the Asymptotic-Preserving methodology, which allows for an accurate resolution of the singularly perturbed problem, uniformly in ε, with little additional computational cost, and using an ε-independent grid which is not necessarily aligned with the magnetic field, so that one can exploit simple Cartesian grids, for example. Initially, AP-techniques were introduced in [14], to deal with singularly perturbed kinetic models. The key idea is to reformulate the singularly perturbed problem into an equivalent problem, which is however well-posed if we set $\varepsilon = 0$ there. The reformulation of the here proposed method is based, similarly as in [18], on introducing a new auxiliary variable, as proposed earlier in an elliptic framework in [7], and replacing the terms of the equation multiplied by $1/\varepsilon$ by the new terms with an $O(1)$ factor. From a numerical point of view, this procedure means transforming the ill-conditioned problem (with condition number $\sim 1/\varepsilon$ on a fixed mesh), into a well-conditioned one, which switches automatically on a given mesh in space in time from the singularly perturbed problem to the limit problem when $\varepsilon \to 0$. In this manner, the mesh sizes τ and h (in, respectively, time and space) can be freely chosen in accordance with the physical phenomenon one is interested in, and are not constrained by the most rapid dynamics induced by ε.

The difference between the method presented in [18] lies first in the treatment of the non-linearity. Instead of fixed point iterations used to approach the nonlinearity $(u^{n+1})^{5/2}$, we choose to implement a much simpler linear extrapolation method (see Sect. 3.2 for details). Moreover, we develop here a robust asymptotic-preserving scheme of second order in time, which has no analogue in the existing literature, to the best of our knowledge. Finally, this paper contains also a detailed mathematical study of the problem.

The paper is organized as follows: Section 2 contains a description of the problem completed by a mathematical study. In Section 3, we present the numerical method based on an asymptotic preserving space discretization and develop three different time-discretizations: implicit Euler, Crank−Nicolson and L-stable Runge−Kutta methods. Finally, in Section 4 we present some numerical results, focusing on the AP-property of the schemes.

2. Description of the problem and mathematical study

We consider a two or three dimensional anisotropic, nonlinear heat problem, given on a sufficiently smooth, bounded domain $\Omega \subset \mathbb{R}^d$, $d = 2, 3$ with boundary Γ. The direction of the anisotropy is defined by the time-independent vector field $b \in (C^\infty(\Omega))^d$, satisfying $|b(x)| = 1$ for all $x \in \Omega$.

Given this vector field b, one can decompose now vectors $v \in \mathbb{R}^d$, gradients $\nabla\phi$, with $\phi(x)$ a scalar function, and divergences $\nabla \cdot v$, with $v(x)$ a vector field, into a part parallel to the anisotropy direction and a part perpendicular to it. These parts are defined as follows:

$$\begin{aligned}
&v_{||} := (v \cdot b)b, v_\perp := (Id - b \otimes b)v, && \text{so that} \quad v = v_{||} + v_\perp,\\
&\nabla_{||}\phi := (\nabla\phi)_{||}, \nabla_\perp\phi := (\nabla\phi)_\perp, && \text{so that} \quad \nabla\phi = \nabla_{||}\phi + \nabla_\perp\phi,\\
&\nabla_{||} \cdot v := \nabla \cdot v_{||}, \nabla_\perp \cdot v := \nabla \cdot v_\perp, && \text{so that} \quad \nabla \cdot v = \nabla_{||} \cdot v + \nabla_\perp \cdot v,
\end{aligned} \tag{2.1}$$

where we denoted by $\otimes$ the vector tensor product.

The boundary Γ can be decomposed into three components following the sign of the intersection with b:

$$\Gamma_{||} := \{x \in \Gamma \;/\; b(x) \cdot n(x) = 0\},$$

$$\Gamma_{\text{in}} := \{x \in \Gamma \;/\; b(x) \cdot n(x) < 0\}, \quad \Gamma_{\text{out}} := \{x \in \Gamma \;/\; b(x) \cdot n(x) > 0\},$$

and $\Gamma_\perp = \Gamma_{\text{in}} \cup \Gamma_{\text{out}}$. The vector n is here the unit outward normal on Γ.

With these notations we can now introduce the mathematical problem, we are interested to study. We are searching for the particle (ions or electrons) temperature $u(t, x)$, solution of the evolution equation

$$(P) \quad \begin{cases}
\partial_t u - \frac{1}{\varepsilon}\nabla_{||} \cdot (A_{||} u^{5/2} \nabla_{||} u) - \nabla_\perp \cdot (A_\perp \nabla_\perp u) = 0, \quad \text{in} \quad [0, T] \times \Omega,\\
\frac{1}{\varepsilon} n_{||} \cdot (A_{||} u^{5/2}(t, \cdot) \nabla_{||} u(t, \cdot)) + n_\perp \cdot (A_\perp \nabla_\perp u(t, \cdot)) = -\gamma\, u(t, \cdot), \quad \text{on} \quad [0, T] \times \Gamma_\perp,\\
\nabla_\perp u(t, \cdot) = 0, \quad \text{on} \quad [0, T] \times \Gamma_{||},\\
u(0, \cdot) = u^0(\cdot), \quad \text{in} \quad \Omega.
\end{cases} \tag{2.2}$$

The coefficient γ is zero for electrons and $\gamma > 0$ for ions [25,28]. The problem (2.2) describes the diffusion of an initial temperature u^0 within the time interval $[0,T]$ and its outflow through the boundary $\Gamma_\perp$. The parameter $0 < \varepsilon \ll 1$ can be very small and is responsible for the high anisotropy of the problem. We shall suppose, all along this paper, that the coefficients $A_{||}$ and $A_\perp$ are of the same order of magnitude, satisfying

Assumption 2.1. Let $\Gamma_\perp$ consist of two connected components $\Gamma_{\text{in}} = \{x \in \Gamma / n \cdot b < 0\}$ and $\Gamma_{\text{out}} = \{x \in \Gamma / n \cdot b > 0\}$ such that $n = -b$ (resp. $n = b$) on Γ_{in} (resp. on Γ_{out}). All the components Γ_{in}, Γ_{out} and $\Gamma_{||}$ are sufficiently smooth. We suppose moreover $0 < \varepsilon \leq 1$ and $\gamma \geq 0$ fixed. The diffusion coefficients $A_\| \in W^{1,\infty}(\bar{\Omega})$ and $A_\perp \in \mathbb{M}_{d\times d}(W^{1,\infty}(\bar{\Omega}))$ are supposed to satisfy

$$0 < A_0 \leq A_\| \leq A_1, \quad \text{on} \;\; \Omega, \tag{2.3}$$

$$A_0 ||v||^2 \leq v^t A_\perp v \leq A_1 ||v||^2, \quad \forall v \in \mathbb{R}^d \;\text{ and everywhere on } \;\; \Omega, \tag{2.4}$$

with some positive constants $A_0 \leq A_1$.

Putting formally $\varepsilon = 0$ in (2.2) leads to the following ill-posed problem, admitting infinitely many solutions

$$\begin{cases} -\nabla_{||} \cdot (A_{||} u^{5/2} \nabla_{||} u) = 0, \quad \text{in} \quad [0,T] \times \Omega, \\ n_{||} \cdot (A_{||} u^{5/2}(t,\cdot) \nabla_{||} u(t,\cdot)) = 0, \quad \text{on} \quad [0,T] \times \Gamma_\perp, \\ \nabla_\perp u(t,\cdot) = 0, \quad \text{on} \quad [0,T] \times \Gamma_{||}, \\ u(0,\cdot) = u^0(\cdot), \quad \text{in} \quad \Omega. \end{cases} \tag{2.5}$$

Indeed, all functions which are constant along the field lines, meaning $\nabla_{||} u \equiv 0$, and satisfying moreover the boundary condition on $\Gamma_{||}$, are solutions of this problem. From a numerical point of view, this ill-posedness in the limit $\varepsilon \to 0$ can be detected by the fact, that trying to solve (2.2) with standard schemes leads to a linear system, which is very ill-conditioned for $0 < \varepsilon \ll 1$, in particular with a condition number of the order of $1/\varepsilon$.

The aim of this paper will be to introduce an efficient numerical method, permitting to solve (2.2) accurately on a coarse Cartesian grid, which has not to be adapted to the field lines of b and whose mesh size is independent of the value of ε. The here proposed scheme belongs to the category of Asymptotic-Preserving schemes, meaning they are stable independently of the small parameter ε and consistent with the limit problem, if ε tends to zero. The construction of the here developed AP-scheme is an adaptation of a method introduced by the authors in an elliptic framework (see [7]), to the here considered non-linear and time-dependent problem, and is based on a reformulation of the singularly perturbed problem (2.2) into an equivalent problem, which appears to be well-posed in the limit $\varepsilon \to 0$. But before introducing the AP-approach, we will start by studying in the following section the mathematical properties of problem (2.2). The test configuration chosen all along this paper is the diffusion of an initial temperature hot spot (see Fig. 1) along arbitrary magnetic field lines b.

2.1. Mathematical properties

Before starting with the presentation of the AP numerical scheme, let us first check some properties of the diffusion problem (2.2) for fixed $\varepsilon > 0$. For notational simplicity, we consider a slightly more general form of problem (P)

$$(P_m) \quad \begin{cases} \partial_t u - \nabla_{||} \cdot (A_{||} |u|^{m-1} \nabla_{||} u) - \nabla_\perp \cdot (A_\perp \nabla_\perp u) = 0, \quad \text{in} \quad [0,T] \times \Omega, \\ A_{||} |u|^{m-1} n_{||} \cdot \nabla_{||} u + A_\perp n_\perp \cdot \nabla_\perp u = -\gamma\, u, \quad \text{on} \quad [0,T] \times \Gamma_\perp, \\ \nabla_\perp u = 0, \quad \text{on} \quad [0,T] \times \Gamma_{||}, \\ u(0,\cdot) = u^0(\cdot), \quad \text{in} \quad \Omega, \end{cases} \tag{2.6}$$

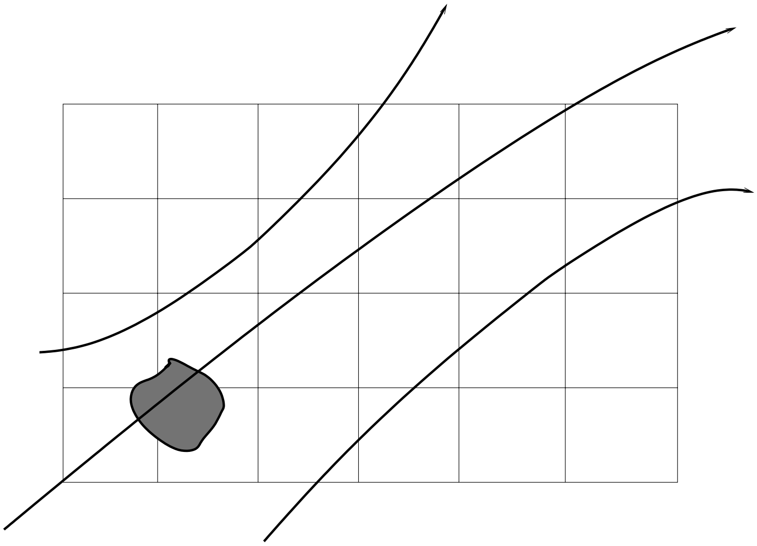


FIGURE 1. Diffusion of a hot temperature spot along the magnetic field lines.

for any $m \geq 1$. We obtain the particular case (2.2) by setting $m = 5/2 + 1$ and redefining $A_{||}$ as $\frac{1}{\varepsilon} A_{||}$ for any $\varepsilon > 0$. Equations of the type (2.6) are rather well studied in the literature. We refer to the classical works [8,9,15] as well as to the more modern literature on "The porous medium equation" as reviewed in [1,26]. However, all these references normally treat only an isotropic version of the problem above, *i.e.* the non-linearity of the type u^{m-1} is present in front of all the derivatives of u. An anisotropic equation of the form (2.6) is studied in [13], but only in the case $m < \frac{d+1}{d-1}$, so that the value of m pertinent to our application is not covered. Another feature of our setting, which is not sufficiently covered in the existing literature, is the prescription of Robin boundary conditions. This is the reason why we wish to study in this paper the existence, uniqueness and positivity of solutions to (2.6).

Remark 2.2. The main difficulty of the theoretical analysis presented in this Section is the possible *lack of diffusion* parallel to b, when u is close to 0. We shall show that u actually remains positive, and even bounded from below by a positive constant, provided this is the case at the initial time. On the other hand, the main difficulty in a numerical simulation of problem (2.2) or in doing a numerical analysis for this problem would be exactly the opposite: it is the *strong diffusion* parallel to b caused by the factor $1/\varepsilon$, leading to ill-conditionned discretizations. Unfortunately, the theoretical analysis in this Section does not allow us to study the behavior of the solutions in the limit $\varepsilon \to 0$. Indeed, the cornerstone of our analysis is the construction of some bounds from below for the weak solutions (*cf.* Lem. 2.9). However, these bounds become useless, *i.e.* they tend to 0, when we restore the dependence on ε by replacing $A_{||}$ with $\frac{1}{\varepsilon} A_{||}$ in (2.6) and study the limit $\varepsilon \to 0$.

We shall first introduce the concept of weak solution of problem (2.6) and state the existence/uniqueness theorem. Note that unlike the literature cited above, we assume from the beginning that the initial conditions are bounded and strictly positive, and prove the same properties for the weak solutions. Our treatment is thus performed under much less general assumptions than usually required, but this is quite enough for our application.

Definition 2.3 (weak solution)**.** Let $u^0 \in L^\infty(\Omega)$, denote $Q_T = (0,T) \times \Omega$ for any $T > 0$ and introduce the functional space

$$\mathcal{W}_T := \{u \in L^\infty(Q_T) \text{ such that } \nabla u \in L^2(Q_T) \text{ and } \quad \partial_t u \in L^2(0,T;(H^1(\Omega))^*)\}.$$

Then $u \in \mathcal{W}_T$ is called a weak solution to (2.6) on Q_T, if $u(0,\cdot) = u^0$ and

$$\int_0^T \langle \partial_t u(t,\cdot), \phi(t,\cdot)\rangle_{(H^1)^*,H^1}\,\mathrm{d}t + \int_0^T \int_\Omega A_{||}|u|^{m-1}\nabla_{||}u \cdot \nabla_{||}\phi \,\mathrm{d}x\mathrm{d}t \\ + \int_0^T \int_\Omega A_\perp \nabla_\perp u \cdot \nabla_\perp \phi \,\mathrm{d}x\mathrm{d}t + \gamma \int_0^T \int_{\Gamma_\perp} u\phi \,\mathrm{d}\sigma\,\mathrm{d}t = 0, \quad \forall \phi \in \mathcal{D}, \tag{2.7}$$

where $\mathcal{D} = L^2(0,T;H^1(\Omega))$.

Finally, a measurable real-valued function on $Q_\infty := (0,\infty)\times\Omega$ is called a weak solution to (2.6) on Q_∞ (or, simply, a weak solution), if $u|_{Q_T} \in \mathcal{W}_T$ satisfies the relation above for all $T>0$. The space of functions $v : Q_\infty \to \mathbb{R}$ such that $v|_{Q_T} \in \mathcal{W}_T$ for all $T>0$ will be denoted by $\mathcal{W}$.

Remark 2.4. All the terms in this variational formulation are well-defined for any $u \in \mathcal{W}$ and $\phi \in \mathcal{D}$. Note, in particular, that $u \in L^\infty(Q_T)$ and $\nabla u \in L^2(Q_T)$ implies $|u|^{m-1}\nabla_{||}u \in L^2(Q_T)$. Moreover, any $u \in \mathcal{W}_T$ has a trace on the boundary, belonging to $L^2((0,T)\times\partial\Omega)$, which justifies the boundary integral in (2.7). We have also

$$\mathcal{W}_T \subset W_2^1(0,T;H^1(\Omega),L^2(\Omega)) := \{u \in L^2(0,T;H^1(\Omega)), \text{ such that } \partial_t u \in L^2(0,T;(H^1(\Omega))^*)\}.$$

Theorem 25.5 of [30] establishes the continuous imbedding $W_2^1(0,T;H^1(\Omega),L^2(\Omega)) \subset C([0,T];L^2(\Omega))$. This shows that $u(t,\cdot)$ is well defined for all $t\in[0,T]$ as a function of $L^2(\Omega)$ and one can thus impose the initial condition $u(0,\cdot) = u^0$ with $u^0 \in L^\infty(\Omega) \subset L^2(\Omega)$.

Theorem 2.5 (existence/uniqueness/positivity)**.** *Let $m \geq 2$ and $u^0 \in L^\infty(\Omega)$ satisfy $0 < \beta \leq u^0 \leq M < \infty$ on Ω, for some $\beta > 0$. Under Assumption 2.1, there exists a unique weak solution $u \in \mathcal{W}$ of (2.6), which satisfies $c\mathrm{e}^{-Kt} \leq u \leq M$ a.e. on Q_∞, with some sufficiently small $c>0$ and some sufficiently large $K>0$. Moreover, one has also $\frac{\mathrm{d}}{\mathrm{d}t}||u(t,\cdot)||_{L^2(\Omega)} \leq 0$.*

Before proving this theorem, let us introduce the notion of sub- and super-solutions to problem (2.6) and establish a comparison principle for them.

Definition 2.6 (sub/super-solutions)**.** A function $u \in \mathcal{W}_T$ is called a weak sub- (resp. super-) solution to problem (2.6) on Q_T if the variational formulation (2.7) is verified for all $\phi \in \mathcal{D}$ with $\phi \geq 0$ on Q_T, and where the equality sign is replaced by $\leq$ (resp. $\geq$).

Lemma 2.7 (comparison principle)**.** *Let $u_1 \in \mathcal{W}_T$ be a non-negative sub-solution and $u_2 \in \mathcal{W}_T$ be a non-negative super-solution to (2.6) on Q_T with $m \geq 2$. Under Assumption 2.1, if initially $u_1(0,x) \leq u_2(0,x)$ for a.a. $x \in \Omega$, then $u_1 \leq u_2$ on Q_T.*

Proof. For any $k>0$, introduce the function $H_k : \mathbb{R} \to \mathbb{R}$ as

$$H_k(u) = \begin{cases} 0, & \text{if} \quad u \leq 0 \\ ku, & \text{if } 0 < u \leq \quad \frac{1}{k} \\ 1, & \text{if} \quad u > \frac{1}{k} \end{cases}$$

and put $\phi = H_k(u_1 - u_2)$. Note that $\phi \in L^2(0,T;H^1(\Omega))$ for $u_1, u_2 \in \mathcal{W}$ and the gradient of ϕ vanishes almost everywhere outside of the set $\omega_T^k = \{(t,x) \in \bar{Q}_T : 0 < u_1 - u_2 < \frac{1}{k}\}$ (*cf.* [10], Sect. 4.2, Thm. 4(*iii*)). Choosing this ϕ as the test function in the inequalities (2.7) for u_1 and u_2 and subtracting the second one form the first one gives

$$\begin{aligned}\int_0^T \langle \partial_t(u_1-u_2), H_k(u_1-u_2)\rangle_{(H^1)^*,H^1}\,\mathrm{d}t &\le -k \iint\limits_{\omega_T^k} A_\perp \nabla_\perp(u_1-u_2)\cdot\nabla_\perp(u_1-u_2)\mathrm{d}x\mathrm{d}t\\ &- k\iint\limits_{\omega_T^k} A_{||}[u_1^{m-1}\nabla_{||}(u_1-u_2)+(u_1^{m-1}-u_2^{m-1})\nabla_{||}u_2]\cdot\nabla_{||}(u_1-u_2)\mathrm{d}x\mathrm{d}t\\ &-\gamma\int_0^T\int_{\Gamma_\perp}(u_1-u_2)H_k(u_1-u_2)\,\mathrm{d}\sigma\mathrm{d}t\\ &\le k\iint\limits_{\omega_T^k}|A_{||}\nabla_{||}(u_1-u_2)\cdot\nabla_{||}u_2|(u_1^{m-1}-u_2^{m-1})\mathrm{d}x\mathrm{d}t \qquad (2.8)\end{aligned}$$

since (u_1-u_2) and $H_k(u_1-u_2)$ are of the same sign. We have moreover on ω_T^k (recall that $u_1 > u_2$ and $u_1-u_2 < \frac{1}{k}$ there)

$$u_1^{m-1}-u_2^{m-1} \le (m-1)u_1^{m-2}(u_1-u_2) \le \frac{C_T}{k} \qquad (2.9)$$

with $C_T = (m-1)||u_1||^{m-2}_{L^\infty(Q_T)}$ which is finite for $u_1 \in \mathcal{W}$ and obviously independent of k. We see now that (2.8) implies

$$\int_0^T \langle \partial_t(u_1-u_2), H_k(u_1-u_2)\rangle_{(H^1)^*,H^1}\,\mathrm{d}t \le C_T\iint\limits_{\omega_T^k}|A_{||}\nabla_{||}(u_1-u_2)\cdot\nabla_{||}u_2|\mathrm{d}x\mathrm{d}t.$$

Let us now take the limit $k\to\infty$ in this inequality. We have $meas(\omega_T^k)\to 0$ since the sets ω_T^k are nested $(\omega_T^1\supset\omega_T^2\supset\cdots)$ and $\cap_{k=1}^\infty\omega_T^k=\varnothing$. It follows that

$$\limsup_{k\to\infty}\int_0^T \langle \partial_t(u_1-u_2), H_k(u_1-u_2)\rangle_{(H^1)^*,H^1}\,\mathrm{d}t \le 0. \qquad (2.10)$$

On the other hand, for any $k\in\mathbb{N}$

$$\int_0^T \langle \partial_t(u_1-u_2), H_k(u_1-u_2)\rangle_{(H^1)^*,H^1}\,\mathrm{d}t = \int_\Omega \Phi_k(u_1-u_2)\mathrm{d}x\Big|_{t=T} - \int_\Omega \Phi_k(u_1-u_2)\mathrm{d}x\Big|_{t=0} \qquad (2.11)$$

where Φ_k is the primitive of H_k such that $\Phi_k(s)=0$ for $s\le 0$. Indeed, (2.11) trivially holds for smooth functions u_1,u_2. A standard density argument shows then that (2.11) actually holds for any $u_1,u_2\in\mathcal{W}$. We recall to this end that $u_1-u_2\in W_2^1$ and thus it can be approximated by a sequence of smooth functions converging in this space, which is continuously imbedded into $C([0,T],L^2(\Omega))$, *cf.* Remark 2.4. This ensures first of all that the right hand side of (2.11) is well defined and secondly that the equality (2.11) holds. Combining (2.10) with (2.11) and recalling that $\Phi_k(u_1-u_2)=0$ for a.a. $x\in\Omega$ at the initial time $t=0$, yields for any $T>0$

$$\limsup_{k\to\infty}\int_\Omega \Phi_k(u_1-u_2)\mathrm{d}x\Big|_{t=T} = \int_\Omega (u_1-u_2)_+\mathrm{d}x\Big|_{t=T} \le 0, \qquad (2.12)$$

with $(u_1-u_2)_+$ denoting the positive part of (u_1-u_2). This means that $u_1\le u_2$ for a.a. $x\in\Omega$ and $t=T$. Clearly, the same conclusion holds at all time $t\in(0,T)$. □

In the forthcoming proof of our main theorem we shall also need a similar result for a regularized version of problem (2.6).

Lemma 2.8 (comparison principle for a regularized problem)**.** *Let* $a_\alpha(u) := [\max(\min(|u|, M), \alpha)]^{m-1}$ *with some* $M > \alpha > 0$ *and* $m \geq 2$*. Consider the problem obtained from* (2.6) *by replacing* $|u|^{m-1}$ *with* $a_\alpha(u)$ *in the non-linear terms, cf. equation* (2.18) *below. Let moreover* $u_1 \in \mathcal{W}_T$ *be a non-negative sub-solution and* $u_2 \in \mathcal{W}_T$ *be a non-negative super-solution to such a problem on* Q_T *in the sense of Definition* 2.6*. Under Assumption* 2.1*, if* $u_1(0,x) \leq u_2(0,x)$ *for a.a.* $x \in \Omega$*, then* $u_1 \leq u_2$ *on* Q_T*.*

Proof. The proof follows that of Lemma 2.7 word by word and we omit it. It suffices to note that the only thing to be re-verified is the analogue of inequality (2.9) which becomes now

$$a_\alpha(u_1) - a_\alpha(u_2) \leq \frac{C_T}{k} \text{ on } \omega_T^k.$$

This follows easily from the observation that for any $u_1 \geq u_2 \geq 0$ one has $a_\alpha(u_1) - a_\alpha(u_2) \leq (u_1^{m-1} - u_2^{m-1})$ combined with the arguments used already to establish (2.9). □

The construction of the following remarkable sub-solution for (2.6) is essentially due to Pierre [21].

Lemma 2.9 (construction of a sub-solution)**.** *Under Assumption* 2.1*, for any* $\beta > 0$*, there exists a weak sub-solution* $w \in \mathcal{W}$ *to problem* (2.6) *with* $m \geq 1$ *satisfying* $c \leq w(0,x) \leq \beta$ *for* $x \in \Omega$ *and* $w(t,x) \geq c\mathrm{e}^{-Kt}$ *for* $(t,x) \in Q_\infty$*, with some constants* $c, K > 0$ *which depend only on* β*.*

Proof. We will construct a smooth sub-solution w satisfying all the announced properties. We thus rewrite the definition of a sub-solution in the strong form supposing from the beginning that $w \geq 0$:

$$\partial_t w - \frac{1}{m}\nabla_{||} \cdot (A_{||}\nabla_{||} w^m) - \nabla_\perp \cdot (A_\perp \nabla_\perp w) \leq 0, \quad \text{on} \quad (0,\infty) \times \Omega \tag{2.13}$$

$$\frac{1}{m} A_{||} n_{||} \cdot \nabla_{||} w^m + A_\perp n_\perp \cdot \nabla_\perp w + \gamma w \leq 0, \quad \text{on} \quad (0,\infty) \times \Gamma_\perp, \tag{2.14}$$

$$n_\perp \cdot \nabla_\perp w \leq 0, \quad \text{on} \quad (0,\infty) \times \Gamma_{||} \tag{2.15}$$

Thanks to our geometrical hypotheses on the domain Ω, namely that $n||b$ on $\Gamma_{\text{in}} \cup \Gamma_{\text{out}}$, *cf.* Assumption 2.1, one can construct a new coordinate system $\xi_1, \ldots, \xi_d$ on Ω such that the coordinate lines ξ_d coincide with the b-field lines and the surfaces $\xi_d = \text{const.}$ are perpendicular to these lines. Domain Ω is represented in these coordinates by a cylinder $\Omega_\xi = D \times (0,1)$ with $\xi' = (\xi_1, \ldots, \xi_{d-1}) \in D$ and $\xi_d \in (0,1)$. We thus have $\nabla_{||} = b\chi\frac{\partial}{\partial \xi_d}$ with some scalar strictly positive field χ. We assume that the component Γ_{in} of the boundary is represented by $D \times \{\xi_d = 0\}$, Γ_{out} is represented by $D \times \{\xi_d = 1\}$ and $\Gamma_{||}$ is represented by $\partial D \times (0,1)$.

We are searching now for a sub-solution under the form $w(t,x) = \delta(t)(\sin(\pi\xi_d) + \eta(t))^{1/m}$ where $\delta(t)$ and $\eta(t)$ are two positive decreasing functions which are yet to be adjusted. We observe immediately that $\nabla_\perp w = 0$ on Ω for all time so that (2.15) is automatically satisfied. The remaining boundary conditions (2.14) should be checked on Γ_{in} and Γ_{out}. We remind that $n = n_{||} = b$ on Γ_{out} $(\xi_d = 1)$. Similarly, $n = n_{||} = -b$ on Γ_{in} $(\xi_d = 0)$. Substituting the Ansatz for w into (2.14) now gives

$$-\frac{1}{m} A_{||}\chi\delta^m\pi + \gamma\delta\eta^{\frac{1}{m}} \leq 0, \text{ for } \xi_d = 0 \text{ and } \xi_d = 1.$$

This holds if one takes $\eta = K_1\delta^{m(m-1)}$ where $K_1 = \left(\min_{\xi \in \bar{\Omega}_\xi} \frac{\pi}{m\gamma} A_{||}\chi\right)^m > 0$.

It remains to check (2.13). Substituting the Ansatz for w, this inequality is reduced to

$$\dot{\delta} - \frac{\delta^m \pi \cos(\pi\xi_d)\chi}{m(\sin(\pi\xi_d) + \eta)^{\frac{1}{m}}} \frac{\partial(A_{||}\chi)}{\partial \xi_d} + \frac{\delta^m}{m} A_{||}\chi^2\pi^2 \frac{\sin(\pi\xi_d)}{(\sin(\pi\xi_d) + \eta)^{\frac{1}{m}}} \leq 0. \tag{2.16}$$

Note that we have denoted the time derivative here by a dot and we neglected a term with $\dot{\eta}$ since it is negative (the function $\eta(t)$ is decreasing). We bound now the terms on the left-hand side as

$$-\frac{\delta^m \pi \cos(\pi\xi_d)\chi}{m(\sin(\pi\xi_d)+\eta)^{\frac{1}{m}}}\frac{\partial(A_{||}\chi)}{\partial\xi_d} \le \frac{\pi\delta^m\chi}{m\eta^{\frac{1}{m}}}\left|\frac{\partial(A_{||}\chi)}{\partial\xi_d}\right| = \frac{\pi\delta^m\chi}{mK_1^{\frac{1}{m}}\delta^{m-1}}\left|\frac{\partial(A_{||}\chi)}{\partial\xi_d}\right| = \delta\frac{\pi}{mK_1^{\frac{1}{m}}}\chi\left|\frac{\partial(A_{||}\chi)}{\partial\xi_d}\right|$$

and

$$\frac{\delta^m}{m}A_{||}\chi^2\pi^2\frac{\sin(\pi\xi_d)}{(\sin(\pi\xi_d)+\eta)^{\frac{1}{m}}} \le \frac{\delta^m}{m}A_{||}\chi^2\pi^2(\sin(\pi\xi_d))^{1-\frac{1}{m}} \le \frac{\delta^m}{m}A_{||}\chi^2\pi^2.$$

We see thus that inequality (2.16) will be satisfied if we require

$$\dot{\delta} + K_2\delta + K_3\delta^m \le 0 \tag{2.17}$$

with

$$K_2 = \frac{\pi}{mK_1^{\frac{1}{m}}}\max_{\xi\in\Omega_\xi}\left|\chi\frac{\partial}{\partial\xi_d}\left(A_{||}\chi\right)\right| \text{ and } K_3 = \frac{\pi^2}{m}\max_{\xi\in\Omega_\xi}\left|A_{||}\chi^2\right|.$$

One can thus take $\delta(t) = \delta_0 \mathrm{e}^{-(K_2+K_3)t}$ with any $\delta_0 \in (0,1]$.

In summary, $w(t,x) = \delta_0\mathrm{e}^{-(K_2+K_3)t}(\sin(\pi\xi_d)+K_1\delta_0^{m(m-1)}\mathrm{e}^{-m(m-1)(K_2+K_3)t})^{\frac{1}{m}}$ is a sub-solution. Clearly, for any $\beta > 0$ one can take δ_0 small enough so that $w(0,x) \le \beta$. Moreover, for any t, $w(t,x) \ge \delta_0^m K_1^{1/m}\mathrm{e}^{-m(K_2+K_3)t}$ so that we have proved the statement of the Lemma putting $c = \delta_0^m K_1^{1/m}$, $K = m(K_2+K_3)$. □

Remark 2.10. In the case of a simple "aligned" geometry, *i.e.* $b = e_d$ and $\Omega = D\times]0,L[$ with D a domain in $\mathbb{R}^{d-1}$, and supposing $A_{||} = \text{const.}$, one can easily construct a sub-solution satisfying a sharper estimate: under the assumptions of the preceding Lemma, there is a sub-solution such that

$$w(t,x) \ge \frac{C}{(1+Kt)^{\frac{m}{m-1}}}\cdot$$

Indeed, one can repeat the proof as in case A of the preceding Lemma, taking $\xi' = (x_1,\ldots,x_{d-1})$, $\xi_d = x_d/L$, up to the differential inequality (2.17). One observes now that $K_2 = 0$ so that one can take $\delta(t) = \frac{\delta_0}{(1+Kt)^{\frac{1}{m-1}}}$ with any $\delta_0 > 0$ and $K = (m-1)K_3\delta_0^{m-1}$. Our sub-solution is thus $w = \frac{\delta_0}{(1+Kt)^{\frac{1}{m-1}}}\left(\sin(\pi\xi_d) + \frac{K_1\delta_0^{m(m-1)}}{(1+Kt)^m}\right)^{\frac{1}{m}}$ and $w \ge \frac{K_1^{\frac{1}{m}}\delta_0^m}{(1+Kt)^{\frac{m}{m-1}}}$ as stated.

Remark 2.11. The hypothesis that $n||b$ on $\Gamma_{\text{in}} \cup \Gamma_{\text{out}}$ is essential for the existence of a positive sub-solution. Indeed, the numerical experiments in Section 4.2.3 suggest that if this hypothesis is violated then the solution to (2.2) starting from a positive initial condition can become zero at a finite time. There is thus no hope to prove an analogue of Lemma 2.9 under more general geometrical assumptions.

Let us now turn to the proof of our main result.

Proof of Theorem 2.5. 1st step. Regularization.
Assume that $M > 0$ is an upper bound for u^0 and pick any $\alpha \in (0,M)$. Introduce the function $a_\alpha : \mathbb{R} \to \mathbb{R}^+$ by $a_\alpha(u) := [\max(\min(|u|,M),\alpha)]^{m-1}$ and consider the regularized version of (2.7): find $u_\alpha \in W_2^1(0,T;H^1(\Omega),L^2(\Omega))$ (*i.e.* $u_\alpha \in L^2(0,T;H^1(\Omega))$ and $\partial_t u_\alpha \in L^2(0,T;(H^1(\Omega))^*)$) such that $u_\alpha(0,\cdot) = u^0$ and

$$\int_0^T \langle\partial_t u_\alpha(t,\cdot),\phi(t,\cdot)\rangle_{(H^1)^*,H^1}\,\mathrm{d}t + \int_0^T\int_\Omega A_{||}a_\alpha(u_\alpha)\nabla_{||}u_\alpha\cdot\nabla_{||}\phi\,\mathrm{d}x\mathrm{d}t$$

$$+\int_0^T\int_\Omega A_\perp\nabla_\perp u_\alpha\cdot\nabla_\perp\phi\,\mathrm{d}x\mathrm{d}t + \gamma\int_0^T\int_{\Gamma_\perp}u_\alpha\phi\,\mathrm{d}\sigma\,\mathrm{d}t = 0, \quad \forall\phi\in\mathcal{D}. \tag{2.18}$$

By standard arguments, this problem has a solution. Indeed, consider the mapping

$$\mathcal{T} : B_R(0) \to L^2(Q_T), \quad B_R(0) := \{v \in L^2(Q_T) \ / \ ||v||_{L^2(Q_T)} \leq R\},$$

where we associate to any $v \in B_R(0)$ the unique solution $u \in W_2^1(0,T;H^1(\Omega),L^2(\Omega))$ of the linearized, regular parabolic problem

$$\begin{aligned}
&\int_0^T \langle \partial_t u(t,\cdot), \phi(t,\cdot)\rangle_{(H^1)^*,H^1}\,\mathrm{d}t + \int_0^T \int_\Omega A_{||} a_\alpha(v) \nabla_{||} u \cdot \nabla_{||}\phi \,\mathrm{d}x\mathrm{d}t \\
&\qquad + \int_0^T \int_\Omega A_\perp \nabla_\perp u \cdot \nabla_\perp \phi \,\mathrm{d}x\mathrm{d}t + \gamma \int_0^T \int_{\Gamma_\perp} u\phi \,\mathrm{d}\sigma\,\mathrm{d}t = 0, \quad \forall \phi \in \mathcal{D}, \\
&u(0,\cdot) = u^0 \text{ on } \Omega.
\end{aligned}$$

Taking the test function $\phi = u$ we see readily that $||u(t,\cdot)||_{L^2(\Omega)} \leq ||u^0||_{L^2(\Omega)}$ for any $t \in (0,T)$ so that $\mathcal{T}(B_R(0)) \subset B_R$ provided $R \geq \sqrt{T}||u^0||_{L^2(\Omega)}$. Moreover, $\mathcal{T}(B_R(0))$ is bounded in $W_2^1(0,T;H^1(\Omega),L^2(\Omega))$ by the usual estimates for the parabolic problems, hence it is relatively compact in $L^2(Q_T)$. Indeed, $W_2^1(0,T;H^1(\Omega),L^2(\Omega))$ is compactly embedded into $L^2(Q_T)$ [24]. It is also easy to see that the mapping $\mathcal{T}$ is continuous. Indeed, for any sequence $v_n \to v$ in $L^2(Q_T)$ denote $u_n = \mathcal{T}(v_n)$. The sequence $\{u_n\}$ is bounded in $W_2^1(0,T;H^1(\Omega),L^2(\Omega))$ so that it contains a weakly convergent sub-sequence $u_{n_k} \rightharpoonup w$ with some w in the same space. Extracting a sub-sequence again, we can also assume that $v_{n_k} \to v$ a.e. in Q_T so that the Lebesgue dominated convergence theorem permits us to pass to the limit in the term of the variational formulation containing $a_\alpha(v)$ and to identify w with $\mathcal{T}(v)$. Since this argument holds for any weakly convergent sub-sequence of $\{u_n\}$, we see that the whole sequence converges weakly, *i.e.* $u_n = \mathcal{T}(v_n) \rightharpoonup \mathcal{T}(v)$ in $W_2^1(0,T;H^1(\Omega),L^2(\Omega))$ so that $\mathcal{T}(v_n) \to \mathcal{T}(v)$ in $L^2(Q_T)$. Thus, $\mathcal{T}$ satisfies all the hypotheses of Schauder fixed point theorem, hence it has a fixed point $\mathcal{T}(u) = u$, which provides a solution to (2.18).

The solution u_α of problem (2.18) satisfies $0 \leq u_\alpha \leq M$, provided we have $0 \leq u^0 \leq M$. Indeed, define $u_\alpha^- := \min(0,u_\alpha) \leq 0$. Then one gets for the initial condition $u_\alpha^-(0,\cdot) \equiv 0$. Taking u_α^- as the test function in the variational formulation (2.18), yields immediately

$$\begin{aligned}
&\frac{1}{2}\int_\Omega |u_\alpha^-(T,x)|^2\mathrm{d}x + \int_0^T \int_\Omega A_{||} a_\alpha(u_\alpha^-) |\nabla_{||} u_\alpha^-|^2 \mathrm{d}x\,\mathrm{d}t \\
&\qquad + \int_0^T \int_\Omega A_\perp \nabla_\perp u_\alpha^- \cdot \nabla_\perp u_\alpha^- \mathrm{d}x\mathrm{d}t + \gamma \int_0^T \int_{\Gamma_\perp} |u_\alpha^-|^2 \,\mathrm{d}\sigma\,\mathrm{d}\tau = 0,
\end{aligned}$$

which shows that $u_\alpha^-(T,\cdot) \equiv 0$.

To prove the estimate from above, define $u_\alpha^+ := \max(0, u_\alpha - M)$. Observe that $u_\alpha^+(0,\cdot) \equiv 0$ and take u_α^+ as the test function in the variational formulation (2.18):

$$\begin{aligned}
&\frac{1}{2}\int_\Omega |u_\alpha^+(T,x)|^2\mathrm{d}x + \int_0^T \int_\Omega A_{||} a_\alpha(u_\alpha) |\nabla_{||} u_\alpha^+|^2 \,\mathrm{d}x\,\mathrm{d}t \\
&\qquad + \int_0^T \int_\Omega A_\perp \nabla_\perp u_\alpha^+ \cdot \nabla_\perp u_\alpha^+ \,\mathrm{d}x\mathrm{d}t + \gamma \int_0^T \int_{\Gamma_\perp} u_\alpha u_\alpha^+ \,\mathrm{d}\sigma\,\mathrm{d}\tau = 0,
\end{aligned}$$

which shows that $u_\alpha^+(T,\cdot) \equiv 0$. Since the same arguments can be applied to any final time, we have $0 \leq u_\alpha \leq M$ at all time $t \in [0,T]$.

2nd step. Bounds on u_α and its identification with u.
Recall the sub-solution to problem (2.6) constructed in Lemma 2.9: $w(t,x)$ such that $w(0,\cdot) \leq \beta \leq u^0(\cdot)$ and $w(t,x) \geq c\mathrm{e}^{-Kt}$ with some constants $c, K > 0$ dependent on β. Fix any final time $T > 0$, take $\alpha := c\mathrm{e}^{-KT}$ and consider the solution u_α of (2.18) constructed above.

On the one hand, one sees immediately that w satisfies $\alpha \le w \le M$ for all $(t,x) \in Q_T$, hence $a_\alpha(w(t,x)) = w(t,x)^{m-1}$ in Q_T which means that w is a non-negative sub-solution to (2.18). On the other hand, u_α is a non-negative super-solution to (2.18) in the sense of Definition 2.6, since we have already proved that $u_\alpha \in L^\infty(Q_T)$ (and thus $u_\alpha \in \mathcal{W}_T$) as well as $u_\alpha \ge 0$. The comparison principle 2.8 entails therefore $u_\alpha(t,x) \ge w(t,x) \ge \alpha$ for a.a. $(t,x) \in Q_T$. It means in turn that $a_\alpha(u_\alpha) = u_\alpha^{m-1}$ on Q_T so that u_α is also a weak solution to (2.6) on Q_T.

This establishes the existence of a solution to (2.6) on Q_T for any final time $T > 0$. In order to prove that the solution actually exists for all time $t > 0$, it remains to observe that if we have a weak solution u to (2.6) on Q_T with some $T > 0$ and another solution u' on $Q_{T'}$ with some $T' > T$ then $u = u'$ on Q_T. This follows from the uniqueness of the weak solution established in the next step.

3rd step. Uniqueness and positivity.

Suppose now that (2.7) admits two weak solutions u_1 and u_2 on Q_T with the same initial condition $u_1 = u_2 = u^0 \ge 0$ at $t = 0$. One can easily prove that u_1 and u_2 are non-negative by the same arguments as those used for non-negativity of u_α in the 1st step. Lemma 2.7 implies now $u_1 \le u_2$ and $u_2 \le u_1$ on Q_T. Thus $u_1 = u_2$ on Q_T.

We conclude that the unique weak solution to (2.6) exists for all time and thus it coincides on any time interval $[0,T]$ with u_α for α small enough. This implies that u is strictly positive, *i.e.* $u \ge c\mathrm{e}^{-Kt}$ with some positive constants c small enough and K large enough, and bounded from above by M. Indeed, these properties are already proven for u_α.

Finally, the inequality $\frac{\mathrm{d}}{\mathrm{d}t}||u(t,\cdot)||_{L^2(\Omega)} \le 0$ can be easily proven by taking the test function $\phi = u$ in (2.7). □

3. Numerical method

3.1. Semi-discretization in space

The variational formulation of the singular perturbation problem (2.2) can be rewritten as follows: find $u \in \mathcal{W}$, *i.e.* $u(t,\cdot) \in \mathcal{V} := H^1(\Omega)$ such that $\partial_t u(t,\cdot) \in \mathcal{V}^*$ and satisfying

$$(P) \quad \langle \partial_t u(t,\cdot),\, v\rangle_{\mathcal{V}^*,\mathcal{V}} + \frac{1}{\varepsilon}\int_\Omega A_{||}|u|^{5/2}\nabla_{||}u(t,\cdot)\cdot\nabla_{||}v\,\mathrm{d}x + \int_\Omega A_\perp\nabla_\perp u(t,\cdot)\cdot\nabla_\perp v\,\mathrm{d}x + \gamma\int_{\Gamma_\perp} u(t,\cdot)v\,\mathrm{d}\sigma = 0, \quad \forall v \in \mathcal{V} \tag{3.1}$$

for almost every $t \in (0,T)$. As mentioned already in Section 2, this problem becomes ill-posed if we take formally the limit $\varepsilon \to 0$. Indeed, only the leading term survives in this limit, so that any function from the space

$$\mathcal{G} := \{p \in \mathcal{V} \;/\; \nabla_{||}p = 0 \text{ in } \Omega\}$$

would be a solution. The well-posed limit problem for the solutions to (P) as $\varepsilon \to 0$ can be however easily established. Indeed, one can restrain the test functions in (P) to be in the space $\mathcal{G}$ so that the ε-dependent term disappears and the correct problem in the limit $\varepsilon \to 0$ reads: find $u(t,\cdot) \in \mathcal{G}$ such that $\partial_t u(t,\cdot) \in \mathcal{G}^*$ and

$$(L) \;\; \langle \partial_t u(t,\cdot),\, v\rangle_{\mathcal{V}^*,\mathcal{V}} + \int_\Omega A_\perp\nabla_\perp u(t,\cdot)\cdot\nabla_\perp v\,\mathrm{d}x + \gamma\int_{\Gamma_\perp} u(t,\cdot)v\,\mathrm{d}\sigma = 0, \quad \forall v \in \mathcal{G}$$

for almost every $t \in (0,T)$. This problem should be accompanied with an initial condition from the space $\mathcal{G}$. The weak formulation suggests (at least formally) that the proper initial condition here is $u(t,\cdot) = P_\mathcal{G}u^0$ where $P_\mathcal{G}$ is the L^2-orthogonal projector on $\mathcal{G}$. In the case of mismatch between the initial conditions for problems (P) and (L) we should thus expect that a sharp boundary layer appears near the initial time $t = 0$ at small values of ε. The derivative in time of the solution to problem (P) will be of order $1/\varepsilon$ there and its width (the duration in time) will be of order ε.

The discussion above shows that a straight-forward discretization of problem (P) may lead to very inaccurate results when $\varepsilon \ll 1$. Indeed, setting $\varepsilon = 0$ would result in a singular problem, so that the problem with

$\varepsilon \ll 1$ would be very ill-conditioned. To cope with this difficulty and to obtain a numerical scheme which is uniformly accurate with respect to ε, we introduce an Asymptotic-Preserving reformulation, very similar to the one introduced in [7]. The idea is to rewrite the singularly perturbed problem (3.1) in an equivalent form, which is however well-posed when one sets there formally $\varepsilon = 0$ and gives moreover the correct limit problem (L). In order to do this, we introduce an auxiliary unknown q by the relation $\varepsilon \nabla_{||} q = u^{5/2} \nabla_{||} u$ in Ω and $q = 0$ on Γ_{in}, which rescales the nasty part of the equation permitting to get rid of the terms of order $O(1/\varepsilon)$. The reformulated problem, called in the sequel the Asymptotic-Preserving reformulation (AP-problem) reads: find $(u(t,\cdot), q(t,\cdot)) \in \mathcal{V} \times \mathcal{L}$, solution of

$$(AP) \quad \begin{cases} \left\langle \dfrac{\partial u}{\partial t}, v \right\rangle_{\mathcal{V}^*,\mathcal{V}} + \displaystyle\int_\Omega (A_\perp \nabla_\perp u) \cdot \nabla_\perp v \,\mathrm{d}x + \int_\Omega A_\| \nabla_\| q \cdot \nabla_\| v \,\mathrm{d}x + \gamma \int_{\Gamma_\perp} uv \,\mathrm{d}s = 0, \\ \hfill \forall v \in \mathcal{V} \\ \displaystyle\int_\Omega A_\| u^{5/2} \nabla_\| u \cdot \nabla_\| w \,\mathrm{d}x - \varepsilon \int_\Omega A_\| \nabla_\| q \cdot \nabla_\| w \,\mathrm{d}x = 0, \quad \forall w \in \mathcal{L}, \end{cases} \tag{3.2}$$

where

$$\mathcal{L} := \{ q \in L^2(\Omega) \ / \ \nabla_\| q \in L^2(\Omega) \text{ and } q|_{\Gamma_{\text{in}}} = 0 \}. \tag{3.3}$$

System (3.2) is an equivalent reformulation (for fixed $\varepsilon > 0$) of the original P-problem (3.1). Putting now formally $\varepsilon = 0$ in (AP) leads to the well-posed limit problem

$$(L') \quad \begin{cases} \left\langle \dfrac{\partial u}{\partial t}, v \right\rangle_{\mathcal{V}^*,\mathcal{V}} + \displaystyle\int_\Omega (A_\perp \nabla_\perp u) \cdot \nabla_\perp v \,\mathrm{d}x + \int_\Omega A_\| \nabla_\| q \cdot \nabla_\| v \,\mathrm{d}x + \gamma \int_{\Gamma_\perp} uv \,\mathrm{d}s = 0, \\ \hfill \forall v \in \mathcal{V} \\ \displaystyle\int_\Omega A_\| u^{5/2} \nabla_\| u \cdot \nabla_\| w \,\mathrm{d}x = 0, \quad \forall w \in \mathcal{L}, \end{cases} \tag{3.4}$$

which is equivalent to problem (L). Note that q acts here as a Lagrange multiplier for the constraint $u \in \mathcal{G}$, which provides the uniqueness of the solution. Hence the AP-reformulation permits a continuous transition from the P-model to the L-model, which enables the uniform accuracy of the scheme with respect to ε.

Imposing q to be zero on Γ_{in} provides the uniqueness of this auxiliary variable in the case of simply connected anisotropy field geometries, *i.e.* every field line entering the computational domain on Γ_{in} must leave on Γ_{out}. Remark that no closed field lines are allowed since their presence would destroy the uniqueness of q. More complicated geometries including magnetic islands in the context of both elliptic and parabolic problems are the subject of the ongoing work [22,23].

Let us now choose a triangulation of the domain Ω with triangles or quadrangles of order h and introduce the finite element spaces $\mathcal{V}_h \subset \mathcal{V}$ and $\mathcal{L}_h \subset \mathcal{L}$ of type $\mathbb{P}_k$ or $\mathbb{Q}_k$ on this mesh. The finite element discretization of (3.2) writes then: find $(u_h, q_h) \in \mathcal{V}_h \times \mathcal{L}_h$ such that

$$(AP)_h \quad \begin{cases} \displaystyle\int_\Omega \frac{\partial u_h}{\partial t} v_h \,\mathrm{d}x + \int_\Omega (A_\perp \nabla_\perp u_h) \cdot \nabla_\perp v_h \,\mathrm{d}x + \int_\Omega A_\| \nabla_\| q_h \cdot \nabla_\| v_h \,\mathrm{d}x + \gamma \int_{\Gamma_\perp} u_h v_h \,\mathrm{d}s = 0, \\ \hfill \forall v_h \in \mathcal{V}_h \\ \displaystyle\int_\Omega A_\| u_h^{5/2} \nabla_\| u_h \cdot \nabla_\| w_h \,\mathrm{d}x - \varepsilon \int_\Omega A_\| \nabla_\| q_h \cdot \nabla_\| w_h \,\mathrm{d}x = 0, \quad \forall w \in \mathcal{L}_h. \end{cases} \tag{3.5}$$

3.2. Semi-discretization in time

Our goal now is to devise time-marching schemes for (3.5) that preserve the AP property announced above on the continuous level in time. This entails, loosely speaking, that the discretization error will depend only on the time step τ, but not on ε, apart from the sharp boundary layer near the initial time, outlined above.

We want to be able to take large time-steps $\tau \gg \varepsilon$ and to capture the slowly varying component of the solution at least for time $t \geq \tau$. If, on the contrary, we were interested in the accurate description of the boundary layer, we would have to choose $\tau < \varepsilon$. We conjecture that our AP-scheme could handle such situations as well, but we do not investigate this in the present paper.

In order to approximate numerically the time derivative in (3.5), we introduce three different schemes: a standard first order, implicit Euler scheme, the Crank−Nicolson scheme and a second order, L-stable Runge−Kutta method. We show in the following that the first order Euler-scheme is stable and asymptotic-preserving. The Crank−Nicolson scheme gives reliable results and second order convergence under certain assumptions, but is not asymptotic-preserving. Thus, if second order accuracy in time is desired, the L-stable Runge−Kutta method has to be applied. All three methods are exposed to numerical tests and compared in Section 4.

3.2.1. Implicit euler scheme

Introducing the forms

$$(\Theta, \chi) := \int_\Omega \Theta \chi \, \mathrm{d}x, \tag{3.6}$$

$$a_{\parallel nl}(\Psi, \Theta, \chi) := \int_\Omega A_\parallel \Psi^{5/2} \nabla_\parallel \Theta \cdot \nabla_\parallel \chi \, \mathrm{d}x, \tag{3.7}$$

$$a_\parallel(\Theta, \chi) := \int_\Omega A_\parallel \nabla_\parallel \Theta \cdot \nabla_\parallel \chi \, \mathrm{d}x, \qquad a_\perp(\Theta, \chi) := \int_\Omega A_\perp \nabla_\perp \Theta \cdot \nabla_\perp \chi \, \mathrm{d}x, \tag{3.8}$$

allows us to write the first order, implicit Euler method in the compact notation: find $(u_h^{n+1}, q_h^{n+1}) \in \mathcal{V}_h \times \mathcal{L}_h$, solution of

$$(E_{AP}) \quad \begin{cases} \left(u_h^{n+1}, v_h\right) + \tau \left(a_\perp \left(u_h^{n+1}, v_h\right) + a_\parallel \left(q_h^{n+1}, v_h\right) + \gamma \int_{\Gamma_\perp} u_h^{n+1} v_h \, \mathrm{d}s \right) = \left(u_h^n, v_h\right) \\ a_{\parallel nl}\left(u_h^n, u_h^{n+1}, w_h\right) - \varepsilon a_\parallel \left(q_h^{n+1}, w_h\right) = 0. \end{cases}, \tag{3.9}$$

Note that the non-linearity is evaluated here explicitly (we take u_h^n rather than (u_h^{n+1}) inside the $|u|^{5/2}$ term).

A slightly different first order AP-scheme was introduced in [18] for the resolution of the same temperature balance problem. There, the (P)-problem was firstly discretized in time (implicit Euler), then linearized by a fixed point mapping, and finally the AP reformulation applied. In contrast to the present scheme, the non-linearity was evaluated there completely implicitly. The numerical results obtained in [18] are similar to the present ones.

3.2.2. Linearized Crank−Nicolson scheme

To construct a scheme, which is second order in time, one can come to the idea to employ the Crank−Nicolson scheme: find $(u_h^{n+1}, q_h^{n+1}) \in \mathcal{V}_h \times \mathcal{L}_h$, solution of

$$\begin{cases} \left(u_h^{n+1}, v_h\right) + \tau \left(a_\perp \left(u_h^{n+1/2}, v_h\right) + a_\parallel \left(q_h^{n+1}, v_h\right) + \gamma \int_{\Gamma_\perp} u_h^{n+1/2} v_h \, \mathrm{d}s \right) = \left(u_h^n, v_h\right) \\ a_{\parallel nl}\left(u_h^{n+1/2}, u_h^{n+1/2}, w_h\right) - \varepsilon a_\parallel \left(q_h^{n+1}, w_h\right) = 0. \end{cases} \tag{3.10}$$

As one can observe, we have to deal for each fixed n, with a nonlinear equation. In the linear terms, one can set $u_h^{n+1/2} = \frac{1}{2}\left(u_h^{n+1} + u_h^n\right)$. To linearize the term $a_{\parallel nl}(u_h^{n+1/2}, u_h^{n+1/2}, w_h)$ however, we shall use the standard linear extrapolation method. In other words, the non-linearity in this last formula, $(u_h^{n+1/2})^{5/2}$, will be replaced by a linearized second order approximation in τ:

$$\left(u_h^{n+1/2}\right)^{5/2} = \left(u_h^n + \frac{1}{2}\left(u_h^n - u_h^{n-1}\right) + O\left(\tau^2\right)\right)^{5/2} = \left(u_h^n + \frac{1}{2}\left(u_h^n - u_h^{n-1}\right)\right)^{5/2} + O\left(\tau^2\right). \tag{3.11}$$

The resulting linear system reads finally: find $(u_h^{n+1}, q_h^{n+1}) \in \mathcal{V}_h \times \mathcal{L}_h$, solution of

$$(CN_{AP}) \quad \begin{cases} \left(u_h^{n+1}, v_h\right) + \frac{\tau}{2}\left(a_\perp\left(u_h^{n+1}, v_h\right) + \gamma \int_{\Gamma_\perp} u_h^{n+1} v_h \, \mathrm{d}s\right) + \tau a_\parallel (q_h^{n+1}, v_h) \\ \qquad\qquad = \left(u_h^n, v_h\right) - \frac{\tau}{2}\left(a_\perp\left(u_h^n, v_h\right) + \gamma \int_{\Gamma_\perp} u_h^n v_h \, \mathrm{d}s\right), \\ \frac{1}{2} a_{\parallel nl}\left(\frac{1}{2}\left(3u_h^n - u_h^{n-1}\right), u_h^{n+1}, w_h\right) - \varepsilon a_\parallel (q_h^{n+1}, w_h) \\ \qquad\qquad = -\frac{1}{2} a_{\parallel nl}\left(\frac{1}{2}\left(3u_h^n - u_h^{n-1}\right), u_h^n, w_h\right). \end{cases} \tag{3.12}$$

Unfortunately, this method is not Asymptotic-Preserving. Indeed, putting $\varepsilon = 0$ does not force u_h^n to lie in a finite element discretization of the space $\mathcal{G}$ of functions constant along the field lines because of the non-zero right hand side in the second equation in (3.12). This should be contrasted with the Euler scheme presented above or the forthcoming DIRK scheme (3.16). The corresponding equations have in both cases zero right hand sides. In the case of the Crank−Nicolson scheme, one may expect therefore some severe numerical artifacts (positivity loss, non physical oscillations) in the regime $\varepsilon \ll 1$ unless one takes exceedingly small time step $\tau \ll \varepsilon$ to resolve the rapid dynamics at time $t < \varepsilon$.

This is closely related to the notion of A-stability and L-stability of numerical schemes for ordinary differential equations (see for example [12]). Let us consider a test problem $y' = ky$ with $k \in \mathbb{C}$. A Runge−Kutta method discretizing this problem on a uniform mesh of step τ can be expressed as $y^{n+1} = \phi(k\tau)y^n$ with ϕ known as the stability function. A method is called A-stable when $|\phi(z)| < 1$ for all z such that $\operatorname{Re} z < 0$, which implies that the numerical solution to the test problem approaches 0 as $t \to \infty$ provided $\operatorname{Re} k < 0$. A method is L-stable if it is A-stable and $|\phi(z)| \to 0$ as $z \to \infty$. Loosely speaking, it means that if the equation is infinitely stiff ($\operatorname{Re} k \to -\infty$) then the L-stable scheme produces the zero numerical solution already at the first time step. This is exactly the behaviour that one would expect from an AP scheme. Among the two schemes introduced above, the implicit Euler is L-stable. On the contrary, the Crank−Nicolson scheme is only A-stable but not L-stable and thus not suitable for small values of ε and large values of τ.

As an example of the non-convergence of the (CN_{AP}) scheme in a general case, we show some numerical results corresponding to a test case defined in the Section 4.2.2. The initial condition is a Gaussian peak located in the center of the computational domain with a maximum of 10^5 K. If the time step τ is too large, u_h^n will immediately reach negative values and thus the numerical algorithm will fail in the next iteration. However, if τ is sufficiently small then the (CN_{AP}) scheme is of second order in time. Unfortunately, the biggest time step that does not provoke oscillations in the numerical solution, is of the order of 10^{-16} s, for an initial Gaussian peak of 10^5 K. This makes the (CN_{AP}) scheme of no practical use in real simulations. These results are plotted on Figure 2.

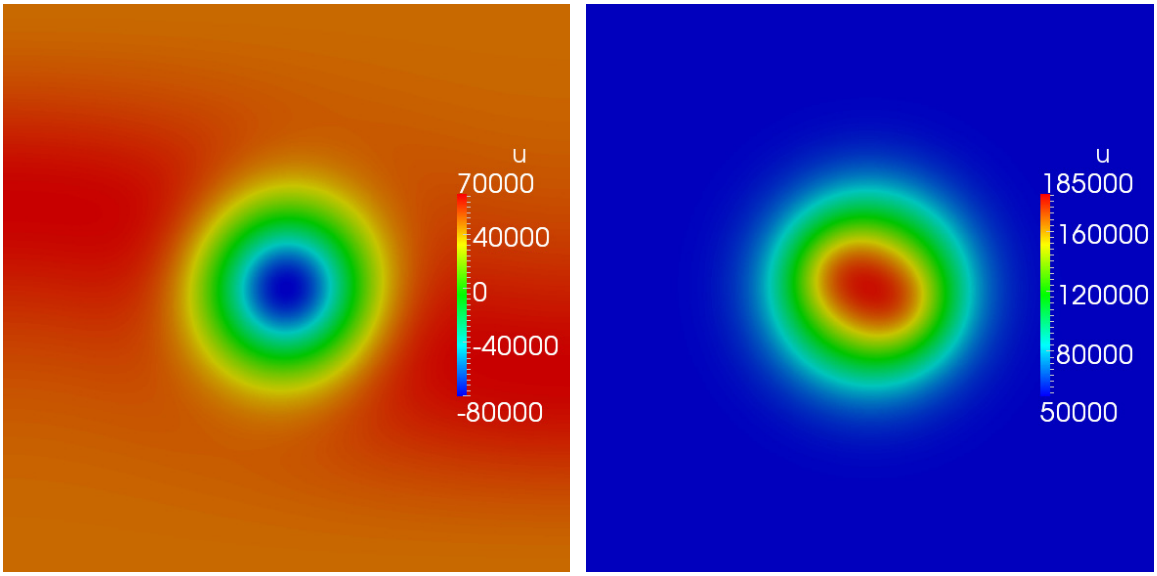
3.2.3. L-stable Runge−Kutta method

As we are interested in an AP-scheme, which is second order accurate in time, we propose to use a two stage Diagonally Implicit Runge−Kutta (DIRK) second order scheme [12], which does not suffer from the limitations of the Crank−Nicolson discretization. In order to explain the construction of this scheme, let us first remind some notations related to the Runge−Kutta method on the example of the ODE of the type

$$\frac{\mathrm{d}u}{\mathrm{d}t} = L(t)u + f(t). \tag{3.13}$$

The coefficients of an s-stage Runge−Kutta method are usually displayed in a Butcher's diagram:

$$\begin{array}{c|ccc} c_1 & a_{11} & \cdots & a_{1s} \\ \vdots & \vdots & & \vdots \\ c_s & a_{s1} & \cdots & a_{ss} \\ \hline & b_1 & \cdots & b_s \end{array}. \tag{3.14}$$



(a) $\tau = 0.1$

(b) $\tau = 10^{-16}$

FIGURE 2. Non convergence of the (CN_{AP}) scheme. Negative values of u_h^n are obtained after one iteration of the method, for big time steps. If the time step is sufficiently small, the method converges.

Applying this method to approximate to the problem above proceeds as follows: for given u^n, being an approximation of $u(t_n)$, u^{n+1} is determined by

$$k_i = L(t_n + c_i\tau)\left(u^n + \tau\sum_{j=1}^{s} a_{ij}k_j\right) + f(t_n + c_i\tau),$$
$$u^{n+1} = u^n + \tau\sum_{j=1}^{s} b_j k_j.$$

The two stage DIRK scheme is developed according to the following Butcher's diagram:

$$\begin{array}{c|cc} \lambda & \lambda & 0 \\ 1 & 1-\lambda & \lambda \\ \hline & 1-\lambda & \lambda \end{array}$$

with $\lambda = 1 - \frac{1}{\sqrt{2}}$ (this together with $\lambda = 1 + \frac{1}{\sqrt{2}}$ are the only values of λ that guarantee the second order accuracy, *cf.* Table 6.4 in [12]). After simple algebraic modifications (introducing $\hat{u}^{n+1} = u^n + \lambda\tau k_1$ and eliminating k_1 and k_2), we can thus write this method, as applied to (3.13), in the following form

$$\begin{aligned} \hat{u}^{n+1} &= u^n + \lambda\tau L(t_n + \lambda\tau)\hat{u}^{n+1} + \lambda\tau f(t_n + \lambda\tau) \\ u^{n+1} &= u^n + \frac{1-\lambda}{\lambda}(\hat{u}^{n+1} - u^n) + \lambda\tau L(t_n + \tau)u^{n+1} + \lambda\tau f(t_n + \tau). \end{aligned} \tag{3.15}$$

To transpose this scheme to the context of the AP-reformulation of the non-linear heat equation (3.5), we should linearize the non-linear coefficient $u_h^{5/2}$. We do it by replacing $u_h(t)$ inside this coefficient for time

$t \in [t_n, t_{n+1})$ by $u_h^n + \frac{t-t_n}{\tau}(u_h^n - u_h^{n-1})$ thus introducing an extra error of order $O(\tau^2)$ which is consistent with the overall order of the scheme. The scheme now reads: find $(\hat{u}_h^{n+1}, \hat{q}_h^{n+1})$, $(u_h^{n+1}, q_h^{n+1}) \in \mathcal{V}_h \times \mathcal{L}_h$, solutions of

$$(RK_{AP}) \quad \begin{cases} \left(\hat{u}_h^{n+1}, v_h\right) + \tau\lambda \left(a_\perp(\hat{u}_h^{n+1}, v_h) + \gamma \int_{\Gamma_\perp} \hat{u}_h^{n+1} v_h \,\mathrm{d}s + a_\parallel(\hat{q}_h^{n+1}, v_h)\right) \\ \qquad\qquad = (u_h^n, v_h) \\ a_{\parallel nl}\left(u_h^n + \lambda(u_h^n - u_h^{n-1}), \hat{u}_h^{n+1}, w_h\right) - \varepsilon a_\parallel\left(\hat{q}_h^{n+1}, w_h\right) = 0 \end{cases} \\ \begin{cases} \left(u_h^{n+1}, v_h\right) + \tau\lambda \left(a_\perp\left(u_h^{n+1}, v_h\right) + \gamma \int_{\Gamma_\perp} u_h^{n+1} v_h \,\mathrm{d}s + a_\parallel(q_h^{n+1}, v_h)\right) \\ \qquad\qquad = (u_h^n, v_h) + \frac{1-\lambda}{\lambda}\left(\hat{u}_h^{n+1} - u_h^n, v_h\right) \\ a_{\parallel nl}\left(u_h^n + (u_h^n - u_h^{n-1}), u_h^{n+1}, w_h\right) - \varepsilon a_\parallel\left(q_h^{n+1}, w_h\right) = 0. \end{cases} \tag{3.16}$$

For each time step we have therefore to assemble and solve two linearized problems. This method is two times slower than the Crank−Nicolson scheme, with the advantage however of maintaining the AP-property of the scheme, advantage which is crucial for $0 < \varepsilon \ll 1$.

Remark 3.1. In some of our test cases below we shall slightly modify the governing equation (2.2) by adding a non-zero source term $f(t, x)$ to the right-hand side. While it is straightforward to adapt the Euler scheme (3.9) to this case, the modifications applied to the DIRK scheme (3.16) deserve a more detailed explanation. In accordance with (3.15) we modify (3.16) in the presence of a source term f as follows: we add

$$\lambda\tau(f(t_n + \lambda\tau), v_h) \qquad \text{resp.} \qquad \lambda\tau(f(t_n + \tau), v_h)$$

to the right-hand side of the first, resp. third, equation in (3.16). The parentheses $(\cdot, \cdot)$ stand here for L^2 scalar product.

4. Numerical results

In this section we compare the proposed implicit Euler-AP and DIRK-AP schemes with a straightforward linearized implicit Euler discretization of the initial singular perturbation problem (2.2), given by

$$(P) \quad \left(u_h^{n+1}, v_h\right) + \tau \left(a_\perp\left(u_h^{n+1}, v_h\right) + \frac{1}{\varepsilon} a_{\parallel nl}\left(u_h^n, u_h^{n+1}, v_h\right) + \gamma \int_{\Gamma_\perp} u_h^{n+1} v_h \,\mathrm{d}s \right) = (u_h^n, v_h). \tag{4.1}$$

This method is called in the sequel the scheme (P).

4.1. Discretization

Let us present the space discretization in a 2D case. We consider a square computational domain $\Omega = [0, 1] \times [0, 1]$. All simulations are performed on structured meshes. Let us introduce the Cartesian, homogeneous grid

$$x_i = i/N_x \ , \ \ 0 \le i \le N_x, \quad y_j = j/N_y \ , \ \ 0 \le j \le N_y, \tag{4.2}$$

where N_x and N_y are positive even constants, corresponding to the number of discretization intervals in the x- resp. y-direction. The corresponding mesh-sizes are denoted by $h_x > 0$ resp. $h_y > 0$. Choosing a $\mathbb{Q}_2$ finite element method ($\mathbb{Q}_2$-FEM), based on the following quadratic base functions (see [29] for more details on rectangular finite elements)

$$\theta_{x_i} = \begin{cases} \dfrac{(x-x_{i-2})(x-x_{i-1})}{2h_x^2} & x \in [x_{i-2}, x_i], \\ \dfrac{(x_{i+2}-x)(x_{i+1}-x)}{2h_x^2} & x \in [x_i, x_{i+2}], \\ 0 & \text{else} \end{cases}, \quad \theta_{y_j} = \begin{cases} \dfrac{(y-y_{j-2})(y-y_{j-1})}{2h_y^2} & y \in [y_{j-2}, y_j], \\ \dfrac{(y_{j+2}-y)(y_{j+1}-y)}{2h_y^2} & y \in [y_j, y_{j+2}], \\ 0 & \text{else} \end{cases} \tag{4.3}$$

for even i, j and

$$\theta_{x_i} = \begin{cases} \dfrac{(x_{i+1}-x)(x-x_{i-1})}{h_x^2} & x \in [x_{i-1}, x_{i+1}], \\ 0 & \text{else} \end{cases}, \quad \theta_{y_j} = \begin{cases} \dfrac{(y_{j+1}-y)(y-y_{j-1})}{h_y^2} & y \in [y_{j-1}, y_{j+1}], \\ 0 & \text{else} \end{cases} \tag{4.4}$$

for odd i, j, we define the space

$$W_h := \{v_h = \sum_{i,j} v_{ij}\, \theta_{x_i}(x)\, \theta_{y_j}(y)\}.$$

The spaces $\mathcal{V}_h$ and $\mathcal{L}_h$ are then defined by

$$\mathcal{V}_h = \mathcal{W}_h, \quad \mathcal{L}_h = \{q_h \in \mathcal{V}_h, \text{ such that } q_h|_{\Gamma_{\text{in}}} = 0\}.$$

The matrix elements are computed using the 2D Gauss quadrature formula, with 3 points in the x and y direction:

$$\int_{-1}^{1}\int_{-1}^{1} f(\xi,\eta)\,\mathrm{d}\xi\,\mathrm{d}\eta = \sum_{i,j=-1}^{1} \omega_i\omega_j f(\xi_i,\eta_j), \tag{4.5}$$

where $\xi_0 = \eta_0 = 0$, $\xi_{\pm1} = \eta_{\pm1} = \pm\sqrt{\frac{3}{5}}$, $\omega_0 = 8/9$ and $\omega_{\pm1} = 5/9$, which is exact for polynomials of degree 5. Linear systems obtained for all methods in these numerical experiments are solved using a LU decomposition, implemented by the MUMPS library.

4.2. Numerical tests

4.2.1. Known analytical solution

We consider first a numerical test case with a known solution to Problem (2.2) modified by the addition of a source term to the right-hand side of the first equation. The important feature is that it has to be constant along the b field lines in the limit $\varepsilon \to 0$. We thus construct the solution u as the sum of a limit solution u_0 and a perturbation proportional to ε:

$$u = u_0 + \varepsilon q \tag{4.6}$$

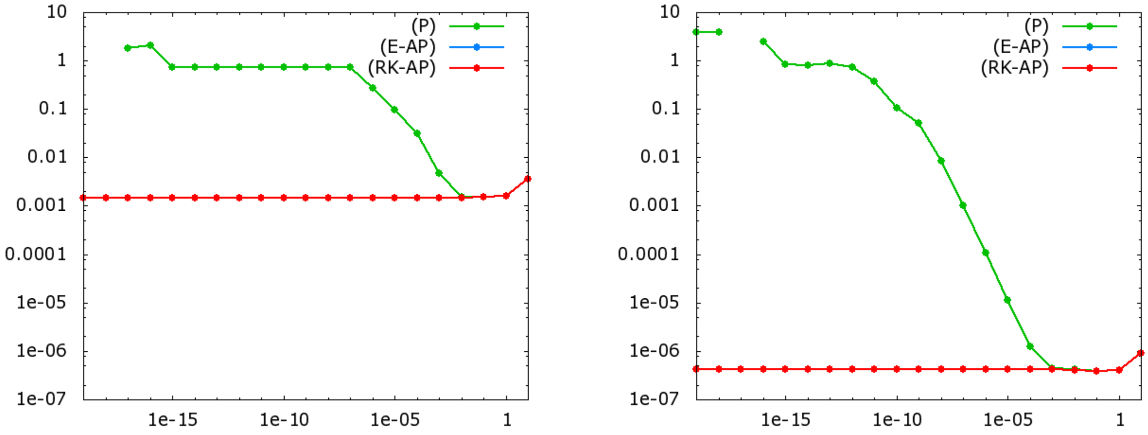
with

$$u_0 = \left(\cos\left(\pi y + \alpha(y^2-y)\cos(\pi x)\right) + 4\right) T_m \mathrm{e}^{-t}, \tag{4.7}$$

$$q = (u_0)^{-3/2}\sin(3\pi x)/3\pi. \tag{4.8}$$

Here α and and T_m are some numerical parameters. Since u_0 is constant along the b field lines, this field can be constructed using the following implication

$$\nabla_{\|} u_0 = 0 \quad \Rightarrow \quad b_x\frac{\partial u_0}{\partial x} + b_y\frac{\partial u_0}{\partial y} = 0, \tag{4.9}$$



(a) $h = 0.1$

(b) $h = 0.00625$

FIGURE 3. Relative L^2-errors between the exact solution u^ε and the computed solution for the scheme (P), Euler AP method (E_{AP}) and DIRK AP scheme (RK_{AP}) as a function of ε and for $h = 0.1$ resp. $h = 0.00625$. The time step is $\tau = 10^{-6}$.

which yields

$$b = \frac{B}{|B|}, \quad B = \begin{pmatrix} \alpha(2y-1)\cos(\pi x) + \pi \\ \pi\alpha(y^2 - y)\sin(\pi x) \end{pmatrix}. \tag{4.10}$$

Note that we have $B \neq 0$ in the computational domain $\Omega = (0, 1)$. In our simulations we set $\alpha = 1$ so that the direction of the anisotropy is variable. The problem (2.2) is slightly modified in order to accommodate the exact solution above. We add the appropriate force term in the first equation of (2.2). Note that this term is of order $\mathcal{O}(1)$ due to the construction of u. We put $\gamma = 1$ in the boundary conditions in (2.2) so that they are automatically satisfied. In this setting we expect both Asymptotic-Preserving methods (E_{AP}) and (RK_{AP}) to converge in the optimal rate, independently on ε and b.

First we test the space convergence of the methods. To do this we choose a small time step such that the time discretization error is much smaller than the space discretization error. We then vary the mesh size and perform simulations for 100 time steps. The results are summarized in Table 1 and Figure 3. All three methods give as expected the third order space convergence in the L_2-norm for large values of ε. Moreover, due to the extremely small time step, the numerical precision is the same, even if one uses second or first order methods. For small values of ε only the Asymptotic Preserving schemes give good numerical solutions.

Finally we test the time convergence of the methods. To do this we choose a small mesh size such that the space discretization error is smaller than the time discretization error. We then vary the time step and perform simulations on a fixed grid. The results are summarized in Table 2 and Figures 4 and 5. Note that the (RK_{AP}) scheme is of second order in time as long as the error due to the time discretization dominates the error induced by the space discretization. The straightforward scheme (P), defined in (4.1), works well and is of first order, as long as ε is close to one. The (E_{AP}) scheme is of first order for all values of the anisotropic parameter. Also note that while the (RK_{AP}) scheme demands twice more computational time than the (E_{AP}) scheme, it gives much better precision. In order to achieve a relative error of the order of 10^{-4} for $\varepsilon = 1$ it suffices to take a time step of $\tau = 0.05$ in the RK-scheme. A comparable accuracy with (E_{AP}) is obtained for a time step 16 times smaller. In the case of $\varepsilon = 10^{-10}$ the ratio is 32.

TABLE 1. The absolute error of u in the L^2-norm for different mesh sizes and $\varepsilon = 1$ or $\varepsilon = 10^{-10}$, using the scheme (P) and the two proposed AP-schemes for a time step of $\tau = 10^{-6}$ s and at instant $t = 10^{-4}$, with $T_m = 1$.

h	L^2-error $\varepsilon = 1$			h	L^2-error $\varepsilon = 10^{-10}$		
	P	E_{AP}	RK_{AP}		P	E_{AP}	RK_{AP}
0.1	1.60×10^{-3}	1.60×10^{-3}	1.60×10^{-3}	0.1	7.3×10^{-1}	1.47×10^{-3}	1.47×10^{-3}
0.05	2.02×10^{-4}	2.02×10^{-4}	2.02×10^{-4}	0.05	7.3×10^{-1}	2.04×10^{-4}	2.04×10^{-4}
0.025	2.55×10^{-5}	2.55×10^{-5}	2.55×10^{-5}	0.025	7.3×10^{-1}	2.65×10^{-5}	2.65×10^{-5}
0.0125	3.2×10^{-6}	3.2×10^{-6}	3.2×10^{-6}	0.0125	4.9×10^{-1}	3.3×10^{-6}	3.3×10^{-6}
0.00625	4.0×10^{-7}	4.0×10^{-7}	4.0×10^{-7}	0.00625	1.04×10^{-1}	4.2×10^{-7}	4.2×10^{-7}

TABLE 2. The absolute error of u in the L^2-norm for different time step using the scheme (P) and two proposed AP-schemes for mesh size 200×200 at time $t = 0.1$ with $T_m = 1$.

τ	L^2-error $\varepsilon = 1$			τ	L^2-error $\varepsilon = 10^{-10}$		
	P	E_{AP}	RK_{AP}		P	E_{AP}	RK_{AP}
0.1	1.57×10^{-2}	1.57×10^{-2}	2.52×10^{-3}	0.1	6.14×10^{-1}	1.57×10^{-2}	2.90×10^{-4}
0.05	8.28×10^{-3}	8.28×10^{-3}	1.93×10^{-4}	0.05	6.30×10^{-1}	8.22×10^{-3}	7.21×10^{-5}
0.025	4.25×10^{-3}	4.25×10^{-3}	2.62×10^{-5}	0.025	6.92×10^{-1}	4.22×10^{-3}	1.80×10^{-5}
0.0125	2.37×10^{-3}	2.37×10^{-3}	6.54×10^{-6}	0.0125	7.08×10^{-1}	2.36×10^{-3}	4.91×10^{-6}
0.00625	1.08×10^{-3}	1.08×10^{-3}	1.50×10^{-6}	0.00625	7.26×10^{-1}	1.08×10^{-3}	1.15×10^{-6}
0.003125	5.44×10^{-4}	5.44×10^{-4}	4.08×10^{-7}	0.003125	7.42×10^{-1}	5.40×10^{-4}	3.43×10^{-7}
0.0015625	2.76×10^{-4}	2.76×10^{-4}	2.07×10^{-7}	0.0015625	6.42×10^{-1}	2.74×10^{-4}	2.05×10^{-7}

To conclude, one can remark that the asymptotic-preserving schemes, (E_{AP}) and (RK_{AP}), are uniformly accurate with respect to the perturbation parameter ε. This essential feature can be very useful in situations where the anisotropy is variable in space, *i.e.* the parameter $\varepsilon(x)$ is x-dependent. No mesh-adaptation is any more needed in these cases, a simple Cartesian grid enables accurate results, with no regard to the ε-values.

4.2.2. Initial Gaussian peak

The second investigated test is the evolution, according to problem (2.2), of the following initial Gaussian peak, located in the middle of the computational domain $\Omega = (0, 1)$:

$$u^0 = \frac{T_m}{2}\left(1 + \mathrm{e}^{-50(x-0.5)^2 - 50(y-0.5)^2}\right), \tag{4.11}$$

where $T_m = 10^5$ K is the maximal temperature in the domain and the anisotropy direction is given as in the previous tests. We perform numerical experiments with the choice of $\varepsilon = 1$. Note that a strong anisotropy is still present in the system due to large value of u^0. In fact one could rescale the problem and look for $\tilde{u_h} = u_h/T_m$. In this case the initial condition $\tilde{u}^0$ would be of the order 1 and the rescaled anisotropy strength would be $\tilde{\varepsilon} = T_m^{-5/2} = 10^{-12.5}$. We choose the time step $\tau = 0.01$ and perform numerical simulations on a fixed 50×50 grid with the final time set to 15 s. The time step is big compared to the time scale induced by the initial condition. Indeed, after the first iteration of the algorithm the numerical solution immediately falls into the space of functions which are almost constant in the direction of the anisotropy (see Fig. 7). The evolution of the numerical solution consists of two phases. The first one, in which the parallel components of the diffusion operator dominate, is characterized by the exponential decay of $||u_h||_{L_2(\Omega)}$, $\min(u_h)$ and $\max(u_h)$ (see Fig. 6). When u_h reaches some critical value, the parallel part of the diffusion operator becomes smaller than the perpendicular one. The direction of the strong diffusion is now inverted and the numerical solution aligns itself rather with the perpendicular direction. The minimum, maximum as well as the L^2-norm of u_h continue

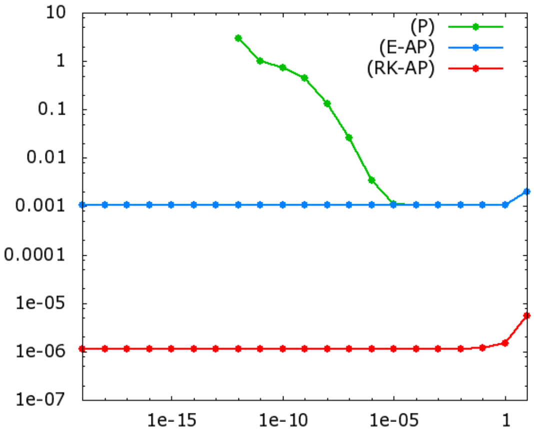
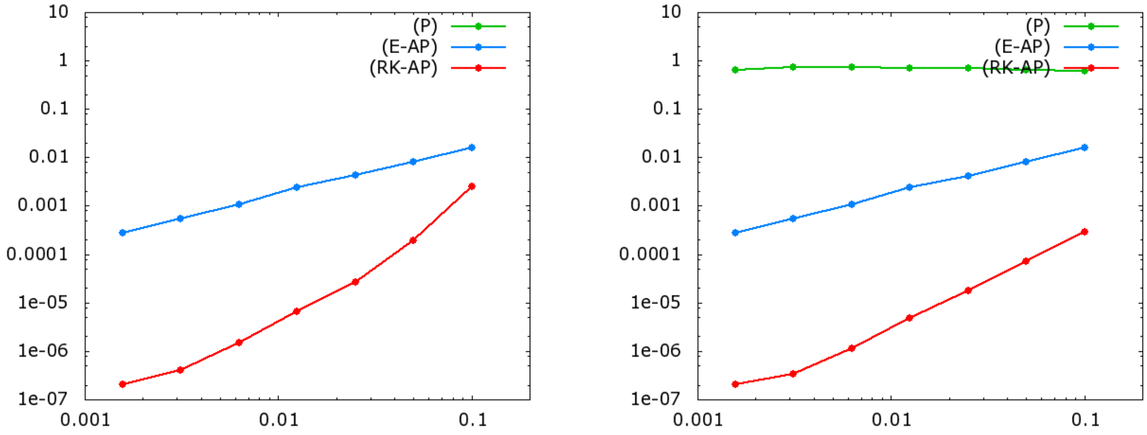


FIGURE 4. Relative L^2-errors between the exact solution u^ε and the computed solution with the scheme (P), the Euler-AP method (E_{AP}) and the DIRK-AP scheme (RK_{AP}) as a function of ε and for $\tau = 0.00625$. The spacial grid is 200×200.



(a) $\varepsilon = 1$

(b) $\varepsilon = 10^{-10}$

FIGURE 5. Relative L^2-errors between the exact solution u^ε and the computed solution with the scheme (P), the Euler-AP method (E_{AP}) and the DIRK-AP scheme (RK_{AP}) as a function of τ, and for $\varepsilon = 1$ resp. $\varepsilon = 10^{-10}$ and a mesh with 200×200 points. Note that for $\varepsilon = 1$ the both schemes (P) and E_{AP} give the same precision.

to approach zero, but the decay is no longer exponential. The L^2-norm and the maximal value remain close to each other and almost constant in time. The minimal value of u_h, as well as the boundary-values decrease much faster.

4.2.3. The case of domain with Γ_{in} or Γ_{out} not perpendicular to the field b

All our numerical experiments so far have been conducted under the geometrical assumptions on the domain Ω and the filed b such that $b||n$ on $\Gamma_{\text{in}} \cup \Gamma_{\text{out}}$. Without this, the solution to (2.2) is not guaranteed to stay positive at the long time. Let us illustrate this by a numerical experiment. Let $\Omega = (0,1)^2$ and the anisotropy

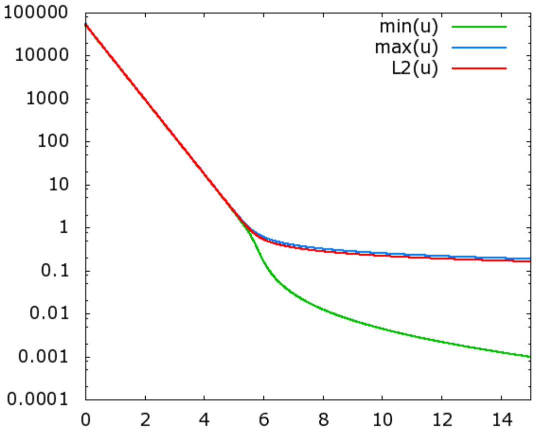


FIGURE 6. $\min(u_h)$, $\max(u_h)$ and $||u_h||_{L^2(\Omega)}$ as a function of time for the Gaussian peak experiment, for $T_m = 10^5$ and $\varepsilon = 1$. Time step is $\tau = 0.01$ s and a mesh size of 50×50.

direction be given as

$$b = \frac{B}{|B|}, \quad B = \begin{pmatrix} 1+x \\ 100y(y-1)(y-0.5) \end{pmatrix},$$

so that Γ_{in} (resp. Γ_{out}) is the part of the boundary at $x = 0$ (resp. $x = 1$), as in our previous numerical experiments, but b is no longer parallel to n $\Gamma_{\text{in}} \cup \Gamma_{\text{out}}$, except for the points $(0, 0.5)$ and $(0, 0.5)$. We perform our numerical simulations taking $u^0 = 1$ as the initial condition. We put $\varepsilon = \gamma = 1$, fix a time step to 0.0125 and perform numerical simulations for different mesh sizes ranging from 20×20 to 1280×1280 points. The simulations are stopped at a critical time t_C, when negative values of u appear somewhere in Ω. It turns out, that it happens at $(1, 0.5)$ which is the only point on Γ_{out} where the b field is perpendicular to the boundary. The value of t_C seems to converge to $t_C \approx 4.65$ as we refine the mesh. Figure 8 shows the field b, numerical results at t_C and a minimal value of u in the computational domain as a function of time for different meshes. The results are comparable for all numerical methods proposed – including the singular perturbation scheme. Decreasing the value of ε seems however to increase the value of t_C.

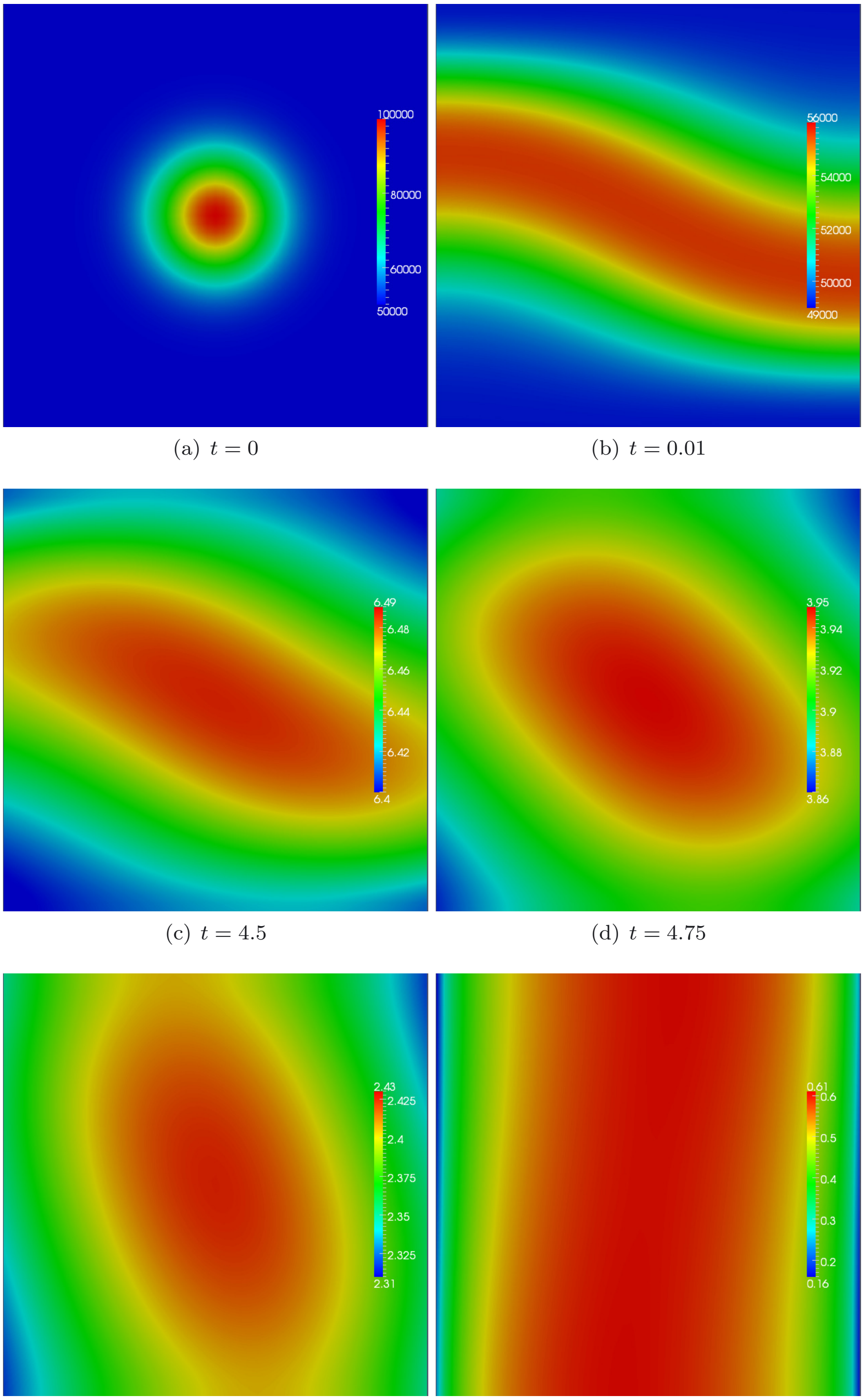
Having said that, we would like to emphasize that out AP method works fine in the general geometry (as long as the solution stays positive). For example, we have run the following test case, which is a slight modification of that of Section 4.2.1. We have kept the domain Ω to be the unit square and changed the field b from (4.10) to

$$b = \frac{B}{|B|}, \quad B = \begin{pmatrix} \alpha(2y-1)\cos(\pi(x-1/2)) + \pi \\ \pi\alpha(y^2-y)\sin(\pi(x-1/2)) \end{pmatrix}$$

so that b is no longer perpendicular to Γ_{in} nor to Γ_{out}. The problem (2.2) is supplied with force terms so that it has the solution $u^\varepsilon = u^0 + \varepsilon q$ with

$$u^0 = \left(\cos\left(\pi y + \alpha(y^2-y)\cos(\pi(x-1/2))\right) + 4\right) T_m \mathrm{e}^{-t}$$

being a slight modification of (4.7) and q defined as in (4.8). The numerical results with this test case are practically indistinguishable from those of Section 4.2.1 and are thus not reported here.



(a) $t = 0$

(b) $t = 0.01$

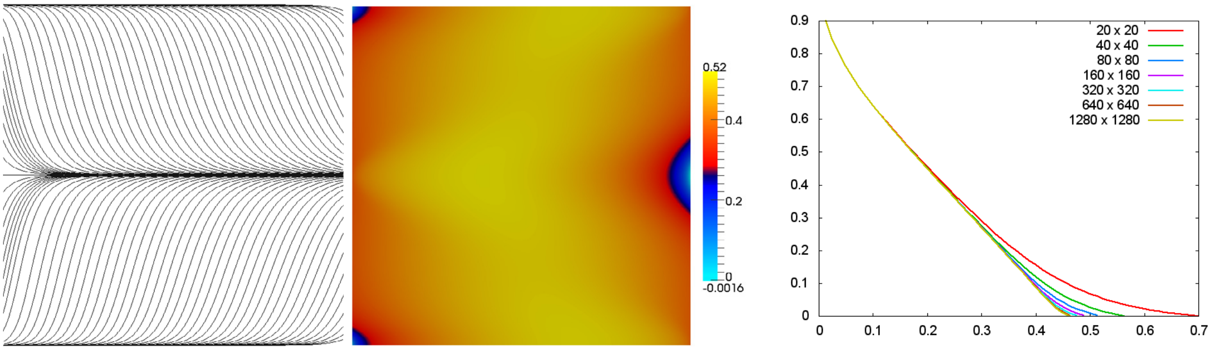
(c) $t = 4.5$

(d) $t = 4.75$

(e) $t = 5$

(f) $t = 6$

FIGURE 7. Numerical solution at different time steps for the Gaussian peak experiment, for $T_m = 10^5$ and $\varepsilon = 1$. Time step is $\tau = 0.01$ s and a mesh size of 50×50.



(a) field b (b) u (c) $\min u$

FIGURE 8. Anisotropy direction, numerical solution at t_C and $\min u$ as a function of time for different mesh sizes in a numerical experiment with negative values of u appearing.

5. Conclusion

The here presented Asymptotic-Preserving scheme proves to be an efficient, general and easy to implement numerical method for solving nonlinear, strongly anisotropic parabolic problems. This kind of problems occur in several important applications, as for example magnetically confined fusion plasmas. The method is based on a reformulation of the problem, initially introduced by the authors in an elliptic framework, and a careful linearization as well as time-discretization of the resulting equation, which does not destroy the AP-properties of the space-discretization. Numerical experiments show clearly the advantages of such an AP-scheme.

Acknowledgements. The authors would like to thank Michel Pierre and Giacomo Dimarco for stimulating and useful discussions.

References

[1] D. Aronson, The porous medium equation. Nonlinear Diffusion Problems, edited by A. Fasano, M. Primicerio. *Lect. Notes Math.* **1224** (1986) 1–46.

[2] S.F. Ashby, W.J. Bosl, R.D. Falgout, S.G. Smith, A.F. Tompson and T.J. Williams, A Numerical Simulation of Groundwater Flow and Contaminant Transport on the CRAY T3D and C90 Supercomputers. *Int. J. High Performance Comput. Appl.* **13** (1999) 80–93.

[3] P. Basser and D. Jones, Diffusion-tensor mri: theory, experimental design and data analysis–a technical review. *NMR Biomedicine* **15** (2002) 456–467.

[4] C. Beaulieu, The basis of anisotropic water diffusion in the nervous system–a technical review. *NMR Biomedicine* **15** (2002) 435–455.

[5] B. Berkowitz, Characterizing flow and transport in fractured geological media: A review. *Adv. Water Resources* **25** (2002) 861–884.

[6] P. Degond, F. Deluzet, A. Lozinski, J. Narski and C. Negulescu, Duality-based asymptotic-preserving method for highly anisotropic diffusion equations. *Commun. Math. Sci.* **10** (2012) 1–31.

[7] P. Degond, A. Lozinski, J. Narski and C. Negulescu, An asymptotic-preserving method for highly anisotropic elliptic equations based on a micro-macro decomposition. *J. Comput. Phys.* **231** (2012) 2724–2740.

[8] Y. Dubinskii, Some integral inequalities and the solvability of degenerate quasi-linear elliptic systems of differential equations. *Matematicheskii Sbornik* **106** (1964) 458–480.

[9] Y. Dubinskii, Weak convergence for nonlinear elliptic and parabolic equations. *Matematicheskii Sbornik* **109** (1965) 609–642.

[10] L.C. Evans and R.F. Gariepy, Measure theory and fine properties of functions. *Stud. Adv. Math.* CRC press (1992).

[11] S. Günter, K. Lackner C. Tichmann, Finite element and higher order difference formulations for modelling heat transport in magnetised plasmas. *J. Comput. Phys.* **226** (2007) 2306–2316.

[12] E. Hairer and G. Wanner, Solving Ordinary Differential Equations II: Stiff and Differential-Algebraic Problems. *Springer Ser. Comput. Math.* Springer-Verlag, New York (1987).

[13] H. Jian and B. Song, Solutions of the anisotropic porous medium equation in $\mathbb{R}^n$ under an l^1-initial value. *Nonlinear Anal.* **64** (2006) 2098–2111.

[14] S. Jin, Efficient asymptotic-preserving (AP) schemes for some multiscale kinetic equations. *SIAM J. Sci. Comput.* **21** (1999) 441–454.

[15] J. Lions, Quelques méthodes de résolution des problèmes aux limites non linéaires. Gauthier-Villars (1969).

[16] J.-L. Lions and E. Magenes, Non-homogeneous boundary value problems and applications. Vol. I. Translated from the French by P. Kenneth, Die Grundlehren der mathematischen Wissenschaften, Band 181. Springer-Verlag, New York (1972).

[17] H. Lutjens and J. Luciani, The xtor code for nonlinear 3d simulations of mhd instabilities in tokamak plasmas. *J. Comput. Phys.* **227** (2008) 6944–6966.

[18] A. Mentrelli and C. Negulescu, Asymptotic preserving scheme for highly anisotropic, nonlinear diffusion equations *J. Comput. Phys.* **231** (2012) 8229–8245.

[19] W. Park, E. Belova, G. Fu, X. Tang, H. Strauss L. Sugiyama, Plasma simulation studies using multilevel physics models. *Phys. Plasmas* **6** (1999) 1796.

[20] P. Perona and J. Malik, Scale-space and edge detection using anisotropic diffusion. Pattern Analysis and Machine Intelligence, IEEE Trans. **12** (1990) 629–639.

[21] M. Pierre, personal e-mail (2011).

[22] J. Narski, Anisotropic finite elements with high aspect ratio for an Asymptotic Preserving method for highly anisotropic elliptic equation. Preprint `arXiv:1302.4269` (2013).

[23] J. Narski and M. Ottaviani, Asymptotic Preserving scheme for strongly anisotropic parabolic equations for arbitrary anisotropy direction. Preprint `arXiv:1303.5219` (2013).

[24] J. Simon, Compact sets in the space $L^p(0,T;B)$. *Ann. Mat. Pura Appl.* **146** (1987) 65–96.

[25] P. Tamain, *Etude des flux de matière dans le plasma de bord des tokamaks*. Ph.D. Thesis, Marseille 1 (2007).

[26] J. Vázquez, The porous medium equation: mathematical theory. Oxford University Press, USA (2007).

[27] J. Weickert, Anisotropic diffusion in image processing. European Consortium for Mathematics in Industry. B.G. Teubner, Stuttgart (1998).

[28] J. Wesson, Tokamaks. Oxford University Press, New York (1987).

[29] O.C. Zienkiewicz and R.L. Taylor, The finite element method. Vol. 1. Butterworth-Heinemann, Oxford (2000).

[30] J. Wloka, Partial diflerential equations. Cambridge University Press (1987).