# Long-time behaviour of hybrid finite volume schemes for advection-diffusion equations: linear and nonlinear approaches

Claire Chainais-Hillairet, Maxime Herda, Simon Lemaire, Julien Moatti

# Long-time behaviour of hybrid finite volume schemes for advection-diffusion equations: linear and nonlinear approaches

Claire Chainais-Hillairet*[1], Maxime Herda†[2], Simon Lemaire‡[2], and Julien Moatti§[2]

[1]Univ. Lille, CNRS, Inria, UMR 8524 - Laboratoire Paul Painlevé, F-59000 Lille, France
[2]Inria, Univ. Lille, CNRS, UMR 8524 - Laboratoire Paul Painlevé, F-59000 Lille, France

July 7, 2021

**Abstract**

We are interested in the long-time behaviour of approximate solutions to heterogeneous and anisotropic linear advection-diffusion equations in the framework of hybrid finite volume (HFV) methods on general polygonal/polyhedral meshes. We consider two linear methods, as well as a new, nonlinear scheme, for which we prove the existence and the positivity of discrete solutions. We show that the discrete solutions to the three schemes converge exponentially fast in time towards the associated discrete steady-states. To illustrate our theoretical findings, we present some numerical simulations assessing long-time behaviour and positivity. We also compare the accuracy of the schemes on some numerical tests in the stationary case.

**Keywords:** Finite volume schemes, general meshes, anisotropic advection-diffusion equations, long-time behaviour, entropy method.

**MSC2020:** 65M08, 35K51, 35Q84, 35B40

## Contents



*claire.chainais@univ-lille.fr
†maxime.herda@inria.fr
‡simon.lemaire@inria.fr
§julien.moatti@inria.fr

# 1 Introduction

We are interested in the numerical approximation of linear advection-diffusion equations on bounded domains. These equations constitute the main building block in the modelling of more complex problems stemming from physics (e.g., porous media flows [3], or corrosion models [2]), biology, or electronics (semi-conductor devices modelling [37]). Thus, designing reliable numerical schemes to approximate their solutions is a pre-requisite before discretising more complex models. Our aim here is the preservation of some key physical properties of these equations at the discrete level, on a large variety of meshes.

Let $\Omega$ be an open, bounded, connected polytopal subset of $\mathbb{R}^d$, $d \geq 2$, with boundary $\partial\Omega$ divided into two disjoint open subsets $\Gamma^D$ and $\Gamma^N$, in such a way that $\partial\Omega = \overline{\Gamma^D} \cup \overline{\Gamma^N}$. We consider the following problem: Find $u : \mathbb{R}_+ \times \Omega \to \mathbb{R}$ solution to

$$\begin{cases} \partial_t u - \operatorname{div}(\Lambda(\nabla u + u\nabla\phi)) = f & \text{in } \mathbb{R}_+ \times \Omega, \\ u = g^D & \text{on } \mathbb{R}_+ \times \Gamma^D, \\ \Lambda(\nabla u + u\nabla\phi) \cdot n = g^N & \text{on } \mathbb{R}_+ \times \Gamma^N, \\ u(0,\cdot) = u^{in} & \text{in } \Omega, \end{cases} \tag{1.1}$$

where $n$ is the unit normal vector to $\partial\Omega$ pointing outward $\Omega$, and the data satisfy:

- $\Lambda \in L^\infty(\Omega; \mathbb{R}^{d\times d})$ is a symmetric and uniformly elliptic diffusion tensor: there exist $\lambda_\flat, \lambda_\sharp$ with $0 < \lambda_\flat \leq \lambda_\sharp < \infty$ such that, for a.e. $x$ in $\Omega$, $\xi \cdot \Lambda(x)\xi \geq \lambda_\flat |\xi|^2$ and $|\Lambda(x)\xi| \leq \lambda_\sharp |\xi|$ for all $\xi \in \mathbb{R}^d$;

- $\phi \in C^1(\overline{\Omega})$ is a regular potential from which derives the advection field $V^\phi := -\Lambda\nabla\phi$, assumed to satisfy $V^\phi \in H(\mathrm{div};\Omega)$;
- $f \in L^2(\Omega)$ is a source term;
- $g^D \in H^{\frac{1}{2}}(\Gamma^D)$ is a Dirichlet datum, assumed to be the trace on $\Gamma^D$ of $u^D \in H^1(\Omega)$ satisfying $\|u^D\|_{H^1(\Omega)} \le C\|g^D\|_{H^{1/2}(\Gamma^D)}$ for a given $C > 0$;
- $g^N \in L^2(\Gamma^N)$ is a Neumann datum;
- $u^{in} \in L^2(\Omega)$ is an initial datum.

When $|\Gamma^D| = 0$, we assume that the compatibility condition $\int_\Omega f + \int_{\partial\Omega} g^N = 0$ holds true. We denote by $M$ the initial mass, $M = \int_\Omega u^{in}$, and we know that it is preserved along time: $\int_\Omega u(t) = M$ for almost every $t > 0$. For further use, we let $u^M := \frac{M}{|\Omega|} \in \mathbb{R}$, and we refer to this quantity as the mass lifting. Advection-diffusion models of the form (1.1) enjoy certain structural properties. First, when the data $f$, $g^D$, $g^N$, and $u^{in}$ are positive, then the solution $u$ is also positive. Second, the asymptotics $t \to \infty$, the so-called long-time behaviour of the solutions, is well understood, see [5, 11, 12, 36] for related models. Indeed, the solution $u$ to (1.1) converges exponentially fast when $t \to \infty$ towards the steady-state $u^\infty$, solution to the stationary problem

$$\begin{cases} -\operatorname{div}\left(\Lambda(\nabla u^\infty + u^\infty\nabla\phi)\right) = f & \text{in } \Omega, \\ u^\infty = g^D & \text{on } \Gamma^D, \\ \Lambda(\nabla u^\infty + u^\infty\nabla\phi)\cdot n = g^N & \text{on } \Gamma^N, \end{cases} \tag{1.2}$$

with additional constraint $\int_\Omega u^\infty = M$ when $|\Gamma^D| = 0$. The question of long-time behavior has been widely studied in the context of many-particle systems, for which the second law of thermodynamics ensures a relaxation of the transient phenomena towards an equilibrium. From a mathematical point of view, this evolution is strongly related to the dissipation of an entropy functional. Such a vision based on entropy dissipation has given birth to the so-called entropy method. As highlighted by Arnold et al. in [1], the successful use of the entropy method in kinetic theory paves the way to extended applications on various dissipative systems. We refer the reader to the book [29] of Jüngel for a presentation of some of these applications. In [5], Bodineau et al. proposed an entropy functional adapted to drift-diffusion equations with non-homogeneous Dirichlet boundary conditions. A direct adaptation of their method allows to conclude in the present case on the exponential convergence in time of the solution to Problem (1.1) towards the solution of Problem (1.2).

Under appropriate assumptions on the data (a sufficient condition, also valid for more general advection fields, is to assume that $\operatorname{div} V^\phi \ge 0$ a.e. in $\Omega$ and $V^\phi \cdot n \le 0$ a.e. on $\Gamma^N$), the stationary Problem (1.2) is coercive and its well-posedness is straightforward. It turns out that such assumptions on the data are not fulfilled, the problem is still coercive in the new unknown $\rho^\infty = u^\infty e^\phi$ so one can conclude on well-posedness invoking the standard inf-sup theory (or equivalently by solving the problem in the new unknown). Concerning the evolution Problem (1.1), the same arguments show the existence and uniqueness of a global weak solution. For general advection fields, we refer the reader to the results of Droniou [17] (for mixed Dirichlet-Neumann boundary conditions), and Droniou and Vázquez [22] (for pure Neumann boundary conditions) for detailed statements about well-posedness and regularity of the solutions.

When it comes to numerical approximation, the accuracy of the method is not the only important feature. In some applications (e.g., in subsurface modelling, where the mesh often results from seismic analysis), the mesh must be taken as a datum of the problem, and the numerical method needs to be adapted so as to handle potentially fairly general meshes. In some other applications (e.g., power plant simulation), the preservation of the positivity of the solutions (or better, of the monotonicity

properties of the equation) is an important quality criterion. In yet some other applications (e.g., nuclear waste repository management), finally, the reliability of the simulations in very large time proves to be crucial for sustainability purposes. The positivity and long-time behaviour of discrete solutions have been closely studied in the context of standard two-point flux approximation (TPFA) finite volume schemes, for isotropic diffusion (i.e., $\Lambda = \lambda I_d$ with $\lambda : \Omega \to \mathbb{R}_+^\star$) on orthogonal meshes. In [26], Filbet and Herda studied the long-time behaviour of a TPFA scheme for nonlinear boundary-driven Fokker–Planck equations, adapting to the discrete setting the arguments of [5]. In [13], Chainais-Hillairet and Herda proved on a variety of models that a whole family of TPFA schemes (the so-called B-schemes) preserves the exponential decay towards discrete steady-states. The results of [26] and [13] are valid for general advection fields, and a choice of data $|\Gamma^D| > 0$, $f = 0$, $g^D > 0$, $g^N = 0$, and $u^{in} \geq 0$. We also refer to [31, 14, 28, 7] for related schemes and similar issues. However, these TPFA schemes suffer from an intrinsic limitation: the mesh needs to be $\Lambda$-orthogonal, which, in practice, restricts their use to isotropic diffusion tensors and (standard) orthogonal meshes. In order to overcome this limitation, several finite volume methods using auxiliary unknowns have been designed (cf. [19] for a presentation of some of these schemes). As highlighted by Droniou in [19], these methods however suffer from a lack of monotonicity, and so do not preserve the positivity of discrete solutions. As a possible remedy, Cancès and Guichard introduced in [10], for a class of models encompassing (1.1) for pure Neumann boundary conditions and a choice of data $f = 0$, $g^N = 0$, and $u^{in} \geq 0$ with $M > 0$, a nonlinear vertex approximate gradient (VAG) scheme, designed so as to preserve at the discrete level the positivity of the solutions and the entropy structure of the models, for arbitrary anisotropic diffusions and general meshes. Following the same ideas, Cancès et al. devised and analysed in [9] a (nonlinear) positivity-preserving discrete duality finite volume (DDFV) scheme, whose discrete entropy structure and long-time behaviour were fully studied in [8], based on the adaptation to the discrete setting of nonlinear functional inequalities. The DDFV scheme at hand is however limited to the two-dimensional case, and its adaptation to a three-dimensional framework seems difficult (cf. [19]). On yet another level, it is known that, provided adequate assumptions hold on the data, the solutions to Problem (1.1) are regular in space (at least locally). This suggests that the use of high-order methods shall be an interesting track in order to increase the accuracy at fixed computational cost. Recently introduced by Di Pietro et al. in [16], hybrid high-order (HHO) methods can be seen as an arbitrary-order generalisation of hybrid finite volume (HFV) schemes, that were introduced by Eymard et al. in [25] as yet another way to overcome the limitations of TPFA schemes. HFV methods hinge on cell and face unknowns (whence the vocable hybrid), and as such benefit from a unified 2D/3D formulation. HFV methods have also been bridged to the larger family of hybrid mimetic mixed (HMM) methods in [21]. In view of the above elements, the study of HFV methods appears to be a natural first step in order to design structure-preserving high-order (HHO) schemes for Problem (1.1), that shall both increase the accuracy at fixed computational burden, and preserve the key properties (positivity and long-time behaviour) of the model at hand.

In this article, we study three different HFV schemes for Problem (1.1). The first one is the HFV variant of the HMM family of schemes introduced and analysed in the stationary setting by Beirão da Veiga et al. in [4, 18] (note that an arbitrary-order (HHO) generalisation of this scheme has been proposed in [15]). It is a linear scheme, based on a discretisation of the diffusive and advective fluxes, that is well-posed under a coercivity condition. The second scheme is also a linear one. Its construction is based on exponential fitting, and takes inspiration from ideas in [6] (it also shares some features with the works [30, 32] and [26], which cover general advection fields). This scheme is unconditionally coercive. These two linear schemes are not expected to preserve positivity, which motivates the introduction of the third method. For pure Neumann boundary conditions, and a choice of data $f = 0$, $g^N = 0$, and $u^{in} \geq 0$ with $M > 0$, we introduce a nonlinear HFV scheme, that is devised along the lines of the nonlinear VAG and DDFV schemes of [10] and [9, 8], so as to

guarantee the positivity of discrete solutions. Our first result, stated in Theorem 1, is the existence of (positive) solutions to this nonlinear scheme. In a second time, we investigate the long-time behaviour of the three schemes at hand. We establish in Theorems 2, 3, and 4 the exponential decay in time of their discrete solutions towards the associated discrete steady-states. We numerically validate our theoretical findings on a set of test-cases and, for completeness, we also compare the accuracy of the three schemes on stationary problems.

The article is organised as follows. In Section 2, we introduce the HFV framework (mesh, discrete unknowns and discrete operators) on a steady variable diffusion problem. In Section 3, we introduce the three schemes for the transient advection-diffusion problem, and we discuss their well-posedness. In Section 4, we study the long-time behaviour of the three schemes, and prove exponential decay to equilibrium. In Section 5, we discuss the implementation of the schemes, and provide a numerical validation of our theoretical results, as well as a comparison of the stationary schemes in terms of accuracy. Appendices A and B finally collect some functional inequalities and the proofs of auxiliary results.

## 2 HFV discretisation of a variable diffusion problem

The aim of this section is to recall the HFV framework on a steady variable diffusion problem, which corresponds to (1.2) without any advection term ($\Lambda\nabla\phi = 0$). For a detailed presentation of the method, we refer the reader to [25].

### 2.1 Mesh and discrete unknowns

The definitions and notation we adopt for the discretisation are essentially the same as in [25]. A discretisation of the (open, bounded, connected) polytopal set $\Omega \subset \mathbb{R}^d$, $d \geq 2$, is defined as a triplet $\mathcal{D} := (\mathcal{M}, \mathcal{E}, \mathcal{P})$, where:

- $\mathcal{M}$ (the mesh) is a partition of $\Omega$, i.e., a finite family of nonempty disjoint (open, connected) subsets $K$ of $\Omega$ (the mesh cells) such that (i) for all $K \in \mathcal{M}$, $|K| > 0$, and (ii) $\overline{\Omega} = \bigcup_{K\in\mathcal{M}} \overline{K}$.
- $\mathcal{E}$ (the set of faces) is a partition of the mesh skeleton $\bigcup_{K\in\mathcal{M}} \partial K$, i.e., a finite family of nonempty disjoint (open, connected) subsets $\sigma$ of $\overline{\Omega}$ (the mesh faces, or mesh edges if $d = 2$) such that (i) for all $\sigma \in \mathcal{E}$, $|\sigma| > 0$ and there exists $\mathcal{H}_\sigma$ affine hyperplane of $\mathbb{R}^d$ such that $\sigma \subset \mathcal{H}_\sigma$, and (ii) $\bigcup_{K\in\mathcal{M}} \partial K = \bigcup_{\sigma\in\mathcal{E}} \overline{\sigma}$. We assume that, for all $K \in \mathcal{M}$, there exists $\mathcal{E}_K \subset \mathcal{E}$ (the set of faces of $K$) such that $\partial K = \bigcup_{\sigma\in\mathcal{E}_K} \overline{\sigma}$. For $\sigma \in \mathcal{E}$, we let $\mathcal{M}_\sigma := \{K \in \mathcal{M} \mid \sigma \in \mathcal{E}_K\}$ be the set of cells whose $\sigma$ is a face. Then, for all $\sigma \in \mathcal{E}$, either $\mathcal{M}_\sigma = \{K\}$ for a cell $K \in \mathcal{M}$, in which case $\sigma$ is a boundary face ($\sigma \subset \partial\Omega$) and we note $\sigma \in \mathcal{E}_{ext}$, or $\mathcal{M}_\sigma = \{K, L\}$ for two cells $K, L \in \mathcal{M}$, in which case $\sigma$ is an interface and we note $\sigma = K|L \in \mathcal{E}_{int}$.
- $\mathcal{P}$ (the set of cell centres) is a finite family $\{x_K\}_{K\in\mathcal{M}}$ of points of $\Omega$ such that, for all $K \in \mathcal{M}$, (i) $x_K \in K$, and (ii) $K$ is star-shaped with respect to $x_K$. Moreover, we assume that the Euclidean (orthogonal) distance $d_{K,\sigma}$ between $x_K$ and the affine hyperplane $\mathcal{H}_\sigma$ containing $\sigma$ is positive (equivalently, the cell $K$ is strictly star-shaped with respect to $x_K$).

For a given discretisation $\mathcal{D}$, we denote by $h_\mathcal{D} > 0$ the size of the discretisation (the meshsize), defined by $h_\mathcal{D} := \sup_{K\in\mathcal{M}} h_K$ where, for all $K \in \mathcal{M}$, $h_K := \sup_{x,y\in\overline{K}} |x - y|$ is the diameter of the cell $K$. For all $\sigma \in \mathcal{E}$, we let $\overline{x}_\sigma \in \sigma$ be the barycentre of $\sigma$. Finally, for all $K \in \mathcal{M}$, and all $\sigma \in \mathcal{E}_K$, we let $n_{K,\sigma} \in \mathbb{R}^d$ be the unit normal vector to $\sigma$ pointing outward $K$, and $P_{K,\sigma}$ be the (open) pyramid of base $\sigma$ and apex $x_K$ (notice that, when $d = 2$, $P_{K,\sigma}$ is always a triangle). Since $|\sigma|$ and $d_{K,\sigma}$ are positive, we have $|P_{K,\sigma}| = \frac{|\sigma|d_{K,\sigma}}{d} > 0$. We depict on Figure 1 an example of discretisation. Notice

that the mesh cells are not assumed to be convex, and $x_K$ is neither assumed to be the barycentre of $K \in \mathcal{M}$. We consider the following measure of regularity for the discretisation (which is slightly

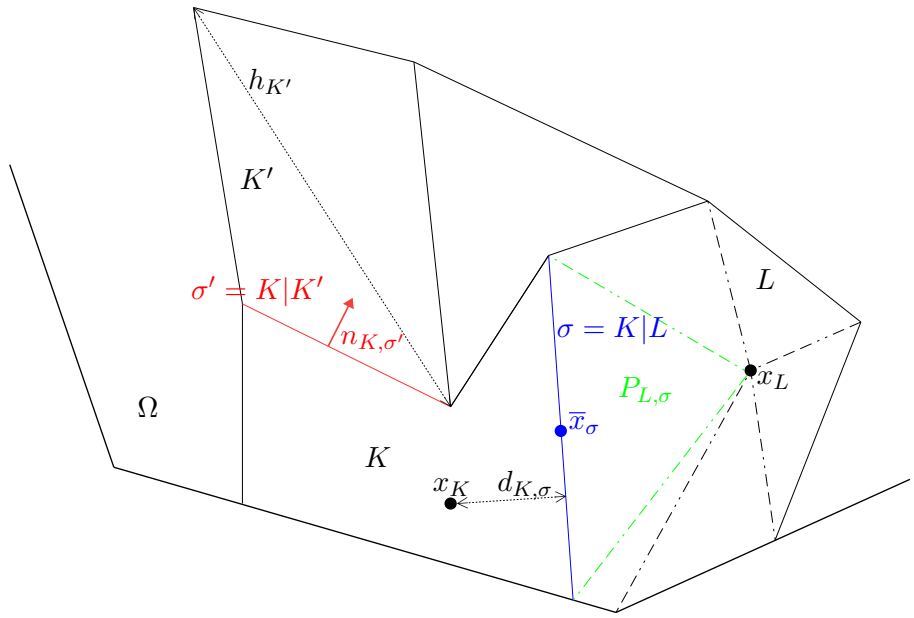



Figure 1: Two-dimensional discretisation and corresponding notations.

stronger than the ones advocated in [25, Eq. (4.1)] or in [20, Eq. (7.8)-(7.9)]):

$$\theta_{\mathcal{D}} := \max\left(\max_{K\in\mathcal{M},\sigma\in\mathcal{E}_K}\frac{h_K}{d_{K,\sigma}}, \max_{\sigma\in\mathcal{E},K\in\mathcal{M}_\sigma}\frac{h_K^{d-1}}{|\sigma|}\right). \tag{2.1}$$

Notice that $\theta_{\mathcal{D}} \geq 1$, and that for all $K \in \mathcal{M}$,

$$h_K^d \geq |K| = \sum_{\sigma\in\mathcal{E}_K}|P_{K,\sigma}| = \sum_{\sigma\in\mathcal{E}_K}\frac{|\sigma|d_{K,\sigma}}{d} \geq \sum_{\sigma\in\mathcal{E}_K}\frac{h_K^d}{d\theta_{\mathcal{D}}^2} = \frac{|\mathcal{E}_K|}{d\theta_{\mathcal{D}}^2}h_K^d.$$

Thus, the number of faces of any mesh cell is uniformly bounded:

$$\forall K \in \mathcal{M}, \qquad |\mathcal{E}_K| \leq d\theta_{\mathcal{D}}^2. \tag{2.2}$$

Also, it is an easy matter to verify that $\max_{\sigma=K|L\in\mathcal{E}_{int}}\max\left(\frac{d_{K,\sigma}}{d_{L,\sigma}}, \frac{d_{L,\sigma}}{d_{K,\sigma}}\right) \leq \theta_{\mathcal{D}}^{\frac{d}{d-1}}$.

We now introduce the set of (hybrid, cell- and face-based) discrete unknowns:

$$\underline{V}_{\mathcal{D}} := \left\{\underline{v}_{\mathcal{D}} = \left((v_K)_{K\in\mathcal{M}}, (v_\sigma)_{\sigma\in\mathcal{E}}\right) : v_K \in \mathbb{R}\ \forall K \in \mathcal{M}, v_\sigma \in \mathbb{R}\ \forall \sigma \in \mathcal{E}\right\}.$$

Given a mesh cell $K \in \mathcal{M}$, we let $\underline{V}_K := \mathbb{R} \times \mathbb{R}^{|\mathcal{E}_K|}$ be the restriction of $\underline{V}_{\mathcal{D}}$ to $K$, and $\underline{v}_K = \left(v_K, (v_\sigma)_{\sigma\in\mathcal{E}_K}\right) \in \underline{V}_K$ be the restriction of a generic element $\underline{v}_{\mathcal{D}} \in \underline{V}_{\mathcal{D}}$ to $K$. Also, for $\underline{v}_{\mathcal{D}} \in \underline{V}_{\mathcal{D}}$, we let $v_{\mathcal{M}} : \Omega \to \mathbb{R}$ and $v_{\mathcal{E}} : \bigcup_{K\in\mathcal{M}}\partial K \to \mathbb{R}$ be the piecewise constant functions such that

$$v_{\mathcal{M}|K} = v_K \text{ for all } K \in \mathcal{M}, \quad \text{and} \quad v_{\mathcal{E}|\sigma} = v_\sigma \text{ for all } \sigma \in \mathcal{E}.$$

In what follows, for any set $X \subset \overline{\Omega}$, we denote by $(\cdot,\cdot)_X$ the inner product in $L^2(X;\mathbb{R}^l)$, for $l \in \{1;d\}$. In particular, we have $(w_{\mathcal{M}}, v_{\mathcal{M}})_\Omega = \sum_{K\in\mathcal{M}} |K| w_K v_K$ and $(w_{\mathcal{E}}, v_{\mathcal{E}})_{\partial\Omega} = \sum_{\sigma\in\mathcal{E}_{ext}} |\sigma| w_\sigma v_\sigma$. For further use, we let $\underline{1}_{\mathcal{D}}$ denote the element of $\underline{V}_{\mathcal{D}}$ with all coordinates equal to 1. Also, given a function $f : \mathbb{R} \to \mathbb{R}$, and with a slight abuse in notation, we denote by $f(\underline{v}_{\mathcal{D}})$ the element of $\underline{V}_{\mathcal{D}}$ whose coordinates are the $(f(v_K))_{K\in\mathcal{M}}$ and the $(f(v_\sigma))_{\sigma\in\mathcal{E}}$. Finally, we let the product $\underline{w}_{\mathcal{D}} \times \underline{v}_{\mathcal{D}}$ denote the element of $\underline{V}_{\mathcal{D}}$ whose $i$-th coordinate is the product of the $i$-th coordinates of $\underline{w}_{\mathcal{D}}$ and $\underline{v}_{\mathcal{D}}$.

When considering mixed Dirichlet-Neumann boundary conditions, we assume that the discretisation $\mathcal{D}$ is compliant with the partition $\partial\Omega = \overline{\Gamma^D} \cup \overline{\Gamma^N}$ of the boundary of the domain, in the sense that the set $\mathcal{E}_{ext}$ can be split into two (necessarily disjoint) subsets $\mathcal{E}_{ext}^D := \{\sigma \in \mathcal{E}_{ext} \mid \sigma \subset \Gamma^D\}$ and $\mathcal{E}_{ext}^N := \{\sigma \in \mathcal{E}_{ext} \mid \sigma \subset \Gamma^N\}$ such that $\mathcal{E}_{ext} = \mathcal{E}_{ext}^D \cup \mathcal{E}_{ext}^N$. Notice that as soon as $|\Gamma^D| > 0$, $|\mathcal{E}_{ext}^D| \geq 1$. We define the following subspace of $\underline{V}_{\mathcal{D}}$, enforcing strongly a homogeneous Dirichlet boundary condition on $\Gamma^D$:

$$\underline{V}_{\mathcal{D},0}^D := \left\{\underline{v}_{\mathcal{D}} \in \underline{V}_{\mathcal{D}} : v_\sigma = 0 \ \forall \sigma \in \mathcal{E}_{ext}^D\right\}.$$

In view of the upcoming analysis, we define a discrete counterpart of the $H^1$ semi-norm. Locally to any cell $K \in \mathcal{M}$, we let, for any $\underline{v}_K \in \underline{V}_K$, $|\underline{v}_K|_{1,K}^2 := \sum_{\sigma\in\mathcal{E}_K} \frac{|\sigma|}{d_{K,\sigma}}(v_K - v_\sigma)^2$. At the global level, for any $\underline{v}_{\mathcal{D}} \in \underline{V}_{\mathcal{D}}$, we let

$$|\underline{v}_{\mathcal{D}}|_{1,\mathcal{D}} := \sqrt{\sum_{K\in\mathcal{M}} |\underline{v}_K|_{1,K}^2}.$$

Notice that $|\cdot|_{1,\mathcal{D}}$ does not define a norm on $\underline{V}_{\mathcal{D}}$, but if $|\underline{v}_{\mathcal{D}}|_{1,\mathcal{D}} = 0$, then there is $c \in \mathbb{R}$ such that $\underline{v}_{\mathcal{D}} = c\,\underline{1}_{\mathcal{D}}$ ($\underline{v}_{\mathcal{D}}$ is constant). Thus, $|\cdot|_{1,\mathcal{D}}$ defines a norm on the space $\underline{V}_{\mathcal{D},0}^D$ as soon as $|\Gamma^D| > 0$, as well as on the space of zero-mass vectors

$$\underline{V}_{\mathcal{D},0}^N := \left\{\underline{v}_{\mathcal{D}} \in \underline{V}_{\mathcal{D}} : \int_\Omega v_{\mathcal{M}} = 0\right\}.$$

For further use, and to allow for a seamless treatment of pure Neumann boundary conditions, we introduce the notation $\underline{V}_{\mathcal{D},0}$, to denote either $\underline{V}_{\mathcal{D},0}^N$ whenever $|\Gamma^D| = 0$, or $\underline{V}_{\mathcal{D},0}^D$ otherwise.

## 2.2 Foundations of the hybrid finite volume discretisation

The HFV method hinges on the definition of a discrete gradient operator $\nabla_{\mathcal{D}}$, that maps any element $\underline{v}_{\mathcal{D}} \in \underline{V}_{\mathcal{D}}$ to a piecewise constant $\mathbb{R}^d$-valued function on the pyramidal submesh of $\mathcal{M}$ formed by all the $P_{K,\sigma}$'s, for $K \in \mathcal{M}$ and $\sigma \in \mathcal{E}_K$. More precisely, for all $K \in \mathcal{M}$, and all $\sigma \in \mathcal{E}_K$,

$$\nabla_{\mathcal{D}}\underline{v}_{\mathcal{D}|K} := \nabla_K \underline{v}_K \quad \text{with} \quad \nabla_K \underline{v}_{K|P_{K,\sigma}} := \nabla_{K,\sigma}\underline{v}_K = G_K\underline{v}_K + S_{K,\sigma}\underline{v}_K \in \mathbb{R}^d,$$

where $G_K\underline{v}_K$ is the consistant part of the gradient given by

$$G_K\underline{v}_K = \frac{1}{|K|}\sum_{\sigma'\in\mathcal{E}_K} |\sigma'|(v_{\sigma'} - v_K)n_{K,\sigma'} = \frac{1}{|K|}\sum_{\sigma'\in\mathcal{E}_K} |\sigma'| v_{\sigma'} n_{K,\sigma'},$$

and $S_{K,\sigma}\underline{v}_K$ is a stabilisation part, given, for some parameter $\eta > 0$, by

$$S_{K,\sigma}\underline{v}_K = \frac{\eta}{d_{K,\sigma}}\big(v_\sigma - v_K - G_K\underline{v}_K \cdot (\overline{x}_\sigma - x_K)\big)n_{K,\sigma}. \tag{2.3}$$

Let us consider the stationary problem (1.2), without any advection term ($\Lambda\nabla\phi = 0$ in $\Omega$). Our aim is to write a HFV discretisation of this steady variable diffusion problem. Locally to any cell $K \in \mathcal{M}$, we introduce the discrete bilinear form $a_K^\Lambda : \underline{V}_K \times \underline{V}_K \to \mathbb{R}$ such that, for all $\underline{u}_K, \underline{v}_K \in \underline{V}_K$,

$$a_K^\Lambda(\underline{u}_K, \underline{v}_K) := \sum_{\sigma\in\mathcal{E}_K} |P_{K,\sigma}| \nabla_{K,\sigma}\underline{v}_K \cdot \Lambda_{K,\sigma}\nabla_{K,\sigma}\underline{u}_K = (\Lambda\nabla_K\underline{u}_K, \nabla_K\underline{v}_K)_K, \tag{2.4}$$

where we set $\Lambda_{K,\sigma} := \frac{1}{|P_{K,\sigma}|}\int_{P_{K,\sigma}} \Lambda$. At the global level, we let $a_{\mathcal{D}}^{\Lambda} : \underline{V}_{\mathcal{D}} \times \underline{V}_{\mathcal{D}} \to \mathbb{R}$ be the discrete bilinear form such that, for all $\underline{u}_{\mathcal{D}}, \underline{v}_{\mathcal{D}} \in \underline{V}_{\mathcal{D}}$,

$$a_{\mathcal{D}}^{\Lambda}(\underline{u}_{\mathcal{D}}, \underline{v}_{\mathcal{D}}) := \sum_{K\in\mathcal{M}} a_K^{\Lambda}(\underline{u}_K, \underline{v}_K) = (\Lambda \nabla_{\mathcal{D}} \underline{u}_{\mathcal{D}}, \nabla_{\mathcal{D}} \underline{v}_{\mathcal{D}})_{\Omega}.$$

The discrete HFV problem then reads: Find $\underline{u}_{\mathcal{D}}^z \in \underline{V}_{\mathcal{D},0}$ such that

$$a_{\mathcal{D}}^{\Lambda}(\underline{u}_{\mathcal{D}}^z, \underline{v}_{\mathcal{D}}) = (f, v_{\mathcal{M}})_{\Omega} + (g^N, v_{\mathcal{E}})_{\Gamma^N} - a_{\mathcal{D}}^{\Lambda}(\underline{u}_{\mathcal{D}}^l, \underline{v}_{\mathcal{D}}) \qquad \forall \underline{v}_{\mathcal{D}} \in \underline{V}_{\mathcal{D},0}, \tag{2.5}$$

where $\underline{u}_{\mathcal{D}}^l \in \underline{V}_{\mathcal{D}}$ is equal

(i) either, when $|\Gamma^D| > 0$, to the HFV interpolate $\underline{u}_{\mathcal{D}}^D$ of the known lifting $u^D$ of the Dirichlet datum $g^D$ (satisfying $|\underline{u}_{\mathcal{D}}^D|_{1,\mathcal{D}} \le C_{l,\Gamma^D} \|g^D\|_{H^{1/2}(\Gamma^D)}$, with $C_{l,\Gamma^D} > 0$ only depending on the discretisation $\mathcal{D}$ through $\theta_{\mathcal{D}}$),

(ii) or, when $|\Gamma^D| = 0$, to $\underline{u}_{\mathcal{D}}^M := u^M \underline{1}_{\mathcal{D}}$, where we recall that $u^M = \frac{M}{|\Omega|}$ is the mass lifting (remark that $a_{\mathcal{D}}^{\Lambda}(\underline{u}_{\mathcal{D}}^M, \underline{v}_{\mathcal{D}}) = 0$ for all $\underline{v}_{\mathcal{D}} \in \underline{V}_{\mathcal{D}}$),

and the approximation of the solution to (1.2) denoted $\underline{u}_{\mathcal{D}}^{\infty} \in \underline{V}_{\mathcal{D}}$ is finally defined as

$$\underline{u}_{\mathcal{D}}^{\infty} := \underline{u}_{\mathcal{D}}^z + \underline{u}_{\mathcal{D}}^l. \tag{2.6}$$

Let us note that the superscript $z$ stands for "zero", while $l$ stands for "lifting".

Problem (2.5) defines a finite volume method, in the sense that it can be equivalently rewritten under a conservative form, with local mass balance, flux equilibration at interfaces, and boundary conditions. For all $K \in \mathcal{M}$, and all $\sigma \in \mathcal{E}_K$, the normal diffusive flux $-\int_\sigma \Lambda \nabla u_{|K}^{\infty} \cdot n_{K,\sigma}$ is approximated by the following numerical flux:

$$F_{K,\sigma}^{\Lambda}(\underline{u}_K) = \sum_{\sigma'\in\mathcal{E}_K} A_K^{\sigma\sigma'}(u_K - u_{\sigma'}), \tag{2.7}$$

where the $A_K^{\sigma\sigma'}$ are defined by

$$A_K^{\sigma\sigma'} := \sum_{\sigma''\in\mathcal{E}_K} |P_{K,\sigma''}|\, y_K^{\sigma''\sigma} \cdot \Lambda_{K,\sigma''} y_K^{\sigma''\sigma'}, \tag{2.8}$$

and the $y_K^{\sigma\sigma'} \in \mathbb{R}^d$ only depend on the geometry of the discretisation $\mathcal{D}$ (see, for example, [25, Eq. (2.22)] for an exact definition with $\eta = \sqrt{d}$). For all $K \in \mathcal{M}$, one can express the local discrete bilinear form $a_K^{\Lambda}$ in terms of the local fluxes $\left(F_{K,\sigma}^{\Lambda}\right)_{\sigma\in\mathcal{E}_K}$: for all $\underline{u}_K, \underline{v}_K \in \underline{V}_K$,

$$a_K^{\Lambda}(\underline{u}_K, \underline{v}_K) = \sum_{\sigma\in\mathcal{E}_K} F_{K,\sigma}^{\Lambda}(\underline{u}_K)(v_K - v_\sigma). \tag{2.9}$$

As for the VAG [10] and DDFV [9, 8] schemes, we can also express the local discrete bilinear form in a different way, which will be useful in the sequel:

$$a_K^{\Lambda}(\underline{u}_K, \underline{v}_K) = \delta_K \underline{v}_K \cdot \mathbb{A}_K \delta_K \underline{u}_K, \tag{2.10}$$

where, for all $\underline{v}_K \in \underline{V}_K$, $\delta_K \underline{v}_K \in \mathbb{R}^{|\mathcal{E}_K|}$ is defined by

$$\delta_K \underline{v}_K := (v_K - v_\sigma)_{\sigma\in\mathcal{E}_K},$$

and $\mathbb{A}_K \in \mathbb{R}^{|\mathcal{E}_K|\times|\mathcal{E}_K|}$ is the symmetric (because $\Lambda$ is) positive semi-definite matrix whose entries are the $A_K^{\sigma\sigma'}$, that can actually be proved to be nonsingular (cf. Lemma 3).

### 2.3 Well-posedness

As for the continuous case, the well-posedness of HFV methods for diffusion problems relies on a coercivity argument. Let $K \in \mathcal{M}$, and reason locally. By definition (2.4) of the local discrete bilinear form $a_K^\Lambda$, and from the bounds on the diffusion coefficient, we have $\lambda_\flat \|\nabla_K \underline{v}_K\|^2_{L^2(K;\mathbb{R}^d)} \le a_K^\Lambda(\underline{v}_K, \underline{v}_K) \le \lambda_\sharp \|\nabla_K \underline{v}_K\|^2_{L^2(K;\mathbb{R}^d)}$ for all $\underline{v}_K \in \underline{V}_K$. Furthermore, the following comparison result holds (cf. [20, Lemma 13.11, $p = 2$] and its proof): there exist $\alpha_\flat, \alpha_\sharp$ with $0 < \alpha_\flat \le \alpha_\sharp < \infty$, only depending on $\Omega$, $d$, and $\theta_\mathcal{D}$ such that $\alpha_\flat |\underline{v}_K|^2_{1,K} \le \|\nabla_K \underline{v}_K\|^2_{L^2(K;\mathbb{R}^d)} \le \alpha_\sharp |\underline{v}_K|^2_{1,K}$ for all $\underline{v}_K \in \underline{V}_K$. Combining both estimates, we infer a local coercivity and boundedness result:

$$\forall \underline{v}_K \in \underline{V}_K, \qquad \lambda_\flat \alpha_\flat |\underline{v}_K|^2_{1,K} \le a_K^\Lambda(\underline{v}_K, \underline{v}_K) \le \lambda_\sharp \alpha_\sharp |\underline{v}_K|^2_{1,K}. \tag{2.11}$$

Summing over $K \in \mathcal{M}$, we get the following global estimates:

$$\forall \underline{v}_\mathcal{D} \in \underline{V}_\mathcal{D}, \qquad \lambda_\flat \alpha_\flat |\underline{v}_\mathcal{D}|^2_{1,\mathcal{D}} \le a_\mathcal{D}^\Lambda(\underline{v}_\mathcal{D}, \underline{v}_\mathcal{D}) \le \lambda_\sharp \alpha_\sharp |\underline{v}_\mathcal{D}|^2_{1,\mathcal{D}}. \tag{2.12}$$

The well-posedness of Problem (2.5)-(2.6) follows.

**Proposition 1** (Well-posedness)**.** *There exists a unique solution $\underline{u}_\mathcal{D}^\infty \in \underline{V}_\mathcal{D}$ to Problem (2.5)-(2.6), which satisfies $|\underline{u}_\mathcal{D}^\infty|_{1,\mathcal{D}} \le C \left( \|f\|_{L^2(\Omega)} + \|g^N\|_{L^2(\Gamma^N)} + \|g^D\|_{H^{1/2}(\Gamma^D)} \right)$, for some $C > 0$ depending on the data, and on the discretisation $\mathcal{D}$ only through $\theta_\mathcal{D}$.*

*Proof.* The existence/uniqueness of $\underline{u}_\mathcal{D}^z \in \underline{V}_{\mathcal{D},0}$ solution to (2.5) (and in turn of $\underline{u}_\mathcal{D}^\infty = \underline{u}_\mathcal{D}^z + \underline{u}_\mathcal{D}^l$) is a direct consequence of the coercivity estimate (2.12), and of the fact that $|\cdot|_{1,\mathcal{D}}$ defines a norm on $\underline{V}_{\mathcal{D},0}$ (recall that $\underline{V}_{\mathcal{D},0}$ denotes either $\underline{V}_{\mathcal{D},0}^N$ when $|\Gamma^D| = 0$, or $\underline{V}_{\mathcal{D},0}^D$ otherwise). To prove the bound on $|\underline{u}_\mathcal{D}^\infty|_{1,\mathcal{D}}$, we use the triangle inequality:

$$|\underline{u}_\mathcal{D}^\infty|_{1,\mathcal{D}} \le |\underline{u}_\mathcal{D}^z|_{1,\mathcal{D}} + |\underline{u}_\mathcal{D}^l|_{1,\mathcal{D}}.$$

Since $|\underline{u}_\mathcal{D}^M|_{1,\mathcal{D}} = 0$, we only need to treat the case $|\Gamma^D| > 0$. The second term in the right-hand side is estimated recalling that $|\underline{u}_\mathcal{D}^D|_{1,\mathcal{D}} \le C_{l,\Gamma^D} \|g^D\|_{H^{1/2}(\Gamma^D)}$. To estimate the first term, we test Problem (2.5) with $\underline{v}_\mathcal{D} = \underline{u}_\mathcal{D}^z \in \underline{V}_{\mathcal{D},0}$, we use (2.12), and we apply the Cauchy–Schwarz inequality. We get

$$\lambda_\flat \alpha_\flat |\underline{u}_\mathcal{D}^z|^2_{1,\mathcal{D}} \le \|f\|_{L^2(\Omega)} \|u_\mathcal{M}^z\|_{L^2(\Omega)} + \|g^N\|_{L^2(\Gamma^N)} \|u_\mathcal{E}^z\|_{L^2(\Gamma^N)} + \lambda_\sharp \alpha_\sharp |\underline{u}_\mathcal{D}^l|_{1,\mathcal{D}} |\underline{u}_\mathcal{D}^z|_{1,\mathcal{D}}.$$

The conclusion follows from a discrete Poincaré inequality recalled in Proposition 7 and applied to $\underline{u}_\mathcal{D}^z \in \underline{V}_{\mathcal{D},0}$, from the discrete trace inequality of [20, Eq. (B.58), $p = 2$] combined with a discrete Poincaré inequality, and from the fact that $|\underline{u}_\mathcal{D}^D|_{1,\mathcal{D}} \le C_{l,\Gamma^D} \|g^D\|_{H^{1/2}(\Gamma^D)}$. □

## 3 Definition of the schemes and well-posedness

In this section, we introduce and study the well-posedness of three HFV schemes, two linear ones and a nonlinear scheme, for the time-dependent advection-diffusion problem (1.1). For the first two (linear) schemes, we introduce and study in the first place their steady versions on Problem (1.2). For the nonlinear scheme, by anticipation of the discrete asymptotic analysis of Section 4.3, we restrict our study to the case where the (positive) solution to Problem (1.1) converges in long time towards the so-called thermal equilibrium (see (3.11)). However, as it will be verified numerically in Section 5.4 in the stationary setting, our scheme is applicable for more general data. We consider a fixed spatial discretisation $\mathcal{D}$ of $\Omega$, which satisfies the conditions detailed in Section 2.1, and a fixed time step $\Delta t > 0$ for the time discretisation.

**Remark 1** (Linear schemes and nonhomogeneous data)**.** *The linearity of Problem (1.1) implies that, (i) if $|\Gamma^D| > 0$, the shifted variable $u^z := u - u^D$ (recall that $u^D$ is a known lifting of the Dirichlet datum $g^D$) satisfies an advection-diffusion equation with zero Dirichlet boundary condition on $\Gamma^D$, and (ii) otherwise, the shifted variable $u^z := u - u^M$ (recall that $u^M = \frac{M}{|\Omega|}$ is the mass lifting) satisfies a (compatible) pure Neumann advection-diffusion equation with zero-mass constraint. Thus, without loss of generality, we can restrict our study to the homogeneous case $g^D = 0$ or $M = 0$. For an example (in the steady, purely diffusive case) on how to handle at the discrete level nonhomogeneous data $g^D$ or $M$, we refer the reader to Problem (2.5) and Proposition 1 above. Notice that such manipulations are possible for linear schemes only.*

## 3.1 Standard HFV scheme

We present the HFV variant of the HMM family of schemes introduced in [4, 18].

### 3.1.1 Stationary problem

We consider Problem (1.2) with $g^D = 0$ when $|\Gamma^D| > 0$, or $M = \int_\Omega u^\infty = 0$ otherwise (cf. Remark 1). Locally to any cell $K \in \mathcal{M}$, we introduce the discrete bilinear form $a_K : \underline{V}_K \times \underline{V}_K \to \mathbb{R}$ such that, for all $\underline{u}_K, \underline{v}_K \in \underline{V}_K$,

$$a_K(\underline{u}_K, \underline{v}_K) := a_K^\Lambda(\underline{u}_K, \underline{v}_K) + a_K^\phi(\underline{u}_K, \underline{v}_K), \tag{3.1}$$

where the diffusive part $a_K^\Lambda$ is defined by (2.4) (and rewrites as (2.9) in terms of the local diffusive fluxes $F_{K,\sigma}^\Lambda(\underline{u}_K)$ given by (2.7)), and the advective part $a_K^\phi$ is defined by

$$a_K^\phi(\underline{u}_K, \underline{v}_K) := \sum_{\sigma \in \mathcal{E}_K} F_{K,\sigma}^\phi(\underline{u}_K)(v_K - v_\sigma), \tag{3.2}$$

with $F_{K,\sigma}^\phi(\underline{u}_K)$ an approximation of the normal advective flux $\int_\sigma u^\infty V_{|K}^\phi \cdot n_{K,\sigma}$. In order to define the numerical advective fluxes, we need to introduce some data. We set $V_{K,\sigma}^\phi := \frac{1}{|\sigma|}\int_\sigma V^\phi \cdot n_{K,\sigma}$, and $\mu_\sigma := \min\left(1, \min_{K\in\mathcal{M}_\sigma} \mathrm{Sp}(\Lambda_K)\right) > 0$ where $\Lambda_K := \frac{1}{|K|}\int_K \Lambda$. We could as well simply define $\mu_\sigma$ as $\min_{K\in\mathcal{M}_\sigma} \mathrm{Sp}(\Lambda_K)$, but we choose here to stick to the formula advocated in [4, 18]. We also consider a Lipschitz continuous function $A : \mathbb{R} \to \mathbb{R}$, satisfying the following conditions:

$$\begin{aligned} A(0) &= 0, \\ \forall s \in \mathbb{R},\ A(-s) - A(s) &= s, \\ \forall s \in \mathbb{R},\ A(-s) + A(s) &\geq 0. \end{aligned} \tag{3.3}$$

Standard choices of $A$ functions include:

- the centred discretisation: $A : s \mapsto -\frac{s}{2}$;
- the upwind discretisation: $A : s \mapsto \max(-s, 0)$;
- the Scharfetter–Gummel discretisation: $A : s \mapsto \begin{cases} \frac{s}{\mathrm{e}^s - 1} - 1 & \text{if } s \neq 0 \\ 0 & \text{if } s = 0 \end{cases}$.

We eventually define, for all $K \in \mathcal{M}$, and all $\sigma \in \mathcal{E}_K$, the numerical advective flux: for all $\underline{u}_K \in \underline{V}_K$,

$$F_{K,\sigma}^\phi(\underline{u}_K) := |\sigma| \frac{\mu_\sigma}{d_{K,\sigma}} \left( A\left(-\frac{d_{K,\sigma}}{\mu_\sigma} V_{K,\sigma}^\phi\right) u_K - A\left(\frac{d_{K,\sigma}}{\mu_\sigma} V_{K,\sigma}^\phi\right) u_\sigma \right). \tag{3.4}$$

Letting $a_\mathcal{D} : \underline{V}_\mathcal{D} \times \underline{V}_\mathcal{D} \to \mathbb{R}$ be the (global) discrete bilinear form such that, for all $\underline{u}_\mathcal{D}, \underline{v}_\mathcal{D} \in \underline{V}_\mathcal{D}$,

$$a_\mathcal{D}(\underline{u}_\mathcal{D}, \underline{v}_\mathcal{D}) := \sum_{K\in\mathcal{M}} a_K(\underline{u}_K, \underline{v}_K), \tag{3.5}$$

and recalling that $\underline{V}_{\mathcal{D},0}$ denotes either $\underline{V}^N_{\mathcal{D},0}$ whenever $|\Gamma^D| = 0$, or $\underline{V}^D_{\mathcal{D},0}$ otherwise, the discrete problem reads: Find $\underline{u}^\infty_\mathcal{D} \in \underline{V}_{\mathcal{D},0}$ such that

$$a_\mathcal{D}(\underline{u}^\infty_\mathcal{D}, \underline{v}_\mathcal{D}) = (f, v_\mathcal{M})_\Omega + (g^N, v_\mathcal{E})_{\Gamma^N} \qquad \forall \underline{v}_\mathcal{D} \in \underline{V}_{\mathcal{D},0}. \tag{3.6}$$

Remark in passing that, for pure Neumann boundary conditions with $M \neq 0$, as opposed to the purely diffusive case of Problem (2.5), $a_\mathcal{D}(\underline{u}^M_\mathcal{D}, \underline{v}_\mathcal{D}) \neq 0$ a priori for $\underline{v}_\mathcal{D} \in \underline{V}^N_{\mathcal{D},0}$. The well-posedness of (3.6) is discussed in the following proposition.

**Proposition 2** (Well-posedness)**.** *Let $A$ be a Lipschitz continuous function satisfying (3.3). If the advection field $V^\phi$ satisfies the two following conditions:*

$$(i)\ \text{almost everywhere on } \Gamma^N,\ V^\phi \cdot n \leq 0, \tag{3.7a}$$

$$(ii)\ \text{there exists } \beta < \frac{2\lambda_\flat \alpha_\flat}{C_P^2} \text{ such that, almost everywhere in } \Omega,\ \operatorname{div} V^\phi \geq -\beta, \tag{3.7b}$$

*where $\lambda_\flat \alpha_\flat$ is the coercivity constant of (2.12), and $C_P$ is either equal to $C_{PW}$ if $|\Gamma^D| = 0$ or to $C_{P,\Gamma^D}$ otherwise (where $C_{PW}, C_{P,\Gamma^D}$ are the Poincaré constants of Proposition 7), then there exists $\kappa > 0$, only depending on $\Lambda$, $\beta$, $\Omega$, $d$, $\Gamma^D$, and $\theta_\mathcal{D}$ such that*

$$\forall \underline{v}_\mathcal{D} \in \underline{V}_{\mathcal{D},0}, \qquad a_\mathcal{D}(\underline{v}_\mathcal{D}, \underline{v}_\mathcal{D}) \geq \kappa |\underline{v}_\mathcal{D}|^2_{1,\mathcal{D}}. \tag{3.8}$$

*Consequently, there exists a unique solution $\underline{u}^\infty_\mathcal{D} \in \underline{V}_{\mathcal{D},0}$ to Problem (3.6). Moreover, one has $|\underline{u}^\infty_\mathcal{D}|_{1,\mathcal{D}} \leq C\big(\|f\|_{L^2(\Omega)} + \|g^N\|_{L^2(\Gamma^N)}\big)$, for some $C > 0$ depending on the data, and on the discretisation $\mathcal{D}$ only through $\theta_\mathcal{D}$.*

*Proof.* Let $K \in \mathcal{M}$, and $\sigma \in \mathcal{E}_K$. Let $s_{K,\sigma} := \frac{d_{K,\sigma}}{\mu_\sigma} V^\phi_{K,\sigma}$ and $\zeta_{K,\sigma} := A(-s_{K,\sigma}) + A(s_{K,\sigma})$. According to (3.3), $\zeta_{K,\sigma} \geq 0$, and we have $A(-s_{K,\sigma}) = \frac{s_{K,\sigma}+\zeta_{K,\sigma}}{2}$ and $-A(s_{K,\sigma}) = \frac{s_{K,\sigma}-\zeta_{K,\sigma}}{2}$. Consequently, for all $\underline{v}_K \in \underline{V}_K$,

$$\big(A(-s_{K,\sigma})v_K - A(s_{K,\sigma})v_\sigma\big)(v_K - v_\sigma) = \frac{1}{2}s_{K,\sigma}(v_K^2 - v_\sigma^2) + \frac{1}{2}\zeta_{K,\sigma}(v_K - v_\sigma)^2 \geq \frac{1}{2}s_{K,\sigma}(v_K^2 - v_\sigma^2).$$

Recalling (3.2) and (3.4), we infer that, for all $\underline{v}_\mathcal{D} \in \underline{V}_\mathcal{D}$,

$$a^\phi_\mathcal{D}(\underline{v}_\mathcal{D}, \underline{v}_\mathcal{D}) := \sum_{K\in\mathcal{M}} a^\phi_K(\underline{v}_K, \underline{v}_K) \geq \frac{1}{2}\sum_{K\in\mathcal{M}}\sum_{\sigma\in\mathcal{E}_K} |\sigma| V^\phi_{K,\sigma}(v_K^2 - v_\sigma^2).$$

Since, for all $K \in \mathcal{M}$, $\sum_{\sigma\in\mathcal{E}_K} |\sigma| V^\phi_{K,\sigma} = \int_K \operatorname{div} V^\phi$ and, for all $\sigma \in \mathcal{E}_{int}$, $\sum_{K\in\mathcal{M}_\sigma} |\sigma| V^\phi_{K,\sigma} = 0$, we have, for all $\underline{v}_\mathcal{D} \in \underline{V}_{\mathcal{D},0}$,

$$a^\phi_\mathcal{D}(\underline{v}_\mathcal{D}, \underline{v}_\mathcal{D}) \geq \frac{1}{2}(\operatorname{div} V^\phi, v^2_\mathcal{M})_\Omega - \frac{1}{2}(V^\phi \cdot n, v^2_\mathcal{E})_{\Gamma^N}. \tag{3.9}$$

Combining (3.5)-(3.1), (3.9)-(3.7), and the coercivity result (2.12), we deduce that, for all $\underline{v}_\mathcal{D} \in \underline{V}_{\mathcal{D},0}$,

$$a_\mathcal{D}(\underline{v}_\mathcal{D}, \underline{v}_\mathcal{D}) \geq \lambda_\flat \alpha_\flat |\underline{v}_\mathcal{D}|^2_{1,\mathcal{D}} - \frac{\beta}{2}\|v_\mathcal{M}\|^2_{L^2(\Omega)}. \tag{3.10}$$

Using a Poincaré inequality from Proposition 7, one has, for all $\underline{v}_\mathcal{D} \in \underline{V}_{\mathcal{D},0}$,

$$a_\mathcal{D}(\underline{v}_\mathcal{D}, \underline{v}_\mathcal{D}) \geq \left(\lambda_\flat \alpha_\flat - \frac{\beta C_P^2}{2}\right) |\underline{v}_\mathcal{D}|^2_{1,\mathcal{D}},$$

therefore the estimate (3.8) holds for $\kappa = \lambda_\flat \alpha_\flat - \frac{\beta C_P^2}{2}$, which is positive according to (3.7b). The existence/uniqueness of $\underline{u}^\infty_\mathcal{D} \in \underline{V}_{\mathcal{D},0}$ solution to (3.6) is a direct consequence of the coercivity estimate (3.8), and of the fact that $|\cdot|_{1,\mathcal{D}}$ defines a norm on $\underline{V}_{\mathcal{D},0}$. The continuous dependency of $\underline{u}^\infty_\mathcal{D}$ with respect to the data can then be proved as in the proof of Proposition 1. □

**Remark 2** (Assumptions on the advection field)**.** *The well-posedness result of Proposition 2 do not use the fact that $V^\phi$ is related to the gradient of a potential, and are thus generalisable to general advection fields. Even better, under a smallness assumption on the meshsize $h_\mathcal{D}$, it is actually possible to prove well-posedness for Problem (3.6) without assumptions (3.7) on the advection field. The starting point to prove so is a discrete Gårding inequality like (3.10), which can be easily obtained in full generality from (3.9) (in the case $|\Gamma^N| > 0$, it is obtained from the multiplicative discrete trace inequality of [20, Eq. (B.57), $p = 2$] and holds for $h_\mathcal{D}$ sufficiently small). The proof then proceeds by contradiction, assuming that a discrete inf-sup condition does not hold in the limit $h_\mathcal{D} \to 0$, and using a compactness argument (cf. [20, Lemmas B.27-B.33, $p = 2$]).*

**Remark 3** (Choice of $A$)**.** *The choice of the function $A$ is of great importance. In particular, the Scharfetter–Gummel approximation is rather classical in various contexts. First introduced in [35] in the framework of TPFA schemes, this approximation of the flux ensures the preservation of the so-called thermal (or Gibbs) equilibrium at the discrete level which has the form:*

$$u_{th}^\infty = \bar\rho\, \mathrm{e}^{-\phi} \tag{3.11}$$

*where $\bar\rho \in \mathbb{R}_+$ is prescribed by the data. For instance, for pure Neumann boundary conditions, $f = 0$, and $g^N = 0$, we have $\bar\rho = M/\int_\Omega e^{-\phi}$. The discrete solution obtained with the TPFA scheme is then the interpolate of $u_{th}^\infty$. This property is no more true for the hybrid scheme and we observe numerically that, for $M \neq 0$, the discrete solution $\underline{u}_\mathcal{D}^\infty \in \underline{V}_\mathcal{D}$ is in general not the HFV interpolate of $u_{th}^\infty$. However, as explained in [18, pp. 553-554], provided the parameters $(\mu_\sigma)_{\sigma\in\mathcal{E}}$ are well-chosen, using the Scharfetter–Gummel $A$ function ensures a sort of automatic upwinding of the scheme.*

### 3.1.2 Evolutive problem

We consider Problem (1.1) with $g^D = 0$ when $|\Gamma^D| > 0$, or $M = \int_\Omega u^{in} = 0$ otherwise (see Remark 1). We use a backward Euler discretisation in time, and the HFV discretisation introduced in Section 3.1.1 in space. The discrete problem reads: Find $\left(\underline{u}_\mathcal{D}^n \in \underline{V}_{\mathcal{D},0}\right)_{n\geq 1}$ such that

$$\begin{cases} \dfrac{1}{\Delta t}(u_\mathcal{M}^n - u_\mathcal{M}^{n-1}, v_\mathcal{M})_\Omega + a_\mathcal{D}(\underline{u}_\mathcal{D}^n, \underline{v}_\mathcal{D}) = (f, v_\mathcal{M})_\Omega + (g^N, v_\mathcal{E})_{\Gamma^N} \qquad \forall \underline{v}_\mathcal{D} \in \underline{V}_{\mathcal{D},0}, & \text{(3.12a)} \\ u_K^0 = \dfrac{1}{|K|}\displaystyle\int_K u^{in} \qquad \forall K \in \mathcal{M}, & \text{(3.12b)} \end{cases}$$

where $a_\mathcal{D}$ is defined by (3.1)–(3.5). Since $a_\mathcal{D}$ is coercive, the bilinear form in (3.12) is also coercive, so the scheme (3.12) is well-posed, under the assumptions (3.7) on the advection field, as a straightforward consequence of Proposition 2.

**Remark 4** (Pure Neumann case)**.** *When considering pure Neumann boundary conditions, and contrary to the stationary case, one can actually seek at each time step for a solution to Problem (3.12) in $\underline{V}_\mathcal{D}$, i.e., it is not necessary to seek for a solution in the constrained space $\underline{V}_{\mathcal{D},0}^N$. Indeed, testing (3.12a) by $\underline{1}_\mathcal{D} \in \underline{V}_\mathcal{D}$, and using that $\int_\Omega f + \int_{\partial\Omega} g^N = 0$ and $a_\mathcal{D}(\underline{u}_\mathcal{D}^n, \underline{1}_\mathcal{D}) = 0$, one can automatically infer that $\int_\Omega u_\mathcal{M}^n = \int_\Omega u_\mathcal{M}^{n-1}$ for all $n \geq 1$, that is, $\int_\Omega u_\mathcal{M}^n = \int_\Omega u_\mathcal{M}^0 = \int_\Omega u^{in} = M = 0$ for all $n \geq 1$, i.e., $\underline{u}_\mathcal{D}^n \in \underline{V}_{\mathcal{D},0}^N$ for all $n \in \mathbb{N}^\star$.*

## 3.2 Exponential fitting HFV scheme

Following ideas in [6] (cf. also [30, 32] and [26] for general advection fields) in the context of finite element methods, we aim to design an unconditionally (i.e., without the need for assumptions (3.7) on the advection field $V^\phi$) coercive scheme for the advection-diffusion problem in the HFV framework.

### 3.2.1 Stationary problem

We consider Problem (1.2). The strategy advocated in [6] is based on the following observation: at the continuous level, if $u^\infty$ is a solution to (1.2), letting

$$\omega = \mathrm{e}^{-\phi}, \tag{3.13}$$

we can introduce the Slotboom change of variable (see [33, 34]): $\rho^\infty = \frac{u^\infty}{\omega}$. Indeed, noticing that $\nabla u^\infty + u^\infty \nabla\phi = \omega\nabla\rho^\infty - \rho^\infty\omega\nabla\phi + \rho^\infty\omega\nabla\phi = \omega\nabla\rho^\infty$, the Slotboom variable $\rho^\infty$ equivalently solves the following pure diffusion problem:

$$\begin{cases} -\operatorname{div}(\omega\Lambda\nabla\rho^\infty) = f & \text{in } \Omega, \\ \rho^\infty = \omega^{-1}_{|\Gamma^D} g^D & \text{on } \Gamma^D, \\ \omega\Lambda\nabla\rho^\infty\cdot n = g^N & \text{on } \Gamma^N, \end{cases} \tag{3.14}$$

with additional constraint $\int_\Omega \omega\rho^\infty = M$ when $|\Gamma^D| = 0$. Following Remark 1 (with $\rho$ instead of $u$, $\rho^D$ lifting of $\omega^{-1}_{|\Gamma^D} g^D$ instead of $u^D$ lifting of $g^D$, and $\rho^M := \frac{M}{\int_\Omega \omega} \in \mathbb{R}$ instead of $u^M$), we consider Problem (3.14) with $\rho^\infty_{|\Gamma^D} = 0$ when $|\Gamma^D| > 0$, or $\int_\Omega \omega\rho^\infty = 0$ otherwise (which is equivalent to consider Problem (1.2) with $g^D = 0$ or $M = \int_\Omega u^\infty = 0$). Since $\phi$ is continuous on $\overline{\Omega}$, there exist $\omega_\flat, \omega_\sharp$ with $0 < \omega_\flat \le \omega_\sharp < \infty$, only depending on $\phi$ and $\Omega$, such that $\omega_\flat \le \omega(x) \le \omega_\sharp$ for all $x \in \overline{\Omega}$. We then denote by $L^2_\omega(\Omega)$ the $\omega$-weighted $L^2$ space on $\Omega$.

At the discrete level, instead of discretising (1.2), we approximate the solution to (3.14). For any $K \in \mathcal{M}$, we let $a^\omega_K : \underline{V}_K \times \underline{V}_K \to \mathbb{R}$ be the discrete bilinear form such that, for all $\underline{\rho}_K, \underline{v}_K \in \underline{V}_K$,

$$a^\omega_K(\underline{\rho}_K, \underline{v}_K) := (\omega\Lambda\nabla_K\underline{\rho}_K, \nabla_K\underline{v}_K)_K, \tag{3.15}$$

and, classically, we let $a^\omega_\mathcal{D} : \underline{V}_\mathcal{D} \times \underline{V}_\mathcal{D} \to \mathbb{R}$ be the corresponding global discrete bilinear form obtained by sum of the local contributions. To account for the change of variable, we let $\underline{V}^\omega_{\mathcal{D},0}$ be the space $\underline{V}^D_{\mathcal{D},0}$ when $|\Gamma^D| > 0$, and the space $\{\underline{v}_\mathcal{D} \in \underline{V}_\mathcal{D} : \int_\Omega \omega v_\mathcal{M} = 0\}$ otherwise. The discrete problem reads: Find $\underline{\rho}^\infty_\mathcal{D} \in \underline{V}^\omega_{\mathcal{D},0}$ such that

$$a^\omega_\mathcal{D}(\underline{\rho}^\infty_\mathcal{D}, \underline{v}_\mathcal{D}) = (f, v_\mathcal{M})_\Omega + (g^N, v_\mathcal{E})_{\Gamma^N} \qquad \forall \underline{v}_\mathcal{D} \in \underline{V}^\omega_{\mathcal{D},0}. \tag{3.16}$$

Remark that, for pure Neumann boundary conditions with $M \neq 0$, as for Problem (2.5), letting $\underline{\rho}^M_\mathcal{D} := \rho^M \underline{1}_\mathcal{D}$, $a^\omega_\mathcal{D}(\underline{\rho}^M_\mathcal{D}, \underline{v}_\mathcal{D}) = 0$ for all $\underline{v}_\mathcal{D} \in \underline{V}_\mathcal{D}$. Letting $\underline{\omega}_\mathcal{D} \in \underline{V}_\mathcal{D}$ be the HFV interpolate of $\omega$, i.e.,

$$\omega_K := \frac{1}{|K|}\int_K \omega \quad \forall K \in \mathcal{M}, \qquad \omega_\sigma := \frac{1}{|\sigma|}\int_\sigma \omega \quad \forall \sigma \in \mathcal{E}, \tag{3.17}$$

the approximation of the solution to Problem (1.2) is finally defined as the product $\underline{u}^\infty_\mathcal{D} := \underline{\omega}_\mathcal{D} \times \underline{\rho}^\infty_\mathcal{D}$, that is

$$u^\infty_K = \omega_K \rho^\infty_K \quad \forall K \in \mathcal{M}, \qquad u^\infty_\sigma = \omega_\sigma \rho^\infty_\sigma \quad \forall \sigma \in \mathcal{E}. \tag{3.18}$$

Remark that $\underline{u}^\infty_\mathcal{D} \in \underline{V}_{\mathcal{D},0}$. Reasoning as in Section 2.3, one can easily prove that, for all $\underline{v}_\mathcal{D} \in \underline{V}_\mathcal{D}$,

$$a^\omega_\mathcal{D}(\underline{v}_\mathcal{D}, \underline{v}_\mathcal{D}) \ge \omega_\flat\lambda_\flat\alpha_\flat\, |\underline{v}_\mathcal{D}|^2_{1,\mathcal{D}}. \tag{3.19}$$

This estimate is instrumental to infer well-posedness for Problem (3.16)-(3.18).

**Proposition 3** (Well-posedness). *There exists a unique solution $\underline{u}^\infty_\mathcal{D} \in \underline{V}_{\mathcal{D},0}$ to Problem (3.16)-(3.18), which satisfies $|\underline{u}^\infty_\mathcal{D}|_{1,\mathcal{D}} \le C\left(\|f\|_{L^2(\Omega)} + \|g^N\|_{L^2(\Gamma^N)}\right)$, for some $C > 0$ depending on the data, and on the discretisation $\mathcal{D}$ only through $\theta_\mathcal{D}$.*

*Proof.* The existence/uniqueness of $\underline{\rho}_{\mathcal{D}}^{\infty} \in \underline{V}_{\mathcal{D},0}^{\omega}$ solution to (3.16) (and in turn of $\underline{u}_{\mathcal{D}}^{\infty} \in \underline{V}_{\mathcal{D},0}$) is a direct consequence of the coercivity estimate (3.19), and of the fact that $|\cdot|_{1,\mathcal{D}}$ clearly defines a norm on $\underline{V}_{\mathcal{D},0}^{\omega}$. To prove the bound on $|\underline{u}_{\mathcal{D}}^{\infty}|_{1,\mathcal{D}}$, we use the fact that $\underline{u}_{\mathcal{D}}^{\infty} = \underline{\omega}_{\mathcal{D}} \times \underline{\rho}_{\mathcal{D}}^{\infty}$. For all $K \in \mathcal{M}$,

$$|\underline{u}_K^{\infty}|_{1,K}^2 = \sum_{\sigma \in \mathcal{E}_K} \frac{|\sigma|}{d_{K,\sigma}} (\omega_K \rho_K^{\infty} - \omega_\sigma \rho_\sigma^{\infty})^2 \le 2 \sum_{\sigma \in \mathcal{E}_K} \frac{|\sigma|}{d_{K,\sigma}} \left( \omega_\sigma^2 (\rho_K^{\infty} - \rho_\sigma^{\infty})^2 + (\rho_K^{\infty})^2 (\omega_K - \omega_\sigma)^2 \right).$$

By definition (3.17) of $\underline{\omega}_{\mathcal{D}}$, and local stability of the HFV interpolant (cf. [20, Proposition B.7, $p = 2$] and its proof), we infer

$$|\underline{u}_K^{\infty}|_{1,K}^2 \le 2\, \omega_\sharp^2\, |\underline{\rho}_K^{\infty}|_{1,K}^2 + 2\, (\rho_K^{\infty})^2 C_{sta}^2 \|\nabla \omega\|_{L^2(K;\mathbb{R}^d)}^2,$$

with $C_{sta} > 0$ only depending on $d$ and $\theta_{\mathcal{D}}$. Since $\nabla \omega = -\omega \nabla \phi$, $\phi \in C^1(\overline{\Omega})$, and $\omega > 0$, we have

$$\|\nabla \omega\|_{L^2(K;\mathbb{R}^d)}^2 \le \omega_\sharp \sup_{x \in \overline{\Omega}} |\nabla \phi(x)|^2 |K|\, \omega_K.$$

Summing over $K \in \mathcal{M}$ then yields

$$|\underline{u}_{\mathcal{D}}^{\infty}|_{1,\mathcal{D}}^2 \le 2\, \omega_\sharp^2\, |\underline{\rho}_{\mathcal{D}}^{\infty}|_{1,\mathcal{D}}^2 + 2\, C_{sta}^2\, \omega_\sharp \sup_{x \in \overline{\Omega}} |\nabla \phi(x)|^2 \|\rho_{\mathcal{M}}^{\infty}\|_{L^2_\omega(\Omega)}^2.$$

When $|\Gamma^D| > 0$, $\|\rho_{\mathcal{M}}^{\infty}\|_{L^2_\omega(\Omega)}^2 \le \omega_\sharp \|\rho_{\mathcal{M}}^{\infty}\|_{L^2(\Omega)}^2 \le \omega_\sharp\, C_{P,\Gamma^D}^2 |\underline{\rho}_{\mathcal{D}}^{\infty}|_{1,\mathcal{D}}^2$, where we have applied the discrete Poincaré inequality (A.2) to $\underline{\rho}_{\mathcal{D}}^{\infty} \in \underline{V}_{\mathcal{D},0}^{D}$. Otherwise, $\underline{\rho}_{\mathcal{D}}^{\infty} \in \underline{V}_{\mathcal{D},0}^{\omega}$ satisfies $\int_\Omega \omega \rho_{\mathcal{M}}^{\infty} = 0$, and one can use [8, Lemma 5.2] to infer that $\|\rho_{\mathcal{M}}^{\infty}\|_{L^2_\omega(\Omega)} \le 2 \|\rho_{\mathcal{M}}^{\infty} - \frac{1}{|\Omega|} \int_\Omega \rho_{\mathcal{M}}^{\infty}\|_{L^2_\omega(\Omega)}$, and finally get that $\|\rho_{\mathcal{M}}^{\infty}\|_{L^2_\omega(\Omega)} \le 2 \sqrt{\omega_\sharp}\, C_{PW} |\underline{\rho}_{\mathcal{D}}^{\infty}|_{1,\mathcal{D}}$ applying the discrete Poincaré inequality (A.1) to $\underline{\rho}_{\mathcal{D}}^{\infty} - \frac{1}{|\Omega|} \int_\Omega \rho_{\mathcal{M}}^{\infty} \underline{1}_{\mathcal{D}} \in \underline{V}_{\mathcal{D},0}^{N}$. In any case, we end up bounding $|\underline{u}_{\mathcal{D}}^{\infty}|_{1,\mathcal{D}}$ by $|\underline{\rho}_{\mathcal{D}}^{\infty}|_{1,\mathcal{D}}$, with multiplicative constant depending on the data, and on the discretisation $\mathcal{D}$ only through $\theta_{\mathcal{D}}$. The rest of the proof consists in bounding $|\underline{\rho}_{\mathcal{D}}^{\infty}|_{1,\mathcal{D}}$, and proceeds as in the proof of Proposition 1, using that $\|\rho_{\mathcal{M}}^{\infty}\|_{L^2(\Omega)} \le 2 \sqrt{\frac{\omega_\sharp}{\omega_\flat}}\, C_{PW} |\underline{\rho}_{\mathcal{D}}^{\infty}|_{1,\mathcal{D}}$ when $|\Gamma^D| = 0$. □

In the sequel, the HFV scheme (3.16)-(3.18) will be referred to as "exponential fitting". Notice that no assumption on $V^\phi$ is needed to ensure its well-posedness since this scheme is unconditionally coercive.

**Remark 5** (Preservation of the thermal equilibrium)**.** *As for the exponential fitting scheme of [6], the method* (3.16)-(3.18) *preserves the thermal equilibrium. This property is analogous to what holds true for the TPFA Scharfetter–Gummel scheme.*

### 3.2.2 Evolutive problem

We consider Problem (1.1). Following the previous strategy, letting $\rho := \frac{u}{\omega}$, one can show that $\rho$ equivalently solves the following transient pure diffusion problem:

$$\begin{cases} \omega \partial_t \rho - \operatorname{div}(\omega \Lambda \nabla \rho) = f & \text{in } \mathbb{R}_+ \times \Omega, \\ \rho = \omega_{|\Gamma^D}^{-1} g^D & \text{on } \mathbb{R}_+ \times \Gamma^D, \\ \omega \Lambda \nabla \rho \cdot n = g^N & \text{on } \mathbb{R}_+ \times \Gamma^N, \\ \rho(0,\cdot) = \rho^{in} & \text{in } \Omega, \end{cases} \tag{3.20}$$

where $\rho^{in} := \frac{u^{in}}{\omega}$. Following Remark 1, we consider Problem (3.20) with $\rho_{|\Gamma^D} = 0$ when $|\Gamma^D| > 0$, or $\int_\Omega \omega \rho^{in} = 0$ otherwise (which is equivalent to consider Problem (1.1) with $g^D = 0$ or $M = \int_\Omega u^{in} = 0$).

At the discrete level, instead of discretising (1.1), we approximate the solution to (3.20). We use a backward Euler discretisation in time, and the HFV discretisation introduced in Section 3.2.1 in space. The discrete problem reads: Find $\left(\underline{\rho}_{\mathcal{D}}^n \in \underline{V}_{\mathcal{D},0}^{\omega}\right)_{n\geq 1}$ such that

$$\begin{cases} \dfrac{1}{\Delta t}(\rho_{\mathcal{M}}^n - \rho_{\mathcal{M}}^{n-1}, \omega_{\mathcal{M}} v_{\mathcal{M}})_\Omega + a_{\mathcal{D}}^{\omega}(\underline{\rho}_{\mathcal{D}}^n, \underline{v}_{\mathcal{D}}) = (f, v_{\mathcal{M}})_\Omega + (g^N, v_{\mathcal{E}})_{\Gamma^N} \qquad \forall \underline{v}_{\mathcal{D}} \in \underline{V}_{\mathcal{D},0}^{\omega}, & \text{(3.21a)} \\ \rho_K^0 = \dfrac{1}{\omega_K |K|} \displaystyle\int_K u^{in} \qquad \forall K \in \mathcal{M}, & \text{(3.21b)} \end{cases}$$

where $a_{\mathcal{D}}^{\omega}$ is defined (locally) by (3.15), and the approximation of the solution $\underline{u}_{\mathcal{D}}^n \in \underline{V}_{\mathcal{D},0}$ to Problem (1.1) is finally defined as $\underline{u}_{\mathcal{D}}^n := \underline{\omega}_{\mathcal{D}} \times \underline{\rho}_{\mathcal{D}}^n$, i.e., according to (3.18) (with superscript $n$ instead of $\infty$). Once again, because of the coercivity of $a_{\mathcal{D}}^{\omega}$, the scheme (3.21)-(3.18) is unconditionally well-posed, as a straightforward consequence of Proposition 3.

**Remark 6** (Pure Neumann case)**.** *When considering pure Neumann boundary conditions, and contrary to the stationary case, one can actually seek at each time step for a solution to Problem (3.21) in $\underline{V}_{\mathcal{D}}$, i.e., it is not necessary to seek for a solution in the constrained space $\left\{\underline{v}_{\mathcal{D}} \in \underline{V}_{\mathcal{D}} : \int_\Omega \omega v_{\mathcal{M}} = 0\right\}$. Indeed, testing (3.21a) by $\underline{1}_{\mathcal{D}} \in \underline{V}_{\mathcal{D}}$, and using that $\int_\Omega f + \int_{\partial\Omega} g^N = 0$ and $a_{\mathcal{D}}^{\omega}(\underline{\rho}_{\mathcal{D}}^n, \underline{1}_{\mathcal{D}}) = 0$, one can automatically infer that $\int_\Omega \omega \rho_{\mathcal{M}}^n = \int_\Omega \omega \rho_{\mathcal{M}}^{n-1}$ for all $n \geq 1$, that is, $\int_\Omega \omega \rho_{\mathcal{M}}^n = \int_\Omega \omega \rho_{\mathcal{M}}^0 = \int_\Omega u^{in} = M = 0$ for all $n \geq 1$, i.e., $\underline{\rho}_{\mathcal{D}}^n \in \left\{\underline{v}_{\mathcal{D}} \in \underline{V}_{\mathcal{D}} : \int_\Omega \omega v_{\mathcal{M}} = 0\right\}$ (equivalently, $\underline{u}_{\mathcal{D}}^n \in \underline{V}_{\mathcal{D},0}^N$) for all $n \in \mathbb{N}^\star$.*

### 3.3 Nonlinear HFV scheme

We are interested in the evolutive problem (1.1). We restrict our study to the pure Neumann case ($|\Gamma^D| = 0$), and to the choice of data $f = 0$, $g^N = 0$, and $u^{in} \geq 0$ with $M = \int_\Omega u^{in} > 0$. Under these assumptions, it is known that the solution $u$ to Problem (1.1) is positive on $\mathbb{R}_+^\star \times \Omega$. Furthermore, in long time, $u(t)$ converges towards the thermal equilibrium. Indeed, we easily verify that the function $u^\infty = \rho^M \, \mathrm{e}^{-\phi}$, where we recall that $\rho^M = \frac{M}{\int_\Omega \mathrm{e}^{-\phi}} > 0$ (cf. Section 3.2.1), solves the steady problem (1.2) with same data. Since $u > 0$ on $\mathbb{R}_+^\star \times \Omega$, we can rewrite the flux $J = -\Lambda(\nabla u + u \nabla \phi)$ under the nonlinear form

$$J = -u\Lambda\nabla\big(\log(u) + \phi\big) = -u\Lambda\nabla \log\left(\frac{u}{u^\infty}\right).$$

At the continuous level, introducing this nonlinearity enables to highlight the following entropy/dissipation structure of the model at hand: testing the equation against $\log\left(\frac{u}{u^\infty}\right)$, we get

$$\frac{\mathrm{d}}{\mathrm{d}t}\mathbb{E}(t) + \mathbb{D}(t) = 0, \tag{3.22}$$

where

$$\mathbb{E}(t) := \int_\Omega u^\infty \, \Phi_1\left(\frac{u(t)}{u^\infty}\right) \qquad \text{and} \qquad \mathbb{D}(t) := \int_\Omega u(t) \, \Lambda\nabla \log\left(\frac{u(t)}{u^\infty}\right) \cdot \nabla \log\left(\frac{u(t)}{u^\infty}\right), \tag{3.23}$$

with $\Phi_1(s) = s\log(s) - s + 1$ for all $s > 0$. Since $\Phi_1 \geq 0$, the relative entropy $\mathbb{E}(t)$ is a non-negative quantity (as well as the relative dissipation $\mathbb{D}(t)$). The entropy/dissipation structure (3.22)-(3.23) is instrumental to prove the exponential convergence in time of the solution $u(t)$ to Problem (1.1) towards the equilibrium $u^\infty$. From the above nonlinear expression of the flux $J$, we build a nonlinear hybrid discretisation of the problem, leading to a scheme designed along the same principles as the nonlinear VAG and DDFV schemes of [10] and [9, 8]. This scheme is devised so as to ensure the positivity of discrete solutions, as well as to preserve at the discrete level the entropy/dissipation structure (and the long-time behaviour) of the model. The choice of designing a nonlinear HFV scheme is driven by the prospect of the design of Hybrid High Order schemes which could have similar features.

### 3.3.1 Definition of the scheme and key properties of discrete solutions

In the sequel, a vector of discrete unknowns $\underline{v}_{\mathcal{D}} \in \underline{V}_{\mathcal{D}}$ will be called positive if and only if, for all $K \in \mathcal{M}$ and all $\sigma \in \mathcal{E}$, $v_K > 0$ and $v_\sigma > 0$. Recall the definition (3.13) of $\omega = \mathrm{e}^{-\phi}$, as well as the definition (3.17) of the HFV interpolate $\underline{\omega}_{\mathcal{D}} \in \underline{V}_{\mathcal{D}}$ of $\omega$. Remark that $\underline{\omega}_{\mathcal{D}}$ is positive. If $\underline{u}_{\mathcal{D}} \in \underline{V}_{\mathcal{D}}$ is positive, one can then define $\underline{w}_{\mathcal{D}}$ as the element of $\underline{V}_{\mathcal{D}}$ such that

$$w_K := \log\left(\frac{u_K}{\omega_K}\right) \quad \forall K \in \mathcal{M}, \qquad w_\sigma := \log\left(\frac{u_\sigma}{\omega_\sigma}\right) \quad \forall \sigma \in \mathcal{E}. \tag{3.24}$$

In what follows, to emphasise the dependency of $\underline{w}_{\mathcal{D}}$ upon $\underline{u}_{\mathcal{D}}$, we sometimes write $\underline{w}_{\mathcal{D}}(\underline{u}_{\mathcal{D}})$. The definition (3.24) of $\underline{w}_{\mathcal{D}}$ will be justified a posteriori by (3.33). Locally to any cell $K \in \mathcal{M}$, we define an approximation of

$$(u, v) \mapsto -\int_K J \cdot \nabla v = \int_K u \, \Lambda \nabla \log\left(\frac{u}{u^\infty}\right) \cdot \nabla v$$

under the form

$$T_K(\underline{u}_K, \underline{w}_K, \underline{v}_K) := \int_K r_K(\underline{u}_K) \, \Lambda \nabla_K \underline{w}_K \cdot \nabla_K \underline{v}_K,$$

for all $\underline{u}_K \in \underline{V}_K$ positive and all $\underline{v}_K \in \underline{V}_K$, where $r_K : (\underline{V}_K)_+^\star \to \mathbb{R}_+^\star$ is a local reconstruction operator. Since $r_K(\underline{u}_K)$ is a (positive) constant on $K$, we have

$$T_K(\underline{u}_K, \underline{w}_K, \underline{v}_K) = r_K(\underline{u}_K) a_K^\Lambda(\underline{w}_K, \underline{v}_K), \tag{3.25}$$

where $a_K^\Lambda$ is defined by (2.4). Following (2.10), one can equivalently reformulate (3.25) using the local matrix $\mathbb{A}_K$ defined by (2.8):

$$T_K(\underline{u}_K, \underline{w}_K, \underline{v}_K) = r_K(\underline{u}_K) \, \delta_K \underline{v}_K \cdot \mathbb{A}_K \delta_K \underline{w}_K. \tag{3.26}$$

As already pointed out in the analysis of the nonlinear DDFV scheme of [9, 8], the definition of the local reconstruction operator is crucial to guarantee the existence of solutions and a good long-time behaviour to the scheme. The most natural choice in the HFV context would obviously be $r_K(\underline{u}_K) = u_K$, however it turns out that such a reconstruction embeds too few information on $\underline{u}_K$ to conclude. Therefore, we use a richer reconstruction, described below, which embeds information from both the local cell and face unknowns. For $\underline{u}_K \in \underline{V}_K$ positive, we let

$$r_K(\underline{u}_K) := f_{|\mathcal{E}_K|}\left(\left(m(u_K, u_\sigma)\right)_{\sigma \in \mathcal{E}_K}\right), \tag{3.27}$$

with $m : (\mathbb{R}_+^\star)^2 \to \mathbb{R}_+^\star$ and, for $k \geq 1$ integer, $f_k : (\mathbb{R}_+^\star)^k \to \mathbb{R}_+^\star$, such that

$$m \text{ is non-decreasing with respect to both its variables,} \tag{3.28a}$$
$$m(x, x) = x \text{ for all } x \in \mathbb{R}_+^\star \text{ and } m(y, x) = m(x, y) \text{ for all } (x, y) \in (\mathbb{R}_+^\star)^2, \tag{3.28b}$$
$$m(\lambda x, \lambda y) = \lambda m(x, y) \text{ for all } \lambda > 0 \text{ and all } (x, y) \in (\mathbb{R}_+^\star)^2, \tag{3.28c}$$
$$\frac{y - x}{\log(y) - \log(x)} \leq m(x, y) \leq \max(x, y) \text{ for all } (x, y) \in (\mathbb{R}_+^\star)^2, \, x \neq y, \tag{3.28d}$$

and

$$f_k(x_1, \ldots, x_k) = \frac{1}{k} \sum_{i=1}^k x_i \qquad \text{or} \qquad f_k(x_1, \ldots, x_k) = \max(x_1, \ldots, x_k). \tag{3.29}$$

Heuristically, $r_K(\underline{u}_K)$ computes an average of the unknowns attached to the cell $K$, especially it contains information about all the local face unknowns. As far as the properties (3.28)-(3.29) are

concerned, they will be instrumental to prove Lemma 2 and Proposition 6 below. As now standard, we finally let $T_{\mathcal{D}}$ be such that, for all $\underline{u}_{\mathcal{D}} \in \underline{V}_{\mathcal{D}}$ positive, and all $\underline{v}_{\mathcal{D}} \in \underline{V}_{\mathcal{D}}$,

$$T_{\mathcal{D}}(\underline{u}_{\mathcal{D}}, \underline{w}_{\mathcal{D}}, \underline{v}_{\mathcal{D}}) := \sum_{K \in \mathcal{M}} T_K(\underline{u}_K, \underline{w}_K, \underline{v}_K), \tag{3.30}$$

where the local contributions $T_K$ are defined by (3.25).

Using a backward Euler discretisation in time, and the HFV discretisation we have just introduced in space, our discrete problem reads: Find $\left(\underline{u}_{\mathcal{D}}^n \in \underline{V}_{\mathcal{D}}\right)_{n \geq 1}$ such that

$$\begin{cases} \dfrac{1}{\Delta t}(u_{\mathcal{M}}^n - u_{\mathcal{M}}^{n-1}, v_{\mathcal{M}})_\Omega + T_{\mathcal{D}}(\underline{u}_{\mathcal{D}}^n, \underline{w}_{\mathcal{D}}(\underline{u}_{\mathcal{D}}^n), \underline{v}_{\mathcal{D}}) = 0 & \forall \underline{v}_{\mathcal{D}} \in \underline{V}_{\mathcal{D}}, \quad (3.31a) \\ u_K^0 = \dfrac{1}{|K|}\displaystyle\int_K u^{in} \qquad \forall K \in \mathcal{M}. & (3.31b) \end{cases}$$

Notice that if $(\underline{u}_{\mathcal{D}}^n)_{n \geq 1}$ solves Problem (3.31), then, necessarily, $\underline{u}_{\mathcal{D}}^n$ is positive for all $n \geq 1$. Therefore, in the sequel, we will speak about the positive solutions to (3.31). Notice also that $u_{\mathcal{M}}^0$ may vanish in some cells of the mesh, since we only impose that $u^{in} \geq 0$ (but $u_{\mathcal{M}}^0$ cannot be identically zero in $\Omega$ since $M > 0$). Notice finally that $u_{\mathcal{E}}^0$ needs not be defined, as the scheme only uses $u_{\mathcal{M}}^0$.

Testing (3.31a) with $\underline{v}_{\mathcal{D}} = \underline{1}_{\mathcal{D}}$, and remarking that $T_{\mathcal{D}}(\underline{u}_{\mathcal{D}}^n, \underline{w}_{\mathcal{D}}(\underline{u}_{\mathcal{D}}^n), \underline{1}_{\mathcal{D}}) = 0$ for all $n \geq 1$, we immediately infer the following discrete mass conservation property.

**Proposition 4** (Mass conservation)**.** *If $\left(\underline{u}_{\mathcal{D}}^n \in \underline{V}_{\mathcal{D}}\right)_{n \geq 1}$ is a (positive) solution to (3.31), then*

$$\forall n \in \mathbb{N}^\star, \qquad \int_\Omega u_{\mathcal{M}}^n = \int_\Omega u_{\mathcal{M}}^0 = \int_\Omega u^{in} = M.$$

Following Proposition 4, a discrete steady-state $\underline{u}_{\mathcal{D}}^\infty \in \underline{V}_{\mathcal{D}}$ of (3.31) shall satisfy

$$T_{\mathcal{D}}(\underline{u}_{\mathcal{D}}^\infty, \underline{w}_{\mathcal{D}}(\underline{u}_{\mathcal{D}}^\infty), \underline{v}_{\mathcal{D}}) = 0 \qquad \forall \underline{v}_{\mathcal{D}} \in \underline{V}_{\mathcal{D}}, \tag{3.32}$$

and $\int_\Omega u_{\mathcal{M}}^\infty = M$. Letting $\underline{w}_{\mathcal{D}}^\infty := \underline{w}_{\mathcal{D}}(\underline{u}_{\mathcal{D}}^\infty)$, and testing (3.32) with $\underline{v}_{\mathcal{D}} = \underline{w}_{\mathcal{D}}^\infty$, by (3.30)-(3.25), since $r_K(\underline{u}_K^\infty) > 0$ for all $K \in \mathcal{M}$, we necessarily have $a_{\mathcal{D}}^\Lambda(\underline{w}_{\mathcal{D}}^\infty, \underline{w}_{\mathcal{D}}^\infty) = 0$, which yields $|\underline{w}_{\mathcal{D}}^\infty|_{1,\mathcal{D}} = 0$ by the coercivity property (2.12). Hence, $\underline{w}_{\mathcal{D}}^\infty = c\,\underline{1}_{\mathcal{D}}$ for some constant $c \in \mathbb{R}$, and since $\int_\Omega \rho^M \omega_{\mathcal{M}} = M = \int_\Omega u_{\mathcal{M}}^\infty$, by (3.24), we necessarily have $\mathrm{e}^c = 1$, that is $c = 0$ and $\underline{w}_{\mathcal{D}}^\infty = \underline{0}$. As a consequence, again by (3.24), $\underline{u}_{\mathcal{D}}^\infty = \rho^M \underline{\omega}_{\mathcal{D}}$, i.e., $\underline{u}_{\mathcal{D}}^\infty$ is the HFV interpolate of $u_{th}^\infty$. Thus, just like the exponential fitting scheme (cf. Remark 5), the nonlinear scheme preserves the thermal equilibrium. An important consequence of the fact that $\underline{u}_{\mathcal{D}}^\infty = \rho^M \underline{\omega}_{\mathcal{D}}$. Moreover, we notice that $\underline{w}_{\mathcal{D}}$ defined first by (3.24) can be modified up to an additive constant without any impact in the scheme (3.31), so that we can redefine $\underline{w}_{\mathcal{D}}$ as

$$w_K = \log\left(\frac{u_K}{u_K^\infty}\right) \quad \forall K \in \mathcal{M}, \qquad w_\sigma = \log\left(\frac{u_\sigma}{u_\sigma^\infty}\right) \quad \forall \sigma \in \mathcal{E}. \tag{3.33}$$

Another important consequence of the fact that $\underline{u}_{\mathcal{D}}^\infty$ is the HFV interpolate of $u_{th}^\infty$ is the following. Letting $u_\flat^\infty := \frac{M\omega_\flat}{|\Omega|\omega_\sharp} > 0$ and $u_\sharp^\infty := \frac{M\omega_\sharp}{|\Omega|\omega_\flat} > 0$ (recall that $\omega_\flat$ and $\omega_\sharp$ only depend on $\phi$ and $\Omega$), we have $u_\flat^\infty \leq u_{th}^\infty \leq u_\sharp^\infty$, but we also have that

$$u_\flat^\infty \underline{1}_{\mathcal{D}} \leq \underline{u}_{\mathcal{D}}^\infty \leq u_\sharp^\infty \underline{1}_{\mathcal{D}}, \tag{3.34}$$

where the inequalities shall be understood coordinate-wise. In other words, the continuous bounds on the steady-state are transferred to the discrete level.

Given a (positive) solution $\left(\underline{u}_{\mathcal{D}}^n \in \underline{V}_{\mathcal{D}}\right)_{n\geq 1}$ to (3.31), we define the following discrete versions of the relative entropy and dissipation introduced in (3.23): for all $n \geq 1$,

$$\mathbb{E}^n := \int_\Omega u_{\mathcal{M}}^\infty \Phi_1\left(\frac{u_{\mathcal{M}}^n}{u_{\mathcal{M}}^\infty}\right) \qquad \text{and} \qquad \mathbb{D}^n := T_{\mathcal{D}}\big(\underline{u}_{\mathcal{D}}^n, \underline{w}_{\mathcal{D}}^n, \underline{w}_{\mathcal{D}}^n\big), \tag{3.35}$$

where we let $w_{\mathcal{D}}^n := \underline{w}_{\mathcal{D}}(\underline{u}_{\mathcal{D}}^n)$, and we recall that $\Phi_1(s) = s\log(s) - s + 1$ for all $s > 0$. Notice that $\mathbb{E}^n \geq 0$ for all $n \geq 1$ since $\Phi_1 \geq 0$. For further use, we extend the function $\Phi_1$ by continuity to 0, letting $\Phi_1(0) = 1$. We can then define, in case there exists $K \in \mathcal{M}$ such that $u_K^0 = 0$, $\mathbb{E}^0 \geq 0$ according to (3.35). As far as $\mathbb{D}^n$ is concerned, by (3.30)-(3.25), for all $n \geq 1$, we have

$$\mathbb{D}^n = \sum_{K\in\mathcal{M}} r_K(\underline{u}_K^n) a_K^\Lambda(\underline{w}_K^n, \underline{w}_K^n) \geq 0.$$

We can now establish the following discrete counterpart of (3.22).

**Proposition 5** (Entropy dissipation). *If $\left(\underline{u}_{\mathcal{D}}^n \in \underline{V}_{\mathcal{D}}\right)_{n\geq 1}$ is a (positive) solution to (3.31), then*

$$\forall n \in \mathbb{N}, \qquad \frac{\mathbb{E}^{n+1} - \mathbb{E}^n}{\Delta t} + \mathbb{D}^{n+1} \leq 0. \tag{3.36}$$

*Proof.* Let $n \in \mathbb{N}$. By the expression (3.35) of the discrete relative entropy, and the convexity of $\Phi_1$, we have

$$\mathbb{E}^{n+1} - \mathbb{E}^n \leq \sum_{K\in\mathcal{M}} |K| u_K^\infty \Phi_1'\left(\frac{u_K^{n+1}}{u_K^\infty}\right)\left(\frac{u_K^{n+1} - u_K^n}{u_K^\infty}\right).$$

Thus, by (3.33), we get

$$\mathbb{E}^{n+1} - \mathbb{E}^n \leq \int_\Omega \left(u_{\mathcal{M}}^{n+1} - u_{\mathcal{M}}^n\right) \log\left(\frac{u_{\mathcal{M}}^{n+1}}{u_{\mathcal{M}}^\infty}\right) = (u_{\mathcal{M}}^{n+1} - u_{\mathcal{M}}^n, w_{\mathcal{M}}^{n+1})_\Omega.$$

By (3.31a) and (3.35), we finally infer

$$\mathbb{E}^{n+1} - \mathbb{E}^n \leq -\Delta t\, T_{\mathcal{D}}(\underline{u}_{\mathcal{D}}^{n+1}, \underline{w}_{\mathcal{D}}^{n+1}, \underline{w}_{\mathcal{D}}^{n+1}) = -\Delta t\, \mathbb{D}^{n+1},$$

which yields (3.36). □

We finally state the main result of Section 3.3, about the existence of (positive) solutions to the nonlinear scheme (3.31). The proof of this result is the subject of the next subsection.

**Theorem 1** (Existence of positive solutions). *Let $u^{in} \in L^2(\Omega)$ be a non-negative function such that $\int_\Omega u^{in} = M > 0$. There exists at least one positive solution $(\underline{u}_{\mathcal{D}}^n \in \underline{V}_{\mathcal{D}})_{n\geq 1}$ to the nonlinear scheme (3.31). Moreover, there exists $\varepsilon > 0$, depending on $\Lambda$, $\phi$, $u^{in}$, $M$, $\Omega$, $d$, $\Delta t$, and $\mathcal{D}$ such that*

$$\forall n \geq 1, \qquad u_K^n \geq \varepsilon \quad \forall K \in \mathcal{M} \qquad \text{and} \qquad u_\sigma^n \geq \varepsilon \quad \forall \sigma \in \mathcal{E}. \tag{3.37}$$

The uniform-in-time positivity result (3.37) on discrete solutions is the equivalent in the HFV context of [10, Lemma 3.7] and [9, Lemma 3.5] obtained, respectively, in the VAG and DDFV contexts.

### 3.3.2 Existence of discrete solutions

The existence of discrete solutions to the nonlinear scheme (3.31) is proved in two steps. First, we introduce a regularised scheme, for which we prove the existence of solutions by a fixed-point argument. Then, we prove that sequences of regularised solutions satisfy uniform a priori bounds, which allows us to pass to the limit in the regularisation parameter. Notice that our proof of existence is based on arguments that are quite different from the topological degree arguments used for proving equivalent results in the VAG [10] and DDFV [9] contexts. Henceforth, we reason in the $\underline{w}_{\mathcal{D}}$ variable, and we recall that $\underline{u}_{\mathcal{D}} = \underline{u}_{\mathcal{D}}^{\infty} \times \exp(\underline{w}_{\mathcal{D}})$ according to (3.33). The advantage of doing so is that we can seek for solutions $\underline{w}_{\mathcal{D}}$ in the whole space $\underline{V}_{\mathcal{D}}$, with bijective correspondence with solutions $\underline{u}_{\mathcal{D}}$ that are automatically positive. Recalling the definition (3.35) of the discrete relative entropy and dissipation, and using (3.30) combined with (3.26), we let, for all $\underline{w}_{\mathcal{D}} \in \underline{V}_{\mathcal{D}}$,

$$\mathbb{E}(w_{\mathcal{M}}) := \sum_{K\in\mathcal{M}} |K| u_K^{\infty} \Phi_1(\mathrm{e}^{w_K}), \qquad \mathbb{D}(\underline{w}_{\mathcal{D}}) := \sum_{K\in\mathcal{M}} r_K\big(\underline{u}_K^{\infty} \times \exp(\underline{w}_K)\big)\, \delta_K \underline{w}_K \cdot \mathbb{A}_K \delta_K \underline{w}_K, \quad (3.38)$$

in such a way that $\mathbb{E}^n = \mathbb{E}\big(w_{\mathcal{M}}(u_{\mathcal{M}}^n)\big)$ and $\mathbb{D}^n = \mathbb{D}\big(\underline{w}_{\mathcal{D}}(\underline{u}_{\mathcal{D}}^n)\big)$ for all $n \geq 1$. Using the fact that $\Phi_1(0) = 1$, we extend the definition of $\mathbb{E}(w_{\mathcal{M}})$ to the case where some $w_K$'s are equal to $-\infty$.

Before proceeding with the proof of Theorem 1, we state two preliminary lemmas. The first one, that can be found, e.g., in [23, Section 9.1], is a corollary of Brouwer's fixed-point theorem. This result is instrumental to show the existence of solutions to the regularised scheme.

**Lemma 1.** *Let $N \in \mathbb{N}^\star$, and let $P : \mathbb{R}^N \to \mathbb{R}^N$ be a continuous vector field. Assume that there is $r > 0$ such that*

$$P(x) \cdot x \geq 0 \qquad \text{if } |x| = r.$$

*Then, there exists a point $x_0 \in \mathbb{R}^N$ such that $P(x_0) = 0$ and $|x_0| \leq r$.*

The second lemma, whose proof is postponed until Appendix B.1, establishes sufficient boundedness conditions on the discrete mass and (relative) dissipation so that a priori bounds hold for vectors of discrete unknowns. This result is instrumental to show that sequences of regularised solutions satisfy (uniform) a priori bounds.

**Lemma 2.** *Let $\underline{w}_{\mathcal{D}} \in \underline{V}_{\mathcal{D}}$, and assume that there exist $C_\sharp \geq 0$, and $M_\sharp \geq M_\flat > 0$ such that*

$$M_\flat \leq \sum_{K\in\mathcal{M}} |K| u_K^{\infty}\, \mathrm{e}^{w_K} \leq M_\sharp \qquad \text{and} \qquad \mathbb{D}(\underline{w}_{\mathcal{D}}) \leq C_\sharp. \quad (3.39)$$

*Then, there exists $C > 0$, depending on $\Lambda$, $u_\flat^{\infty}$, $u_\sharp^{\infty}$, $M_\flat$, $M_\sharp$, $C_\sharp$, $\Omega$, $d$, and $\mathcal{D}$ such that*

$$|w_K| \leq C \quad \forall K \in \mathcal{M} \qquad \text{and} \qquad |w_\sigma| \leq C \quad \forall \sigma \in \mathcal{E}.$$

We can now proceed with the proof of Theorem 1. Let us first define the following inner product and corresponding norm on the space $\underline{V}_{\mathcal{D}}$: for all $\underline{w}_{\mathcal{D}}, \underline{v}_{\mathcal{D}} \in \underline{V}_{\mathcal{D}}$,

$$\langle \underline{w}_{\mathcal{D}}, \underline{v}_{\mathcal{D}} \rangle := \sum_{K\in\mathcal{M}} w_K v_K + \sum_{\sigma\in\mathcal{E}} w_\sigma v_\sigma \qquad \text{and} \qquad \|\underline{v}_{\mathcal{D}}\| := \sqrt{\langle \underline{v}_{\mathcal{D}}, \underline{v}_{\mathcal{D}} \rangle}.$$

Letting $N := |\mathcal{M}| + |\mathcal{E}|$, and identifying $\underline{V}_{\mathcal{D}}$ to $\mathbb{R}^N$, the inner product $\langle \cdot, \cdot \rangle$ is nothing but the standard inner product on $\mathbb{R}^N$. For all $K \in \mathcal{M}$, and all $\sigma \in \mathcal{E}_K$, we let, for $\underline{u}_K \in \underline{V}_K$ positive,

$$\mathcal{F}_{K,\sigma}^{nl}(\underline{u}_K) := r_K(\underline{u}_K) \sum_{\sigma'\in\mathcal{E}_K} A_K^{\sigma\sigma'} \left( \log\left(\frac{u_K}{u_K^{\infty}}\right) - \log\left(\frac{u_{\sigma'}}{u_{\sigma'}^{\infty}}\right) \right), \quad (3.40)$$

where the $A_K^{\sigma\sigma'}$ are defined by (2.8). Combining (3.25)-(3.33) with (2.9)-(2.7) and (3.40), we have that $T_K(\underline{u}_K, \underline{w}_K, \underline{v}_K) = \sum_{\sigma\in\mathcal{E}_K} \mathcal{F}^{nl}_{K,\sigma}(\underline{u}_K)(v_K - v_\sigma)$ for all $\underline{v}_K \in \underline{V}_K$. In what follows, we let $n \in \mathbb{N}^\star$ and $u_{\mathcal{M}}^{n-1} \geq 0$ be given. We assume that $M^{n-1} := \int_\Omega u_{\mathcal{M}}^{n-1} > 0$ and that $u_{\mathcal{M}}^{n-1} > 0$ if $n > 1$. Letting, for any $\underline{u}_{\mathcal{D}} \in \underline{V}_{\mathcal{D}}$ positive, $\underline{\mathcal{G}}^n_{\mathcal{D}}(\underline{u}_{\mathcal{D}})$ be the element of $\underline{V}_{\mathcal{D}}$ such that

$$\mathcal{G}^n_K(\underline{u}_{\mathcal{D}}) := |K|\frac{u_K - u_K^{n-1}}{\Delta t} + \sum_{\sigma\in\mathcal{E}_K} \mathcal{F}^{nl}_{K,\sigma}(\underline{u}_K) \qquad \forall K \in \mathcal{M}, \tag{3.41a}$$

$$\mathcal{G}^n_\sigma(\underline{u}_{\mathcal{D}}) := -\big(\mathcal{F}^{nl}_{K,\sigma}(\underline{u}_K) + \mathcal{F}^{nl}_{L,\sigma}(\underline{u}_L)\big) \qquad \forall \sigma = K \mid L \in \mathcal{E}_{int}, \tag{3.41b}$$

$$\mathcal{G}^n_\sigma(\underline{u}_{\mathcal{D}}) := -\mathcal{F}^{nl}_{K,\sigma}(\underline{u}_K) \qquad \forall \sigma \in \mathcal{E}_{ext}, \tag{3.41c}$$

we infer that, for all $\underline{v}_{\mathcal{D}} \in \underline{V}_{\mathcal{D}}$,

$$\frac{1}{\Delta t}\left(u_{\mathcal{M}} - u_{\mathcal{M}}^{n-1}, v_{\mathcal{M}}\right)_\Omega + T_{\mathcal{D}}(\underline{u}_{\mathcal{D}}, \underline{w}_D(\underline{u}_D), \underline{v}_{\mathcal{D}}) = \left\langle \underline{\mathcal{G}}^n_{\mathcal{D}}(\underline{u}_{\mathcal{D}}), \underline{v}_{\mathcal{D}}\right\rangle. \tag{3.42}$$

Hence, a positive vector $\underline{u}^n_{\mathcal{D}} \in \underline{V}_{\mathcal{D}}$ is a solution to the nonlinear equation (3.31a) if and only if $\underline{\mathcal{G}}^n_{\mathcal{D}}(\underline{u}^n_{\mathcal{D}}) = \underline{0}$. With this observation in hand, we now detail the two steps of the proof.

**Step 1:** Using the relation $\underline{u}_{\mathcal{D}} = \underline{u}^\infty_{\mathcal{D}} \times \exp(\underline{w}_{\mathcal{D}})$, we define the vector field $\underline{\mathcal{P}}^{n,\mu}_{\mathcal{D}} : \underline{V}_{\mathcal{D}} \to \underline{V}_{\mathcal{D}}$ such that, for all $\underline{w}_{\mathcal{D}} \in \underline{V}_{\mathcal{D}}$,

$$\underline{\mathcal{P}}^{n,\mu}_{\mathcal{D}}(\underline{w}_{\mathcal{D}}) := \underline{\mathcal{G}}^n_{\mathcal{D}}\big(\underline{u}^\infty_{\mathcal{D}} \times \exp(\underline{w}_{\mathcal{D}})\big) + \mu\,\underline{w}_{\mathcal{D}}, \tag{3.43}$$

with $\underline{\mathcal{G}}^n_{\mathcal{D}}$ defined by (3.41) and $\mu \geq 0$. Notice that, unlike $\underline{\mathcal{G}}^n_{\mathcal{D}}$, the vector field $\underline{\mathcal{P}}^{n,\mu}_{\mathcal{D}}$ is continuous on the whole space $\underline{V}_{\mathcal{D}}$ for any $\mu \geq 0$. If $\underline{w}^n_{\mathcal{D}} \in \underline{V}_{\mathcal{D}}$ satisfies $\underline{\mathcal{P}}^{n,0}_{\mathcal{D}}(\underline{w}^n_{\mathcal{D}}) = \underline{0}$, then letting $\underline{u}^n_{\mathcal{D}} := \underline{u}^\infty_{\mathcal{D}} \times \exp(\underline{w}^n_{\mathcal{D}})$, we have $\underline{\mathcal{G}}^n_{\mathcal{D}}(\underline{u}^n_{\mathcal{D}}) = \underline{0}$, therefore $\underline{u}^n_{\mathcal{D}}$ is a (positive) solution to (3.31a). For $\mu > 0$, the problem of finding $\underline{w}^{n,\mu}_{\mathcal{D}} \in \underline{V}_{\mathcal{D}}$ such that $\underline{\mathcal{P}}^{n,\mu}_{\mathcal{D}}(\underline{w}^{n,\mu}_{\mathcal{D}}) = \underline{0}$ can thus be seen as a regularisation of the original problem. By (3.43) and (3.42), for all $\underline{w}_{\mathcal{D}} \in \underline{V}_{\mathcal{D}}$, we have

$$\begin{aligned}\left\langle \underline{\mathcal{P}}^{n,\mu}_{\mathcal{D}}(\underline{w}_{\mathcal{D}}), \underline{w}_{\mathcal{D}}\right\rangle = &\sum_{K\in\mathcal{M}} \frac{|K|}{\Delta t}\left(u_K^\infty\,\mathrm{e}^{w_K} - u_K^{n-1}\right) w_K \\ &+ \sum_{K\in\mathcal{M}} r_K\big(\underline{u}^\infty_K \times \exp(\underline{w}_K)\big)\delta_K\underline{w}_K \cdot \mathbb{A}_K\delta_K\underline{w}_K + \mu\|\underline{w}_{\mathcal{D}}\|^2.\end{aligned} \tag{3.44}$$

By (3.38), we recognise in the second term of the right-hand side the quantity $\mathbb{D}(\underline{w}_{\mathcal{D}}) \geq 0$. As far as the first term is concerned, for $n > 1$, by positivity of the $(u_K^{n-1})_{K\in\mathcal{M}}$, there exist real numbers $(w_K^{n-1})_{K\in\mathcal{M}}$ such that $u_K^{n-1} = u_K^\infty\,\mathrm{e}^{w_K^{n-1}}$ for all $K \in \mathcal{M}$, and since $\Phi_1$ is convex,

$$\begin{aligned}\sum_{K\in\mathcal{M}} \frac{|K|}{\Delta t}\left(u_K^\infty\,\mathrm{e}^{w_K} - u_K^{n-1}\right) w_K &= \sum_{K\in\mathcal{M}} \frac{|K|}{\Delta t} u_K^\infty w_K\left(\mathrm{e}^{w_K} - \mathrm{e}^{w_K^{n-1}}\right) \\ &\geq \sum_{K\in\mathcal{M}} \frac{|K|}{\Delta t} u_K^\infty\left(\Phi_1\left(\mathrm{e}^{w_K}\right) - \Phi_1\left(\mathrm{e}^{w_K^{n-1}}\right)\right) = \frac{\mathbb{E}(w_{\mathcal{M}}) - \mathbb{E}(w_{\mathcal{M}}^{n-1})}{\Delta t},\end{aligned} \tag{3.45}$$

where we have used the definition (3.38) of $\mathbb{E}(w_{\mathcal{M}})$. For $n = 1$, now, it may happen that $u_K^0$ be zero for some $K \in \mathcal{M}$, and then $w_K^0$ such that $u_K^0 = u_K^\infty\,\mathrm{e}^{w_K^0}$ cannot be defined. However, letting in that case $w_K^0 := -\infty$, the inequality above still holds since $\Phi_1(0) = 1$ and $\Phi_1(\mathrm{e}^s) - 1 \leq s\,\mathrm{e}^s$ for all $s \in \mathbb{R}$. By non-negativity of $\mathbb{E}(w_{\mathcal{M}})$, we finally infer from (3.44) and (3.45) that

$$\left\langle \underline{\mathcal{P}}^{n,\mu}_{\mathcal{D}}(\underline{w}_{\mathcal{D}}), \underline{w}_{\mathcal{D}}\right\rangle \geq \mu\|\underline{w}_{\mathcal{D}}\|^2 - \frac{\mathbb{E}(w_{\mathcal{M}}^{n-1})}{\Delta t},$$

so that we have $\langle \underline{\mathcal{P}}_{\mathcal{D}}^{n,\mu}(\underline{w}_{\mathcal{D}}), \underline{w}_{\mathcal{D}} \rangle \geq 0$ if $\|\underline{w}_{\mathcal{D}}\| = \sqrt{\frac{\mathbb{E}(w_{\mathcal{M}}^{n-1})}{\mu \Delta t}}$. By Lemma 1, we conclude about the existence of solutions to the regularised scheme. There exists $\underline{w}_{\mathcal{D}}^{n,\mu} \in \underline{V}_{\mathcal{D}}$ such that

$$\underline{\mathcal{P}}_{\mathcal{D}}^{n,\mu}(\underline{w}_{\mathcal{D}}^{n,\mu}) = \underline{0} \quad \text{and} \quad \|\underline{w}_{\mathcal{D}}^{n,\mu}\| \leq \sqrt{\frac{\mathbb{E}(w_{\mathcal{M}}^{n-1})}{\mu \Delta t}}. \tag{3.46}$$

**Step 2:** Since $\langle \underline{\mathcal{P}}_{\mathcal{D}}^{n,\mu}(\underline{w}_{\mathcal{D}}^{n,\mu}), \underline{w}_{\mathcal{D}}^{n,\mu} \rangle = 0$, by (3.44) and (3.45), we have

$$\frac{\mathbb{E}(w_{\mathcal{M}}^{n,\mu})}{\Delta t} + \mathbb{D}(\underline{w}_{\mathcal{D}}^{n,\mu}) + \mu \|\underline{w}_{\mathcal{D}}^{n,\mu}\|^2 \leq \frac{\mathbb{E}(w_{\mathcal{M}}^{n-1})}{\Delta t}.$$

The three terms on the left-hand side being non-negative, we infer that

$$\mathbb{D}(\underline{w}_{\mathcal{D}}^{n,\mu}) \leq C_\sharp, \tag{3.47}$$

with $C_\sharp := \frac{\mathbb{E}(w_{\mathcal{M}}^{n-1})}{\Delta t} \geq 0$. Moreover, since $\langle \underline{\mathcal{P}}_{\mathcal{D}}^{n,\mu}(\underline{w}_{\mathcal{D}}^{n,\mu}), \underline{1}_{\mathcal{D}} \rangle = 0$, by (3.43) and (3.42), we have

$$\sum_{K \in \mathcal{M}} |K| u_K^\infty \, \mathrm{e}^{w_K^{n,\mu}} - M^{n-1} = -\mu \Delta t \left\langle \underline{w}_{\mathcal{D}}^{n,\mu}, \underline{1}_{\mathcal{D}} \right\rangle.$$

Applying a Cauchy–Schwarz inequality, and recalling the bound (3.46), we obtain

$$\left| \sum_{K \in \mathcal{M}} |K| u_K^\infty \, \mathrm{e}^{w_K^{n,\mu}} - M^{n-1} \right| \leq \mu \Delta t \|\underline{1}_{\mathcal{D}}\| \|\underline{w}_{\mathcal{D}}^{n,\mu}\| \leq \sqrt{\mu} \sqrt{N \Delta t \, \mathbb{E}(w_{\mathcal{M}}^{n-1})},$$

so that, letting $\mu_0 := \frac{(M^{n-1})^2}{4N \Delta t \, \mathbb{E}(w_{\mathcal{M}}^{n-1})} > 0$, the following holds for all $0 \leq \mu \leq \mu_0$:

$$\frac{M^{n-1}}{2} =: M_\flat \leq \sum_{K \in \mathcal{M}} |K| u_K^\infty \, \mathrm{e}^{w_K^{n,\mu}} \leq M_\sharp := \frac{3M^{n-1}}{2}. \tag{3.48}$$

Remark that if $\mathbb{E}(w_{\mathcal{M}}^{n-1}) = 0$, then $\sum_{K \in \mathcal{M}} |K| u_K^\infty \, \mathrm{e}^{w_K^{n,\mu}} = M^{n-1}$ and the inequality (3.48) remains valid for all $\mu \geq 0$. By (3.48) and (3.47), we infer that $\underline{w}_{\mathcal{D}}^{n,\mu}$ satisfies (3.39) for $\mu$ sufficiently small with constants that are uniform in $\mu$, so that by Lemma 2 the family $(\underline{w}_{\mathcal{D}}^{n,\mu})_{0 \leq \mu \leq \mu_0}$ is bounded uniformly in $\mu$. As a consequence, by compactness, there is $\underline{w}_{\mathcal{D}}^{n} \in \underline{V}_{\mathcal{D}}$ such that, up to extraction, $\underline{w}_{\mathcal{D}}^{n,\mu}$ converges towards $\underline{w}_{\mathcal{D}}^{n}$ when $\mu$ tends to zero. Since $\underline{\mathcal{P}}_{\mathcal{D}}^{n,\mu}$ converges to $\underline{\mathcal{P}}_{\mathcal{D}}^{n,0}$ as $\mu$ tends to zero, we finally infer that $\underline{\mathcal{P}}_{\mathcal{D}}^{n,0}(\underline{w}_{\mathcal{D}}^{n}) = \underline{0}$ (also, $\sum_{K \in \mathcal{M}} |K| u_K^\infty \, \mathrm{e}^{w_K^{n}} = M^{n-1}$).

**Conclusion:** Letting $\underline{u}_{\mathcal{D}}^{n} := \underline{u}_{\mathcal{D}}^{\infty} \times \exp(\underline{w}_{\mathcal{D}}^{n})$, we have $\underline{\mathcal{G}}_{\mathcal{D}}^{n}(\underline{u}_{\mathcal{D}}^{n}) = \underline{0}$, therefore $\underline{u}_{\mathcal{D}}^{n}$ is a (positive) solution to (3.31a). By Propositions 4 and 5, and since $\mathbb{D}(\underline{w}_{\mathcal{D}}^{n}) = \mathbb{D}^n$, $\mathbb{E}^n$ is non-negative, and $\mathbb{E}^n$ is non-increasing in $n$ according to (3.36), we deduce that

$$\sum_{K \in \mathcal{M}} |K| u_K^\infty \, \mathrm{e}^{w_K^{n}} = M \quad \text{and} \quad \mathbb{D}(\underline{w}_{\mathcal{D}}^{n}) \leq \frac{\mathbb{E}^0}{\Delta t}.$$

By Lemma 2 (recall also that $u_\flat^\infty, u_\sharp^\infty$ only depend on $M$, $\phi$, and $\Omega$), there exists $C > 0$, depending on $\Lambda$, $\phi$, $u^{in}$, $M$, $\Omega$, $d$, $\Delta t$, and $\mathcal{D}$, but not on $n \in \mathbb{N}^\star$, such that $-C\underline{1}_{\mathcal{D}} \leq \underline{w}_{\mathcal{D}}^{n} \leq C\underline{1}_{\mathcal{D}}$. By (3.34), we finally infer that $\underline{u}_{\mathcal{D}}^{n} \geq \varepsilon \underline{1}_{\mathcal{D}}$, with $\varepsilon := u_\flat^\infty \, \mathrm{e}^{-C} > 0$ still independent of $n \in \mathbb{N}^\star$. This concludes the proof of Theorem 1.

# 4 Long-time behaviour

In this section, we analyse the long-time behaviour of the three HFV schemes we have introduced in Section 3, thereby proving the main results of this paper.

**Remark 7** (Linear schemes and nonhomogeneous data)**.** *In order to stay consistent with Section 3, we here below state our asymptotic results of Theorems 2 and 3, which respectively concern the (linear) standard and exponential fitting schemes, for discrete problems that feature homogeneous data (i.e., $g^D = 0$ when $|\Gamma^D| > 0$, or $M = 0$ otherwise). Nonetheless, Theorems 2 and 3 remain valid in the general case of nonhomogeneous data (we refer to Remark 1 for the straightforward adaptation of the schemes to this situation). Indeed, the proofs of the latter results solely hinge on the fact that the difference between the discrete transient and steady-state solutions belongs to the homogeneous space $\underline{V}_{\mathcal{D},0}$ or $\underline{V}^{\omega}_{\mathcal{D},0}$, which remains true in the general case. This remark does not apply, however, to the nonlinear scheme.*

## 4.1 Asymptotic behaviour of the standard HFV scheme

We recall that $u$ is the solution to Problem (1.1), and that $u^\infty$ is the corresponding steady-state, solution to Problem (1.2), and we consider the following definition of the relative entropy and dissipation:

$$\mathbb{E}(t) := \frac{1}{2}\left\|u(t) - u^\infty\right\|^2_{L^2(\Omega)}, \qquad \mathbb{D}(t) := \int_\Omega \Lambda\big(\nabla(u(t) - u^\infty) + (u(t) - u^\infty)\nabla\phi\big) \cdot \nabla(u(t) - u^\infty).$$

It can be easily verified that the following entropy/dissipation relation holds at the continuous level:

$$\frac{\mathrm{d}}{\mathrm{d}t}\mathbb{E}(t) + \mathbb{D}(t) = 0.$$

At the discrete level, recalling that $\big(\underline{u}^n_{\mathcal{D}} \in \underline{V}_{\mathcal{D},0}\big)_{n\geq 1}$ is the solution to Problem (3.12), and that $\underline{u}^\infty_{\mathcal{D}} \in \underline{V}_{\mathcal{D},0}$ is the corresponding steady-state, solution to Problem (3.6), we consider the following equivalents of the relative entropy and dissipation: for all $n \in \mathbb{N}^\star$,

$$\mathbb{E}^n := \frac{1}{2}\left\|u^n_{\mathcal{M}} - u^\infty_{\mathcal{M}}\right\|^2_{L^2(\Omega)}, \qquad \mathbb{D}^n := a_{\mathcal{D}}\big(\underline{u}^n_{\mathcal{D}} - \underline{u}^\infty_{\mathcal{D}}, \underline{u}^n_{\mathcal{D}} - \underline{u}^\infty_{\mathcal{D}}\big),$$

where the discrete bilinear form $a_{\mathcal{D}}$ is defined by (3.5)-(3.1). The definition of the relative entropy is seamlessly extended to the case $n = 0$. Our main result on the standard HFV scheme is the following.

**Theorem 2** (Asymptotic stability)**.** *Assume that the advection field $V^\phi$ satisfies the conditions (3.7) of Proposition 2, with constant $\beta$. Then, the following discrete entropy/dissipation relation holds true:*

$$\forall n \in \mathbb{N}, \qquad \frac{\mathbb{E}^{n+1} - \mathbb{E}^n}{\Delta t} + \mathbb{D}^{n+1} \leq 0. \tag{4.1}$$

*Furthermore, the discrete entropy decays exponentially fast in time: there is $\nu := \frac{2\kappa}{C_P^2} > 0$, where $\kappa$ is the constant of (3.8) (only depending on $\Lambda$, $\beta$, $\Omega$, $d$, $\Gamma^D$, and $\theta_{\mathcal{D}}$), and $C_P$ is either equal to $C_{PW}$ if $|\Gamma^D| = 0$ or to $C_{P,\Gamma^D}$ otherwise (where $C_{PW}$, $C_{P,\Gamma^D}$ are the Poincaré constants of Proposition 7), such that*

$$\forall n \in \mathbb{N}, \qquad \mathbb{E}^{n+1} \leq (1 + \nu\Delta t)^{-1}\,\mathbb{E}^n. \tag{4.2}$$

*Consequently, the discrete solution converges exponentially fast in time towards its associated discrete steady-state: for all $n \in \mathbb{N}^\star$,*

$$\left\|u^n_{\mathcal{M}} - u^\infty_{\mathcal{M}}\right\|_{L^2(\Omega)} \leq (1 + \nu\Delta t)^{-\frac{n}{2}}\left\|u^0_{\mathcal{M}} - u^\infty_{\mathcal{M}}\right\|_{L^2(\Omega)}. \tag{4.3}$$

*Proof.* Let $n \in \mathbb{N}$. First, remark that

$$\mathbb{E}^n = \sum_{K\in\mathcal{M}} |K| \left( \frac{(u_K^n)^2 - (u_K^\infty)^2}{2} - u_K^\infty \big(u_K^n - u_K^\infty\big) \right).$$

Hence, we can infer that

$$\mathbb{E}^{n+1} - \mathbb{E}^n = \sum_{K\in\mathcal{M}} |K| \left( \frac{(u_K^{n+1})^2 - (u_K^n)^2}{2} - u_K^\infty \big(u_K^{n+1} - u_K^n\big) \right).$$

Since $x \mapsto x^2$ is convex, for all $x, y \in \mathbb{R}$, we have $y^2 - x^2 \le 2y(y-x)$, therefore

$$\mathbb{E}^{n+1} - \mathbb{E}^n \le \sum_{K\in\mathcal{M}} |K| \big(u_K^{n+1} - u_K^n\big)\big(u_K^{n+1} - u_K^\infty\big) = \big(u_\mathcal{M}^{n+1} - u_\mathcal{M}^n, u_\mathcal{M}^{n+1} - u_\mathcal{M}^\infty\big)_\Omega. \tag{4.4}$$

Now, testing (3.12a) with $\underline{v}_\mathcal{D}^{n+1} := \underline{u}_\mathcal{D}^{n+1} - \underline{u}_\mathcal{D}^\infty \in \underline{V}_{\mathcal{D},0}$ yields

$$\big(u_\mathcal{M}^{n+1} - u_\mathcal{M}^n, v_\mathcal{M}^{n+1}\big)_\Omega = -\Delta t a_\mathcal{D}\big(\underline{u}_\mathcal{D}^{n+1}, \underline{v}_\mathcal{D}^{n+1}\big) + \Delta t\big((f, v_\mathcal{M}^{n+1})_\Omega + (g^N, v_\mathcal{E}^{n+1})_{\Gamma^N}\big).$$

By definition (3.6) of the discrete steady-state $\underline{u}_\mathcal{D}^\infty$, we also have

$$(f, v_\mathcal{M}^{n+1})_\Omega + (g^N, v_\mathcal{E}^{n+1})_{\Gamma^N} = a_\mathcal{D}\big(\underline{u}_\mathcal{D}^\infty, \underline{v}_\mathcal{D}^{n+1}\big),$$

whence, by bilinearity of $a_\mathcal{D}$, we infer

$$\big(u_\mathcal{M}^{n+1} - u_\mathcal{M}^n, v_\mathcal{M}^{n+1}\big)_\Omega = -\Delta t\, a_\mathcal{D}\big(\underline{u}_\mathcal{D}^{n+1} - \underline{u}_\mathcal{D}^\infty, \underline{v}_\mathcal{D}^{n+1}\big) = -\Delta t\, \mathbb{D}^{n+1}.$$

Combined to (4.4), this proves the entropy/dissipation relation (4.1). Now, since the advection field $V^\phi$ satisfies (3.7), we can invoke (3.8) from Proposition 2 to infer that

$$\mathbb{D}^{n+1} \ge \kappa |\underline{u}_\mathcal{D}^{n+1} - \underline{u}_\mathcal{D}^\infty|_{1,\mathcal{D}}^2,$$

where $\kappa > 0$ only depends on $\Lambda$ ,$\beta$, $\Omega$, $d$, $\Gamma^D$, and $\theta_\mathcal{D}$. Combining this estimate with a discrete Poincaré inequality from Proposition 7 (applied to $\underline{u}_\mathcal{D}^{n+1} - \underline{u}_\mathcal{D}^\infty \in \underline{V}_{\mathcal{D},0}$), and with the definition of the discrete (relative) entropy, yields

$$\mathbb{D}^{n+1} \ge \frac{\kappa}{C_P^2} \big\|u_\mathcal{M}^{n+1} - u_\mathcal{M}^\infty\big\|_{L^2(\Omega)}^2 = \frac{2\kappa}{C_P^2} \mathbb{E}^{n+1},$$

where $C_P$ is either equal to $C_{PW}$ if $|\Gamma^D| = 0$ or to $C_{P,\Gamma^D}$ otherwise. This last inequality, combined with the entropy/dissipation relation (4.1), implies the entropy decay (4.2). The inequality (4.3) is then a straightforward consequence of the definition of $\mathbb{E}^n$. □

### 4.2 Asymptotic behaviour of the exponential fitting scheme

We recall that $\rho = \frac{u}{\omega}$, with $\omega = \mathrm{e}^{-\phi}$, is the solution to Problem (3.20), and $\rho^\infty$ is the corresponding steady-state, solution to Problem (3.14). We consider the following $\omega$-weighted definitions of the relative entropy and dissipation:

$$\mathbb{E}_\omega(t) := \frac{1}{2} \|\rho(t) - \rho^\infty\|_{L^2_\omega(\Omega)}^2, \qquad \mathbb{D}_\omega(t) := \int_\Omega \omega \Lambda \nabla(\rho(t) - \rho^\infty) \cdot \nabla(\rho(t) - \rho^\infty).$$

It can be easily verified that the following entropy/dissipation relation holds at the continuous level:

$$\frac{\mathrm{d}}{\mathrm{d}t} \mathbb{E}_\omega(t) + \mathbb{D}_\omega(t) = 0.$$

Let us recall that, at the discrete level, $\big(\underline{\rho}_{\mathcal{D}}^n \in \underline{V}_{\mathcal{D},0}^{\omega}\big)_{n\geq 1}$ is the solution to Problem (3.21). We then set $\underline{u}_{\mathcal{D}}^n = \underline{\omega}_{\mathcal{D}} \times \underline{\rho}_{\mathcal{D}}^n$ with $\underline{\omega}_{\mathcal{D}}$ defined by (3.17), such that $\underline{u}_{\mathcal{D}}^n \in \underline{V}_{\mathcal{D},0}$. Similarly, as $\underline{\rho}_{\mathcal{D}}^{\infty} \in \underline{V}_{\mathcal{D},0}^{\omega}$ is the corresponding steady-state, solution to Problem (3.16) , we set $\underline{u}_{\mathcal{D}}^{\infty} = \underline{\omega}_{\mathcal{D}} \times \underline{\rho}_{\mathcal{D}}^{\infty}$. Then, we consider the following equivalents of the $\omega$-weighted (relative) entropy and dissipation: for all $n \in \mathbb{N}^\star$,

$$\mathbb{E}_{\omega}^n := \frac{1}{2}\left\|\rho_{\mathcal{M}}^n - \rho_{\mathcal{M}}^{\infty}\right\|_{L^2_{\omega}(\Omega)}^2, \qquad \mathbb{D}_{\omega}^n := a_{\mathcal{D}}^{\omega}\big(\underline{\rho}_{\mathcal{D}}^n - \underline{\rho}_{\mathcal{D}}^{\infty}, \underline{\rho}_{\mathcal{D}}^n - \underline{\rho}_{\mathcal{D}}^{\infty}\big),$$

where the discrete bilinear form $a_{\mathcal{D}}^{\omega}$ is defined (locally) by (3.15). The definition of the relative entropy is seamlessly extended to the case $n = 0$. Our main result on the exponential fitting HFV scheme is the following, whose proof is very similar to the one of Theorem 2.

**Theorem 3** (Asymptotic stability)**.** *The following discrete entropy/dissipation relation holds true:*

$$\forall n \in \mathbb{N}, \qquad \frac{\mathbb{E}_{\omega}^{n+1} - \mathbb{E}_{\omega}^n}{\Delta t} + \mathbb{D}_{\omega}^{n+1} \leq 0. \tag{4.5}$$

*Furthermore, the discrete entropy decays exponentially fast in time: there is* $\nu_{\omega} := \frac{2\omega_\flat \lambda_\flat \alpha_\flat}{\omega_\sharp C_P^2} > 0$*, where* $\lambda_\flat \alpha_\flat$ *is the coercivity constant of* (2.12) *(only depending on* $\Lambda$*,* $\Omega$*,* $d$*, and* $\theta_{\mathcal{D}}$*), (we recall that) the bounds* $\omega_\flat, \omega_\sharp$ *only depend on* $\phi$ *and* $\Omega$*, and* $C_P$ *is either equal to* $2C_{PW}$ *if* $|\Gamma^D| = 0$ *or to* $C_{P,\Gamma^D}$ *otherwise (where* $C_{PW}$*,* $C_{P,\Gamma^D}$ *are the Poincaré constants of Proposition 7), such that*

$$\forall n \in \mathbb{N}, \qquad \mathbb{E}_{\omega}^{n+1} \leq (1 + \nu_{\omega}\Delta t)^{-1}\, \mathbb{E}_{\omega}^n. \tag{4.6}$$

*Consequently, the discrete solution converges exponentially fast in time towards its associated discrete steady-state: for all* $n \in \mathbb{N}^\star$*,*

$$\|u_{\mathcal{M}}^n - u_{\mathcal{M}}^{\infty}\|_{L^2(\Omega)} \leq \sqrt{\frac{\omega_\sharp}{\omega_\flat}}\,(1 + \nu_{\omega}\Delta t)^{-\frac{n}{2}}\, \|u_{\mathcal{M}}^0 - u_{\mathcal{M}}^{\infty}\|_{L^2(\Omega)}. \tag{4.7}$$

*Proof.* Let $n \in \mathbb{N}$. By convexity of $x \mapsto x^2$, we have

$$\begin{aligned}
\mathbb{E}_{\omega}^{n+1} - \mathbb{E}_{\omega}^n &= \sum_{K\in\mathcal{M}} |K|\omega_K \left(\frac{(\rho_K^{n+1})^2 - (\rho_K^n)^2}{2} - \rho_K^{\infty}\big(\rho_K^{n+1} - \rho_K^n\big)\right) \\
&\leq \sum_{K\in\mathcal{M}} |K|\omega_K \left(\rho_K^{n+1}\big(\rho_K^{n+1} - \rho_K^n\big) - \rho_K^{\infty}\big(\rho_K^{n+1} - \rho_K^n\big)\right) \\
&= \big(\rho_{\mathcal{M}}^{n+1} - \rho_{\mathcal{M}}^n, \omega_{\mathcal{M}}\big(\rho_{\mathcal{M}}^{n+1} - \rho_{\mathcal{M}}^{\infty}\big)\big)_{\Omega}.
\end{aligned}$$

Testing (3.21a) with $\big(\underline{\rho}_{\mathcal{D}}^{n+1} - \underline{\rho}_{\mathcal{D}}^{\infty}\big) \in \underline{V}_{\mathcal{D},0}^{\omega}$, and using the definition (3.16) of $\underline{\rho}_{\mathcal{D}}^{\infty}$ along with the bilinearity of $a_{\mathcal{D}}^{\omega}$, we then get

$$\mathbb{E}_{\omega}^{n+1} - \mathbb{E}_{\omega}^n \leq -\Delta t\, a_{\mathcal{D}}^{\omega}\big(\underline{\rho}_{\mathcal{D}}^{n+1} - \underline{\rho}_{\mathcal{D}}^{\infty}, \underline{\rho}_{\mathcal{D}}^{n+1} - \underline{\rho}_{\mathcal{D}}^{\infty}\big) = -\Delta t\, \mathbb{D}_{\omega}^{n+1},$$

which proves (4.5). By the coercivity estimate (3.19), we infer that

$$\mathbb{D}_{\omega}^{n+1} \geq \omega_\flat \lambda_\flat \alpha_\flat \big|\underline{\rho}_{\mathcal{D}}^{n+1} - \underline{\rho}_{\mathcal{D}}^{\infty}\big|_{1,\mathcal{D}}^2,$$

where $\alpha_\flat > 0$ from Section 2.3 only depends on $\Omega$, $d$, and $\theta_{\mathcal{D}}$, and $\omega_\flat > 0$ only depends on $\phi$ and $\Omega$. Reasoning as in the proof of Proposition 3, and using a discrete Poincaré inequality from Proposition 7 (combined with [8, Lemma 5.2] in the case $|\Gamma^D| = 0$), we infer that

$$\|\rho_{\mathcal{M}}^{n+1} - \rho_{\mathcal{M}}^{\infty}\|_{L^2_{\omega}(\Omega)} \leq \sqrt{\omega_\sharp}\, C_P |\underline{\rho}_{\mathcal{D}}^{n+1} - \underline{\rho}_{\mathcal{D}}^{\infty}|_{1,\mathcal{D}},$$

where $\omega_\sharp > 0$ only depends on $\phi$ and $\Omega$, and $C_P$ is either equal to $2C_{PW}$ if $|\Gamma^D| = 0$ or to $C_{P,\Gamma^D}$ otherwise. Thus, we finally get that

$$\mathbb{D}_{\omega}^{n+1} \geq \frac{2\omega_\flat \lambda_\flat \alpha_\flat}{\omega_\sharp C_P^2} \mathbb{E}_{\omega}^{n+1}.$$

Combined to (4.5), this yields (4.6). Deriving (4.7) is then straightforward. □

### 4.3 Asymptotic behaviour of the nonlinear scheme

Recall that $u > 0$ is the solution to Problem (1.1) endowed with pure Neumann boundary conditions ($|\Gamma^D| = 0$), and data $f = 0$, $g^N = 0$, and $u^{in} \geq 0$ with $M = \int_\Omega u^{in} > 0$, and that $u^\infty > 0$, solution to Problem (1.2) with same data, is the thermal equilibrium $u^\infty_{th}$ given by (3.11). The analysis of the nonlinear scheme relies on the entropy/dissipation structure (3.22)-(3.23) introduced in Section 3.3. Notice that the relative dissipation (or relative Fisher information) of (3.23) can be equivalently rewritten

$$\mathbb{D}(t) = \int_\Omega u(t)\Lambda\nabla\log\left(\frac{u(t)}{u^\infty}\right)\cdot\nabla\log\left(\frac{u(t)}{u^\infty}\right) = 4\int_\Omega u^\infty\Lambda\nabla\sqrt{\frac{u(t)}{u^\infty}}\cdot\nabla\sqrt{\frac{u(t)}{u^\infty}}.$$

At the discrete level, recalling that $\left(\underline{u}^n_{\mathcal{D}} \in \underline{V}_{\mathcal{D}}\right)_{n\geq 1}$ is a (positive) solution to Problem (3.31), and that $\underline{u}^\infty_{\mathcal{D}} \in \underline{V}_{\mathcal{D}}$ is the corresponding steady-state, solution to Problem (3.32), that is equal to the HFV interpolate of $u^\infty_{th}$, we consider the discrete entropy $\mathbb{E}^n$ and dissipation $\mathbb{D}^n$ defined by (3.35), and we define a discrete equivalent of the relative dissipation written in root-form: for all $n \geq 1$,

$$\hat{\mathbb{D}}^n := 4\sum_{K\in\mathcal{M}} u^\infty_{K,\flat}\int_K \Lambda\nabla_K\underline{\xi}^n_K\cdot\nabla_K\underline{\xi}^n_K = 4\sum_{K\in\mathcal{M}} u^\infty_{K,\flat}\delta_K\underline{\xi}^n_K\cdot\mathbb{A}_K\delta_K\underline{\xi}^n_K, \tag{4.8}$$

where, for all $K \in \mathcal{M}$, $u^\infty_{K,\flat} := \min\left(u^\infty_K, \min_{\sigma\in\mathcal{E}_K} u^\infty_\sigma\right)$ and the matrix $\mathbb{A}_K$ is defined by (2.8), and $\underline{\xi}^n_{\mathcal{D}}$ is the element of $\underline{V}_{\mathcal{D}}$ such that

$$\xi^n_K := \sqrt{\frac{u^n_K}{u^\infty_K}} \quad \forall K\in\mathcal{M}, \qquad \xi^n_\sigma := \sqrt{\frac{u^n_\sigma}{u^\infty_\sigma}} \quad \forall\sigma\in\mathcal{E}. \tag{4.9}$$

At the discrete level, and as opposed to the continuous level, the quantities $\mathbb{D}^n$ and $\hat{\mathbb{D}}^n$ are not equal, therefore we need to compare them. The definition of $u^\infty_{K,\flat}$ results from the following observation: according to the structures of $\mathbb{D}^n$ and $\hat{\mathbb{D}}^n$, locally, we expect to have to compare $u^\infty_{K,\flat}$ with $r_K(\underline{u}^n_K)$, which depends on $u^n_K$ and on the $(u^n_\sigma)_{\sigma\in\mathcal{E}_K}$.

**Proposition 6** (Fisher information)**.** *There is $C_F > 0$, only depending on $\Lambda$, $\Omega$, $d$, and $\theta_{\mathcal{D}}$ such that*

$$\forall n \geq 1, \qquad \hat{\mathbb{D}}^n \leq C_F\mathbb{D}^n. \tag{4.10}$$

*Proof.* Let $n \in \mathbb{N}^\star$, and $K \in \mathcal{M}$. By (B.5) from Lemma 3, we first have that

$$4u^\infty_{K,\flat}\,\delta_K\underline{\xi}^n_K\cdot\mathbb{A}_K\delta_K\underline{\xi}^n_K \leq 4u^\infty_{K,\flat}\,\delta_K\underline{\xi}^n_K\cdot\mathbb{B}_K\delta_K\underline{\xi}^n_K, \tag{4.11}$$

where the matrix $\mathbb{B}_K$ is the diagonal matrix defined by (B.4). Since for all $(x,y) \in \left(\mathbb{R}^\star_+\right)^2$, $\int_x^y \frac{\mathrm{d}z}{\sqrt{z}} = 2\left(\sqrt{y}-\sqrt{x}\right)$, the Cauchy–Schwarz inequality yields

$$4\left(\sqrt{y}-\sqrt{x}\right)^2 \leq (y-x)\big(\log(y)-\log(x)\big).$$

By the property (3.28d) of the function $m$, we then get that

$$\forall(x,y)\in\left(\mathbb{R}^\star_+\right)^2, \qquad 4\left(\sqrt{y}-\sqrt{x}\right)^2 \leq m(x,y)\big(\log(y)-\log(x)\big)^2. \tag{4.12}$$

Since $\mathbb{B}_K$ is diagonal, the combination of (4.11), (4.9), and (4.12), yields

$$\begin{aligned}
4u^\infty_{K,\flat}\,\delta_K\underline{\xi}^n_K\cdot\mathbb{A}_K\delta_K\underline{\xi}^n_K &\leq \sum_{\sigma\in\mathcal{E}_K} 4u^\infty_{K,\flat}B^{\sigma\sigma}_K\left(\sqrt{\frac{u^n_K}{u^\infty_K}}-\sqrt{\frac{u^n_\sigma}{u^\infty_\sigma}}\right)^2\\
&\leq \sum_{\sigma\in\mathcal{E}_K} u^\infty_{K,\flat}B^{\sigma\sigma}_K m\left(\frac{u^n_K}{u^\infty_K},\frac{u^n_\sigma}{u^\infty_\sigma}\right)\left(\log\left(\frac{u^n_K}{u^\infty_K}\right)-\log\left(\frac{u^n_\sigma}{u^\infty_\sigma}\right)\right)^2.
\end{aligned}$$

By definition of $u^\infty_{K,\flat}$, and monotonicity (3.28a) and homogeneity (3.28c) of $m$, we infer that, for all $\sigma \in \mathcal{E}_K$,

$$u^\infty_{K,\flat} m\left(\frac{u^n_K}{u^\infty_K}, \frac{u^n_\sigma}{u^\infty_\sigma}\right) \le u^\infty_{K,\flat} m\left(\frac{u^n_K}{u^\infty_{K,\flat}}, \frac{u^n_\sigma}{u^\infty_{K,\flat}}\right) = m\left(u^n_K, u^n_\sigma\right).$$

By definition (3.29) of $f_{|\mathcal{E}_K|}$, and the bound (2.2) on $|\mathcal{E}_K|$, we then have

$$\max_{\sigma\in\mathcal{E}_K} u^\infty_{K,\flat} m\left(\frac{u^n_K}{u^\infty_K}, \frac{u^n_\sigma}{u^\infty_\sigma}\right) \le \max_{\sigma\in\mathcal{E}_K} m\left(u^n_K, u^n_\sigma\right) \le |\mathcal{E}_K| r_K(\underline{u}^n_K) \le d\theta^2_{\mathcal{D}}\, r_K(\underline{u}^n_K).$$

We deduce that

$$\begin{aligned} 4u^\infty_{K,\flat}\, \delta_K \underline{\xi}^n_K \cdot \mathbb{A}_K \delta_K \underline{\xi}^n_K &\le d\theta^2_{\mathcal{D}}\, r_K(\underline{u}^n_K) \sum_{\sigma\in\mathcal{E}_K} B^{\sigma\sigma}_K \left(\log\left(\frac{u^n_K}{u^\infty_K}\right) - \log\left(\frac{u^n_\sigma}{u^\infty_\sigma}\right)\right)^2 \\ &= d\theta^2_{\mathcal{D}}\, r_K(\underline{u}^n_K)\ \delta_K \underline{w}^n_K \cdot \mathbb{B}_K \delta_K \underline{w}^n_K, \end{aligned}$$

where $\underline{w}^n_K \in \underline{V}_K$ is such that $\underline{u}^n_K = \underline{u}^\infty_K \times \exp(\underline{w}^n_K)$. Using again (B.5) from Lemma 3, we finally infer that

$$4u^\infty_{K,\flat}\, \delta_K \underline{\xi}^n_K \cdot \mathbb{A}_K \delta_K \underline{\xi}^n_K \le d\theta^2_{\mathcal{D}} C_B\, r_K(\underline{u}^n_K)\ \delta_K \underline{w}^n_K \cdot \mathbb{A}_K \delta_K \underline{w}^n_K,$$

with $C_B > 0$ only depending on $\Lambda$, $\Omega$, $d$, and $\theta_{\mathcal{D}}$. Summing over $K \in \mathcal{M}$, and recalling the definitions (4.8) of $\hat{\mathbb{D}}^n$, and (3.35) of $\mathbb{D}^n$, eventually yields (4.10) with $C_F = d\theta^2_{\mathcal{D}} C_B$. □

The long-time behaviour of the nonlinear HFV scheme is studied in the following result.

**Theorem 4** (Asymptotic stability)**.** *Recall the discrete entropy/dissipation relation of Proposition 5. The discrete entropy decays exponentially fast in time: there is* $\nu_{nl} := \frac{4u^\infty_\flat \lambda_\flat \alpha_\flat}{C_F C^2_{LS}\sqrt{M u^\infty_\sharp}} > 0$*, depending on* $\Lambda$*,* $\phi$*,* $M$*,* $\Omega$*,* $d$*, and* $\theta_{\mathcal{D}}$ *such that*

$$\forall n \in \mathbb{N}, \qquad \mathbb{E}^{n+1} \le (1 + \nu_{nl}\,\Delta t)^{-1} \mathbb{E}^n. \tag{4.13}$$

*Consequently, the discrete solution converges exponentially fast in time towards its associated discrete steady-state: for all* $n \in \mathbb{N}^\star$*,*

$$\|u^n_{\mathcal{M}} - u^\infty_{\mathcal{M}}\|_{L^1(\Omega)} \le \sqrt{2M\mathbb{E}^0}\,(1 + \nu_{nl}\,\Delta t)^{-\frac{n}{2}}. \tag{4.14}$$

*Proof.* Let $n \in \mathbb{N}$. By definition (4.8) of $\hat{\mathbb{D}}^n$, and from the coercivity estimate (2.12), we first infer that

$$\hat{\mathbb{D}}^{n+1} = 4\sum_{K\in\mathcal{M}} u^\infty_{K,\flat} a^\Lambda_K\left(\underline{\xi}^{n+1}_K, \underline{\xi}^{n+1}_K\right) \ge 4u^\infty_\flat a^\Lambda_{\mathcal{D}}\left(\underline{\xi}^{n+1}_{\mathcal{D}}, \underline{\xi}^{n+1}_{\mathcal{D}}\right) \ge 4u^\infty_\flat \lambda_\flat \alpha_\flat \left|\underline{\xi}^{n+1}_{\mathcal{D}}\right|^2_{1,\mathcal{D}},$$

which, combined with (4.10), implies that

$$\mathbb{D}^{n+1} \ge \frac{1}{C_F}\hat{\mathbb{D}}^{n+1} \ge \frac{4u^\infty_\flat \lambda_\flat \alpha_\flat}{C_F}\left|\underline{\xi}^{n+1}_{\mathcal{D}}\right|^2_{1,\mathcal{D}}.$$

In order to compare this quantity with the entropy, we use the discrete log-Sobolev inequality (A.3) of Proposition 8, applied to the couple $(\underline{u}^{n+1}_{\mathcal{D}}, \underline{u}^\infty_{\mathcal{D}})$ (which satisfies the mass condition owing to Proposition 4). We get

$$\mathbb{E}^{n+1} = \int_\Omega u^\infty_{\mathcal{M}} \Phi_1\left(\frac{u^{n+1}_{\mathcal{M}}}{u^\infty_{\mathcal{M}}}\right) \le C^2_{LS}\sqrt{M u^\infty_\sharp}\,\left|\underline{\xi}^{n+1}_{\mathcal{D}}\right|^2_{1,\mathcal{D}},$$

which, combined with the previous estimate, yields

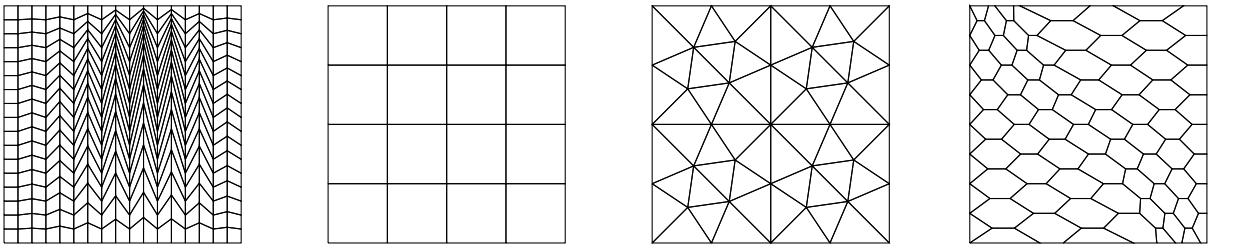
$$\mathbb{D}^{n+1} \geq \frac{4u_\flat^\infty \lambda_\flat \alpha_\flat}{C_F C_{LS}^2 \sqrt{M u_\sharp^\infty}} \mathbb{E}^{n+1}.$$

Combined with (3.36) from Proposition 5, this shows (4.13). The $L^1$-norm estimate (4.14) is then a direct consequence of (4.13) and of the Csiszár–Kullback lemma (cf., e.g., [8, Lemma 5.6]) applied to the probability measure $\mu(x)\,\mathrm{d}x = v_{\mathcal{M}}^\infty(x)\,\frac{\mathrm{d}x}{M}$ and to the function $g = \frac{u_{\mathcal{M}}^n}{u_{\mathcal{M}}^\infty}$. □

# 5 Numerical results

## 5.1 Implementation

In this Section, we discuss some practical details concerning the implemention of the schemes described in this paper. In all the test cases presented below, the two-dimensional domain is taken to be $\Omega = (0,1)^2$. The meshes used for the numerical tests are classical Cartesian, triangular, Kershaw meshes - see for example [27] for the first three ones - and a tilted hexagonal mesh, presented on Figure 2. These meshes have convex cells, hence we always choose $x_K$ as the barycentre of $K$. We



(a) Kershaw mesh (b) Cartesian mesh (c) Triangular mesh (d) Tilted hexagonal mesh

Figure 2: **Implementation.** Coarsest meshes of each family used in the numerical tests.

compute the meshsize as $\tilde{h}_{\mathcal{D}} = \max_{K\in\mathcal{M}} \frac{|K|}{|\partial K|}$, which is easier to evaluate than $h_{\mathcal{D}}$ in practice. Observe that $\tilde{h}_{\mathcal{D}}/h_{\mathcal{D}}$ is framed by constants depending only on the regularity of the mesh. Moreover, we use a fixed stabilisation parameter $\eta = 1.5$ (see (2.3)). In the sequel, we will denote by HMM the classical HFV linear scheme for advection-convection, and we restrict our attention to the Scharfetter–Gummel discretisation of the flux in (3.4), namely $A(s) = \frac{s}{\mathrm{e}^s - 1} - 1$, extended by continuity at $s = 0$.

### 5.1.1 Linear systems and static condensation

The two linear schemes (HMM and exponential fitting) are implemented in the same way. To fix ideas, we consider the stationary problem with pure Neumann boundary condition, of unknown $\underline{u} \in \underline{V}$, and we denote by $U_{\mathcal{M}} \in \mathbb{R}^{|\mathcal{M}|}$ and $U_{\mathcal{E}} \in \mathbb{R}^{|\mathcal{E}|}$ the unknown vectors $(u_K)_{K\in\mathcal{M}}$ and $(u_\sigma)_{\sigma\in\mathcal{E}}$. Testing the scheme against a basis of $\underline{V}$, we can write the linear scheme as the following block system:

$$\begin{pmatrix} \mathbb{M}_{\mathcal{M}} & \mathbb{M}_{\mathcal{M},\mathcal{E}} \\ \mathbb{M}_{\mathcal{E},\mathcal{M}} & \mathbb{M}_{\mathcal{E}} \end{pmatrix} \begin{pmatrix} U_{\mathcal{M}} \\ U_{\mathcal{E}} \end{pmatrix} = \begin{pmatrix} S_{\mathcal{M}} \\ S_{\mathcal{E}} \end{pmatrix}, \tag{5.1}$$

where $\mathbb{M}_{\mathcal{M}} \in \mathbb{R}^{|\mathcal{M}|\times|\mathcal{M}|}$, $\mathbb{M}_{\mathcal{M},\mathcal{E}} \in \mathbb{R}^{|\mathcal{M}|\times|\mathcal{E}|}$, $\mathbb{M}_{\mathcal{E},\mathcal{M}} \in \mathbb{R}^{|\mathcal{E}|\times|\mathcal{M}|}$, $\mathbb{M}_{\mathcal{E}} \in \mathbb{R}^{|\mathcal{E}|\times|\mathcal{E}|}$, and $S_{\mathcal{M}}$ and $S_{\mathcal{E}}$ are vectors of size $|\mathcal{E}|$ and $|\mathcal{M}|$ issued from the loading term and the boundary conditions.
By construction, the matrix $\mathbb{M}_{\mathcal{M}}$ is diagonal with non-zero diagonal entries and can therefore be inverted at a very low computational cost. Thus, we can eliminate the cell unknowns, noticing that

$$U_{\mathcal{M}} = \mathbb{M}_{\mathcal{M}}^{-1}\left(S_{\mathcal{M}} - \mathbb{M}_{\mathcal{M},\mathcal{E}} U_{\mathcal{E}}\right). \tag{5.2}$$

Using this relation, one shows that $U_{\mathcal{E}}$ is the solution to the following linear system

$$\left(\mathbb{M}_{\mathcal{E}} - \mathbb{M}_{\mathcal{E},\mathcal{M}}\mathbb{M}_{\mathcal{M}}^{-1}\mathbb{M}_{\mathcal{M},\mathcal{E}}\right) U_{\mathcal{E}} = S_{\mathcal{E}} - \mathbb{M}_{\mathcal{E},\mathcal{M}}\mathbb{M}_{\mathcal{M}}^{-1}S_{\mathcal{M}}, \tag{5.3}$$

where $\mathbb{M}_{\mathcal{D}} = \mathbb{M}_{\mathcal{E}} - \mathbb{M}_{\mathcal{E},\mathcal{M}}\mathbb{M}_{\mathcal{M}}^{-1}\mathbb{M}_{\mathcal{M},\mathcal{E}}$, the Schur complement of the matrix of (5.1), is an invertible matrix of size $|\mathcal{E}|$. In practice, we solve the linear system (5.3), using a LU factorization algorithm, and we use the solution $U_{\mathcal{E}}$ in order to directly compute $U_{\mathcal{M}}$ from the equation (5.2). This method, called *static condensation*, allows one to inverse a system of size $|\mathcal{E}|$ instead of a system of size $|\mathcal{E}| + |\mathcal{M}|$ and to get a substantial decrease of the computational cost.

For mixed Dirichlet-Neumann boundary conditions, we solve the system in the lifted unknown $\underline{u}_{\mathcal{D}}^z = u_{\mathcal{D}} - \underline{u}_{\mathcal{D}}^l$, and we eliminate the rows and columns corresponding to the Dirichlet boundary unknowns, since the associated unknowns are fixed to 0.

### 5.1.2 Exponential fitting scheme: choice of unknown and harmonic averaging

The exponential fitting scheme can be expressed in either the $u$ or $\rho = ue^{\phi}$ variable. In the $\rho$ variable the resulting linear system is symmetric. One can take advantage of this property in order to improve computational efficiency (using for example Choleski factorization or conjugate gradient method). However, the formulation in $\rho$ is also ill-conditioned. Indeed, in our numerical experiments, the ratio between condition numbers of the linear system in $\rho$ and in $u$ exceeds $10^5$. Because of this issue, we choose and recommend to solve the system in the unknown $u$.

In order to implement the exponential fitting scheme, one needs to evaluate averages of the diffusion matrix $\int_{P_{K,\sigma}} \omega\Lambda$. Observe that $\omega(x)/\omega(x_K)$ is of order $e^{h_K\|\nabla\phi\|_{L^\infty(P_{K,\sigma})}}$ on $P_{K,\sigma}$. Therefore with large advection fields, the diffusion problem (3.14) becomes strongly heterogeneous. It is pointed out in [6] that an (empirical) solution to improve robustness to this heterogeneity is to use harmonic averages to approximmate integrals of the diffusion tensor. In the tests of the following subsections, we compare the "classical" exponential fitting scheme (where the integral is approached by standard quadrature) with the "harmonic" one, in which case we choose to use the following approximation

$$\int_{P_{K,\sigma}} \omega\Lambda \approx |\mathcal{E}_{P_{K,\sigma}}||P_{K,\sigma}| \left( \sum_{F\in\mathcal{E}_{P_{K,\sigma}}} \frac{1}{\omega(\overline{x}_F)} \right)^{-1} \Lambda(x_K),$$

where $\mathcal{E}_{P_{K,\sigma}}$ denote the set of faces (or edges if $d = 2$) of $P_{K,\sigma}$ and $\overline{x}_F$ is the barycentre of $F$.

### 5.1.3 Nonlinear scheme and Newton method

The implementation of the nonlinear scheme relies on the following formulation: given $\underline{u}_{\mathcal{D}}^{n-1} \in \underline{V}_{\mathcal{D}}$ positive, we want to solve the nonlinear system $\underline{\mathcal{G}}_{\mathcal{D}}^{n,\delta t}(\underline{u}_{\mathcal{D}}^n) = 0$, where $\underline{\mathcal{G}}_{\mathcal{D}}^{n,\delta t}$ is defined in (3.41), with a step time $\delta t$ instead of $\Delta t$. The resolution of this system relies on a Newton method.

First, one initialises the method with $\max(\underline{u}_{\mathcal{D}}^{n-1}, \varepsilon\underline{1}_{\mathcal{D}}) \in \underline{V}_{\mathcal{D}}$ (where the maximum is taken coordinate by coordinate), in order to avoid potential problems due to the singularity of the log near 0. In order to compute the residues, we perform a static condensation, since the linear systems have the same structure as (5.1). As a stopping criterion, we compare the $l^\infty$ relative norm of the residue with a threshold $tol$. If the method does not converge after $i_{max}$ iterations, we divide the time step by 2 and restart the resolution. When the method converges, one can proceed to the approximation of $\underline{u}_{\mathcal{D}}^{n+1}$, with an initial time step of $\min(\Delta t, 2\delta t)$.
The initial time step (used to compute $\underline{u}_{\mathcal{D}}^1$) is $\Delta t$. In practice, we use $\varepsilon = 10^{-11}$, $i_{max} = 50$ and $tol = 10^{-11}$.

The implementation of the non linear scheme relies on the computation of $\log(\omega_K)$ and $\log(\omega_\sigma)$. Since we use meshes such that $x_K$ is the barycenter of $K$, we choose to approximate $\frac{1}{|K|}\int_K \mathrm{e}^{-\phi}$ by $\mathrm{e}^{-\phi(x_K)}$. Therefore, $\log(\omega_K)$ is computed as $\log(\mathrm{e}^{-\phi(x_K)}) = -\phi(x_K)$. The same holds for $\log(\omega_\sigma)$.

In the simulation showed below, we use the arithmetic means for the reconstruction defined in (3.28) - (3.29). Therefore, for $\underline{u} \in \underline{V}$ and $K \in \mathcal{M}$,

$$r_K(\underline{u}_K) = \frac{1}{2}\left(u_K + \frac{1}{|\mathcal{E}_K|}\sum_{\sigma\in\mathcal{E}_K} u_\sigma\right).$$

For a discussion on other choices of reconstructions we refer to [8, Section 6.2].

### 5.2 Long-time behaviour of discrete solutions

In this subsection, we present some numerical illustrations of the long-time behaviour of discrete solutions.

We start with a test case used in [10, 9, 8]. We consider homogeneous Neumann boundary conditions ($\Gamma^D = \emptyset$ and $g^N = 0$) and no loading term ($f = 0$). The potential and diffusion tensor are set to $\phi(x, y) = -x$ and $\Lambda = \begin{pmatrix} l_x & 0 \\ 0 & 1 \end{pmatrix}$. The exact solution is defined by

$$u(t, x, y) = C_1\, \mathrm{e}^{-\alpha t + \frac{x}{2}}\,(2\pi\cos(\pi x) + \sin(\pi x)) + 2C_1\pi\, \mathrm{e}^{x-\frac{1}{2}},$$

where $C_1 > 0$ and $\alpha = l_x\left(\frac{1}{4} + \pi^2\right)$. Note that $u^{in}$ vanishes on $\{x = 1\}$, but for any $t > 0$, $u(t, \cdot) > 0$. The steady sate is

$$u^\infty(x, y) = 2C_1\pi\, \mathrm{e}^{x-\frac{1}{2}}.$$

The experiments are performed using the following values:

$$l_x = 10^{-2} \text{ and } C_1 = 10^{-1}.$$

We compute the solution on the time interval $[0, T_f]$ and we denote by $(\underline{u}_{\mathcal{D}}^n)_{1\le n\le N_f+1}$ the corresponding approximate solution. Note that the number of time steps $N_f$ may differ between linear and nonlinear schemes, because of the adaptive time step refinement procedure for the nonlinear scheme.

We set $T_f = 350$ in order to see the complete evolution. Since the long-time behaviour of the schemes does not depend on the size of the discretisation, we can explore the evolution using a large time step, $\Delta t = 10^{-1}$. We perform the numerical experiments on two Kershaw meshes (see Figure 2a) of sizes 0.02 and 0.006. On Figure 3, we show the $L^1$ distance between $\underline{u}_D^n$ and $u^\infty$ (the real steady state) computed as

$$\sum_{K\in\mathcal{M}} |K||u_K^n - u^\infty(x_K)|.$$

We observe the exponential decay towards the steady state, until some precision is reached. Note that for the HMM scheme, some saturation occurs at precision of magnitude $10^{-6}$ and $10^{-7}$ : the scheme does not preserve the thermal equilibrium (see Remark 3). This saturation corresponds to the accuracy of the stationary scheme (see next subsection), so the threshold is lower on the refined mesh. Note that one could also consider $\sum_{K\in\mathcal{M}} |K||u_K^n - u_K^\infty|$, which measures the error between the discrete solution and the discrete steady state : this quantity decays exponentially, with a lower saturation of magnitude $10^{-12}$. The other schemes (nonlinear and exponential fitting) have the same decay rate, and the saturation occurs at the machine precision. For the four schemes, the rates of convergence observed are similar to the real one. Especially, the use of harmonic averages in the exponential fitting scheme does not have impact on the long-time behaviour.

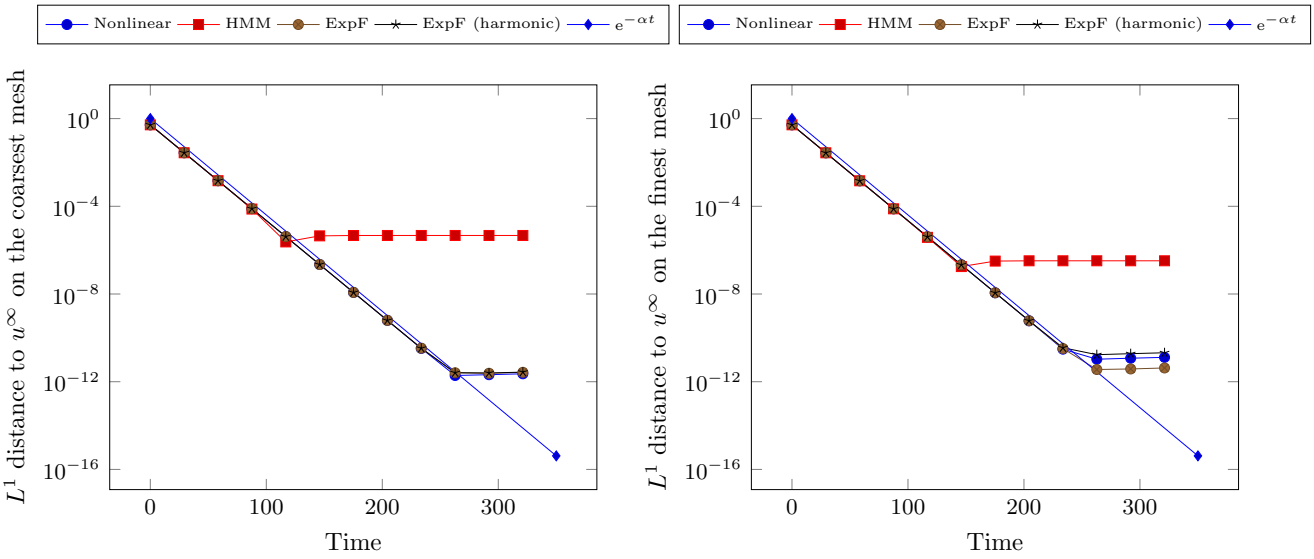



Figure 3: **Long-time behaviour of discrete solutions.** Comparison of long-time behaviour on Kershaw meshes, for $T_f = 350$ and $\Delta t = 0.1$.

### 5.3 Positivity of discrete solutions

We are now interested in the positivity of the discrete solutions. Here, we use a test case with anisotropic diffusion, and homogeneous Neumann boundary conditions. We set $\Gamma^D = \emptyset$, $f = 0$, $g^N = 0$,

$$\phi(x,y) = -\left((x-0.4)^2 + (y-0.6)^2\right) \text{ and } \Lambda = \begin{pmatrix} 0.8 & 0 \\ 0 & 1 \end{pmatrix}.$$

For the initial data, we take

$$u^{in} = 10^{-3}.\mathbb{1}_B + \mathbb{1}_{\Omega\setminus B},$$

where $B$ is the Euclidean ball $\left\{(x,y) \in \mathbb{R}^2 \mid (x-0.5)^2 + (y-0.5)^2 \leq 0.2^2\right\}$. These data ensure that the solution $u$ is positive on $\mathbb{R}_+ \times \Omega$. The experiment is performed on a tilted hexagonal mesh (see Figure 2d) of size $4.3 \cdot 10^{-3}$, made up of 4192 cells and 12512 edges. Since we deal with a diffusive phenomenon, the smallest values of $u$ are observed for a small time, so that we perform the simulation with a relatively small final time $T_f = 5.10^{-4}$, alongside with a time step of $\Delta t = 10^{-5}$.

The results are presented on Table 1. The cost is defined as the number of linear systems solved in order to compute the solution $(\underline{u}^n_{\mathcal{D}})_{1\leq n\leq N_f+1}$, and the minimum values mincells and minfaces are defined as

$$\min\left\{u^n_K \mid 1 \leq n \leq N_f+1,\ K \in \mathcal{M}\right\} \text{ and } \min\left\{u^n_\sigma \mid 1 \leq n \leq N_f+1,\ \sigma \in \mathcal{E}\right\}.$$

As expected, the nonlinear scheme has positive discrete solutions, whereas the linear ones exhibit

| | cost | mincells | minfaces | # negative unknowns |
|---|---|---|---|---|
| Nonlinear | 175 | 9.93e-04 | 7.36e-04 | 0 |
| HMM | 50 | -5e-03 | –7.74e-02 | 5.93e+02 |
| ExpF | 50 | -4.98e-03 | -7.72e-02 | 5.90e+02 |
| ExpF (harmonic) | 50 | -4.98e-03 | -7.74e-02 | 5.88e+02 |

Table 1: **Positivity of discrete solutions.** Numerical results for $T_f = 5.10^{-4}$, $\Delta t = 10^{-5}$ on a tilted hexagonal mesh. At each time step, there are 4192 cells unknowns and 12512 faces unknowns.

a default of positivity (the value of $\eta$ seems to have some influence over the positivity, see [24]). Note that the use of harmonic averages for the exponential fitting scheme has no impact on the undershoots.

We observe that the nonlinear scheme requires approximately 3.5 times more linear system inversions than the linear schemes. However, this value depends strongly on the final time of simulation $T_f$. Indeed, the number of linear systems solved at the step $n$ decreases when $n$ increases. The first iteration costs 9 resolutions, but this number rapidly decreases as the solution approaches the steady state (the second and the third iterations respectively cost 5 and 4 resolutions).

### 5.4 Accuracy of stationary solutions

In this subsection, we aim to exhibit the convergence rates of the different schemes for the stationary problem, in order to compare their respective accuracy. To do so, we define the $L^2$ and $H^1$ errors to be the corresponding norm of the difference between the discrete solution $\underline{u}_\mathcal{D}$ and the projection of the continuous solution $\underline{\Pi}_\mathcal{D} u := ((u(x_K))_K, (u(\overline{x}_\sigma))_\sigma)$ divided by the norm of $\underline{\Pi}_\mathcal{D} u$.

The nonlinear scheme is extended to a more general setting in order to take into account a loading term $f > 0$ and mixed Dirichlet-Neumann boundary conditions with $g^D > 0$ and $g^N \geq 0$. The scheme writes

$$\text{find } \underline{u}_\mathcal{D} \in \underline{V}_\mathcal{D} \text{ positive such that } \underline{\mathcal{G}}_\mathcal{D}(\underline{u}_\mathcal{D}) = 0, \tag{5.4}$$

where $\underline{\mathcal{G}}_\mathcal{D} : (\mathbb{R}_+^*)^{|\mathcal{M}|+|\mathcal{E}|} \to \underline{V}_\mathcal{D}$ is the vector field defined by

$$\forall K \in \mathcal{M},\ \mathcal{G}_K(\underline{u}_\mathcal{D}) = \sum_{\sigma \in \mathcal{E}_K} \mathcal{F}_{K,\sigma}(\underline{u}_K) - \int_K f, \tag{5.5a}$$

$$\forall \sigma = K|L \in \mathcal{E}_{int},\ \mathcal{G}_\sigma(\underline{u}_\mathcal{D}) = -\left(\mathcal{F}_{K,\sigma}(\underline{u}_K) + \mathcal{F}_{L,\sigma}(\underline{u}_K)\right), \tag{5.5b}$$

$$\forall \sigma \in \mathcal{E}_{ext}^N,\ \mathcal{G}_\sigma(\underline{u}_\mathcal{D}) = \int_\sigma g^N - \mathcal{F}_{K,\sigma}(\underline{u}_K), \tag{5.5c}$$

$$\forall \sigma \in \mathcal{E}_{ext}^D,\ \mathcal{G}_\sigma(\underline{u}_\mathcal{D}) = \frac{1}{|\sigma|}\int_\sigma g^D - u_\sigma, \tag{5.5d}$$

$$\forall K \in \mathcal{M}, \forall \sigma \in \mathcal{E}_K,\ \mathcal{F}_{K,\sigma}(\underline{u}_K) = r^K(\underline{u}_K) \sum_{\sigma' \in \mathcal{E}_K} A_K^{\sigma\sigma'} \left(\log(u_K) + \phi(x_K) - \log(u_\sigma) - \phi(x_\sigma)\right). \tag{5.5e}$$

The implementation of this scheme still relies on a Newton method similar to the one used for the evolutionary scheme. It is initialised with a discretisation of the thermal equilibrium $\mathrm{e}^{-\phi}$.

The first test case is the same as in [4]. It is an isotropic problem, with $\Lambda = I$, $\Gamma^N = \emptyset$, $\phi(x, y) = -(2x + 3y)$ and the exact solution is

$$u(x, y) = \left(x - \mathrm{e}^{2(x-1)}\right)\left(y - \mathrm{e}^{3(y-1)}\right),$$

the other data are set according to this exact solution. Note that for this test case, the diffusion and advection terms are of same order. The numerical experiments are performed on triangular meshes.

In Figure 4, we show the convergence results. As expected, the two linear schemes are of order two in $L^2$ norm, and one in $H^1$ norm. The same holds for the nonlinear scheme, whose accuracy seems to be rather the same as the classical HMM scheme. On this test case, the use of harmonic averages for the exponential fitting scheme does not have a significant impact.

The second test case is a advection-dominated problem, with anisotropic diffusion and mixed Dirichlet-Neumann boundary conditions. We set $\Gamma^D = (\{0\} \times [0, 1]) \cup (\{1\} \times [0, 1])$, $\Gamma^N = ([0, 1] \times \{0\}) \cup ([0, 1] \times \{1\})$, $g^D = 1$, $g^N = 0$ and $f = 0$. The diffusion tensor and the potential are defined

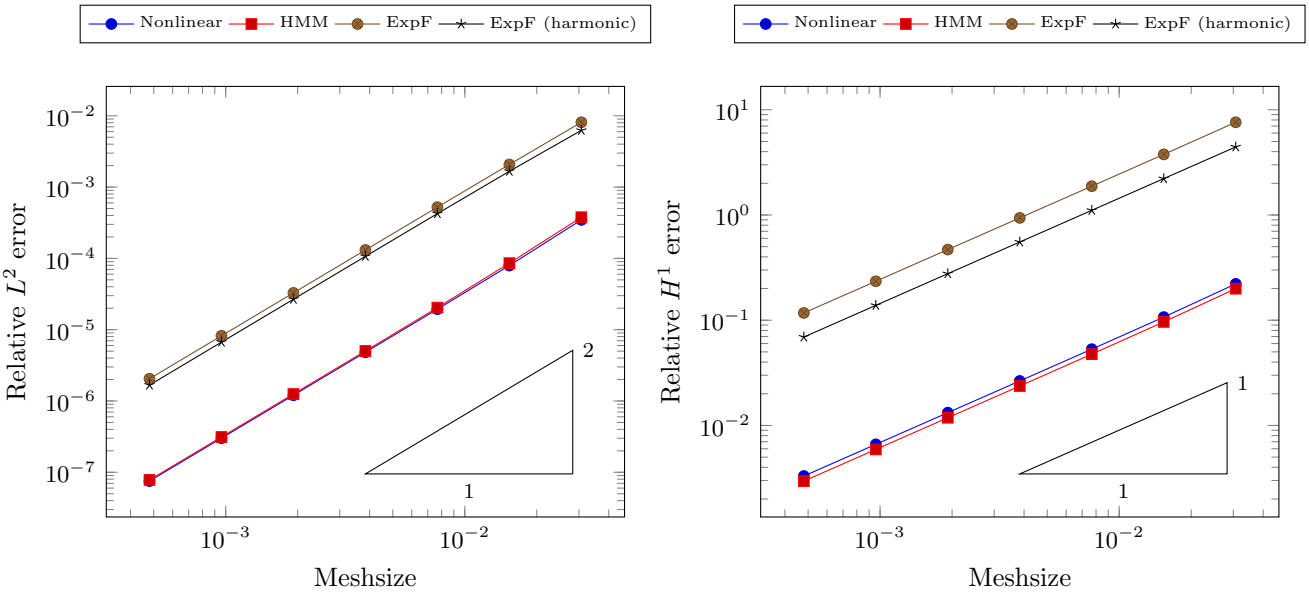



Figure 4: **Accuracy of stationary solutions.** Errors in $L^2$ and $H^1$ norms for the first test case, on triangular meshes.

by the following expressions

$$\Lambda = \begin{pmatrix} 1 & 0 \\ 0 & l_y \end{pmatrix} \text{ and } \phi(x, y) = \log\left(\frac{1}{v} + x\right),$$

with $v > 0$. Note that the advection field $-\Lambda\nabla\phi = \begin{pmatrix} \frac{v}{1+vx} \\ 0 \end{pmatrix}$ has a norm of order $v$ when $x$ is small. Thus, near $\{0\} \times [0, 1]$, the problem is advection dominated if $v$ is big enough. Moreover, $-\operatorname{div}(\Lambda\nabla\phi) = \dfrac{v^2}{(1+vx)^2} > 0$, so the problem is coercive. The explicit solution is given by

$$u(x, y) = \frac{v}{1+vx}\left(\frac{2vx}{2+v}\left(\frac{1}{v} + \frac{x}{2}\right) + \frac{1}{v}\right).$$

We perform the numerical experiments on Cartesian meshes, with

$$l_y = 100 \text{ and } v = 200.$$

The results are presented on Figure 5. They show that the HMM scheme exhibits an order of convergence inferior to one in $H^1$ norm and two in $L^2$ norm. This deteriorated convergence appears because the advective term predominates over the diffusive term, at least in some part of the domain. The other schemes are of order one in $H^1$ norm and two in $L^2$ norm. Moreover, on this test case, their effective accuracy is better than that of the HMM scheme on refined meshes.

Note that, by the expression of the potential, $\omega = \mathrm{e}^{-\phi} = \frac{v}{1+vx}$ will have small variations on the cells even if $v$ is big. Therefore, the diffusion tensor $\omega\Lambda$ of the problem in the $\rho$ unknown for the exponential fitting schemes is not too heterogeneous (locally). It could explain the performance of the exponential fitting schemes in this case. Moreover, on this test case, using harmonic averages in the exponential fitting scheme gives a substantial gain of accuracy for both $L^2$ and $H^1$ relative errors, of magnitude $10^1$.

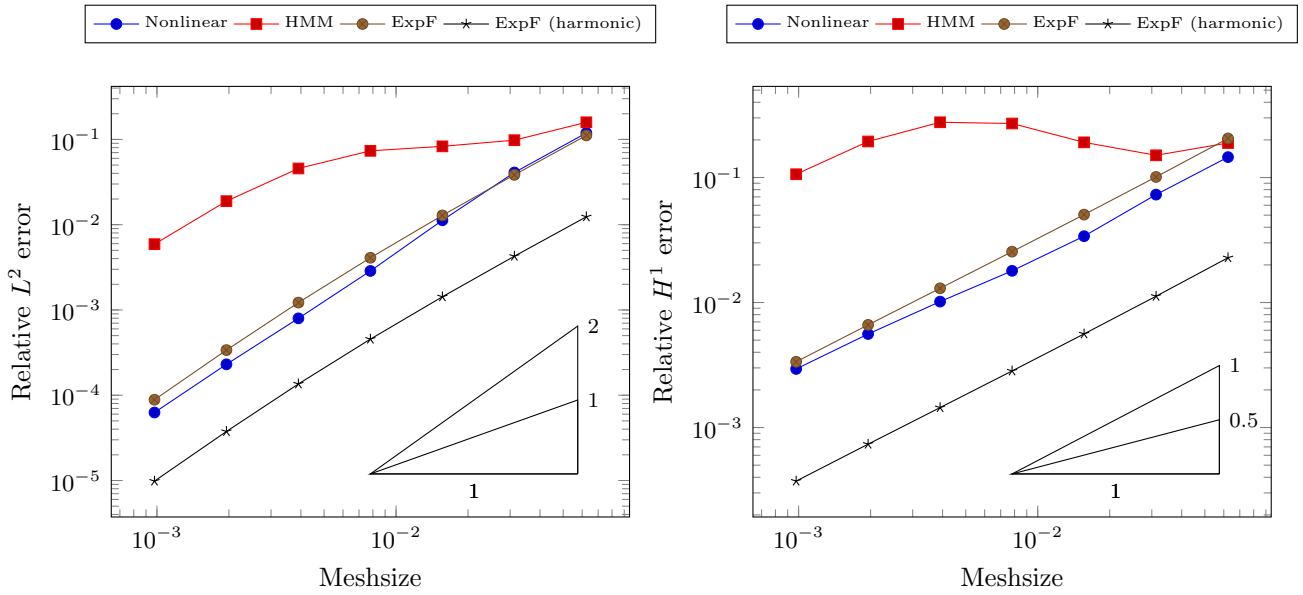



Figure 5: **Accuracy of stationary solutions.** Errors in $L^2$ and $H^1$ norms for the second test case, on Cartesian meshes.

# 6 Conclusion

In this paper, by means of discrete entropy methods, we have analysed the long-time behaviour of three hybrid finite volume schemes for linear advection-diffusion equations. We have proved that the solutions to all schemes converge exponentially fast in time towards the associated discrete steady-states. Two schemes among the three are new, that are the (linear) exponential fitting scheme (adapting known ideas to the HFV context) and the nonlinear scheme, for which we have proved the existence of solutions. All schemes can handle anisotropy and general meshes. The two linear schemes can deal with general data and mixed Dirichlet-Neumann boundary conditions, however they do not preserve the positivity of solutions. On the other hand, the nonlinear scheme preserves positivity and can be used in practice with general boundary conditions. However, at the moment, its analysis is limited to systems that converge in time towards the thermal equilibrium, restricting the admissible data. We have finally validated our theoretical findings on different numerical tests, assessing positivity, long-time behaviour, and spatial accuracy of the schemes.

# A Discrete functional inequalities

## A.1 Discrete Poincaré inequalities

We recall the hybrid discrete Poincaré inequalities (cf. [20, Lemmas B.25 and B.32, $p = 2$]).

**Proposition 7** (Discrete Poincaré inequalities)**.** *Let $\mathcal{D}$ be a given discretisation of $\Omega$, with regularity parameter $\theta_\mathcal{D}$. There exists $C_{PW} > 0$, only depending on $\Omega$, $d$, and $\theta_\mathcal{D}$ such that*

$$\forall \underline{v}_\mathcal{D} \in \underline{V}^N_{\mathcal{D},0}, \qquad \|v_\mathcal{M}\|_{L^2(\Omega)} \le C_{PW} |\underline{v}_\mathcal{D}|_{1,\mathcal{D}}. \tag{A.1}$$

*Assume that $|\Gamma^D| > 0$. Then, there exists $C_{P,\Gamma^D} > 0$, only depending on $\Omega$, $d$, $\Gamma^D$, and $\theta_\mathcal{D}$ such that*

$$\forall \underline{v}_\mathcal{D} \in \underline{V}^D_{\mathcal{D},0}, \qquad \|v_\mathcal{M}\|_{L^2(\Omega)} \le C_{P,\Gamma^D} |\underline{v}_\mathcal{D}|_{1,\mathcal{D}}. \tag{A.2}$$

### A.2 A discrete logarithmic Sobolev inequality

Adapting to the hybrid setting the arguments of the proof of [8, Proposition 5.3] (in turn based on the proof of [8, Theorem 5.1]), we prove a hybrid discrete log-Sobolev inequality.

**Proposition 8** (Discrete logarithmic Sobolev inequality)**.** *Let $\mathcal{D}$ be a given discretisation of $\Omega$, with regularity parameter $\theta_{\mathcal{D}}$. Let $\underline{v}_{\mathcal{D}}, \underline{v}_{\mathcal{D}}^{\infty} \in \underline{V}_{\mathcal{D}}$ be two positive vectors of unknowns such that*

$$\int_\Omega v_{\mathcal{M}} = \int_\Omega v_{\mathcal{M}}^{\infty} =: M,$$

*and set $v_{\mathcal{M},\sharp}^{\infty} := \sup_{K\in\mathcal{M}} v_K^{\infty}$. Define $\underline{\xi}_{\mathcal{D}}$ as the element of $\underline{V}_{\mathcal{D}}$ such that*

$$\xi_K := \sqrt{\frac{v_K}{v_K^{\infty}}} \quad \forall K \in \mathcal{M}, \qquad \xi_\sigma := \sqrt{\frac{v_\sigma}{v_\sigma^{\infty}}} \quad \forall \sigma \in \mathcal{E}.$$

*There exists $C_{LS} > 0$, only depending on $\Omega$, $d$, and $\theta_{\mathcal{D}}$ such that*

$$\int_\Omega v_{\mathcal{M}}^{\infty} \Phi_1\left(\frac{v_{\mathcal{M}}}{v_{\mathcal{M}}^{\infty}}\right) \le C_{LS}^2 \sqrt{M v_{\mathcal{M},\sharp}^{\infty}}\, |\underline{\xi}_{\mathcal{D}}|_{1,\mathcal{D}}^2, \tag{A.3}$$

*where we recall that $\Phi_1(s) = s\log(s) - s + 1$ for all $s > 0$.*

*Proof.* By [8, Eq. (5.7)] applied to the probability measure $\mu(x)\,\mathrm{d}x = v_{\mathcal{M}}^{\infty}(x)\,\frac{\mathrm{d}x}{M}$ and the function $f = \sqrt{\frac{v_{\mathcal{M}}}{v_{\mathcal{M}}^{\infty}}} = \xi_{\mathcal{M}}$, we first infer that

$$\int_\Omega v_{\mathcal{M}} \log\left(\frac{v_{\mathcal{M}}}{v_{\mathcal{M}}^{\infty}}\right) \le 8\sqrt{M v_{\mathcal{M},\sharp}^{\infty}}\, \|\xi_{\mathcal{M}} - \overline{\xi_{\mathcal{M}}}\|_{L^4(\Omega)}^2,$$

where we let $\overline{\xi_{\mathcal{M}}} := \frac{1}{|\Omega|}\int_\Omega \xi_{\mathcal{M}}$. The conclusion then falls in two steps. On the one hand, since $\int_\Omega v_{\mathcal{M}} = \int_\Omega v_{\mathcal{M}}^{\infty}$, we remark that

$$\int_\Omega v_{\mathcal{M}}^{\infty} \Phi_1\left(\frac{v_{\mathcal{M}}}{v_{\mathcal{M}}^{\infty}}\right) = \int_\Omega v_{\mathcal{M}} \log\left(\frac{v_{\mathcal{M}}}{v_{\mathcal{M}}^{\infty}}\right).$$

On the other hand, we invoke the following hybrid discrete Poincaré–Sobolev inequality (cf. [20, Lemma B.25, $p = 2$, $q = 4$]):

$$\forall \underline{w}_{\mathcal{D}} \in \underline{V}_{\mathcal{D},0}^N, \qquad \|w_{\mathcal{M}}\|_{L^4(\Omega)} \le C_{PS} |\underline{w}_{\mathcal{D}}|_{1,\mathcal{D}},$$

where $C_{PS} > 0$ only depends on $\Omega$, $d$, and $\theta_{\mathcal{D}}$, that we apply to $\underline{w}_{\mathcal{D}} = \underline{\xi}_{\mathcal{D}} - \overline{\xi_{\mathcal{M}}}\,\underline{1}_{\mathcal{D}} \in \underline{V}_{\mathcal{D},0}^N$. This proves (A.3) with $C_{LS} = 2\sqrt{2}\,C_{PS}$. □

## B Proofs of technical results

### B.1 Discrete boundedness by mass and dissipation

We prove Lemma 2 from Section 3.3.2. To ease the reading, we first recall the result.

**Lemma 2.** *Let $\underline{w}_{\mathcal{D}} \in \underline{V}_{\mathcal{D}}$, and assume that there exist $C_\sharp \ge 0$, and $M_\sharp \ge M_\flat > 0$ such that*

$$M_\flat \le \sum_{K\in\mathcal{M}} |K| u_K^{\infty}\, \mathrm{e}^{w_K} \le M_\sharp \qquad \text{and} \qquad \mathbb{D}(\underline{w}_{\mathcal{D}}) \le C_\sharp. \tag{3.39}$$

*Then, there exists $C > 0$, depending on $\Lambda$, $u_\flat^{\infty}$, $u_\sharp^{\infty}$, $M_\flat$, $M_\sharp$, $C_\sharp$, $\Omega$, $d$, and $\mathcal{D}$ such that*

$$|w_K| \le C \quad \forall K \in \mathcal{M} \quad \text{and} \quad |w_\sigma| \le C \quad \forall \sigma \in \mathcal{E}.$$

*Proof.* For $K \in \mathcal{M}$, using (2.10) and (2.11), we first infer that

$$\delta_K \underline{w}_K \cdot \mathbb{A}_K \delta_K \underline{w}_K = a_K^\Lambda(\underline{w}_K, \underline{w}_K) \geq \lambda_\flat \alpha_\flat |\underline{w}_K|_{1,K}^2 = \lambda_\flat \alpha_\flat \sum_{\sigma \in \mathcal{E}_K} \frac{|\sigma|}{d_{K,\sigma}} \big(w_K - w_\sigma\big)^2.$$

By definition (2.1) of the regularity parameter $\theta_\mathcal{D}$, we have that $\frac{|\sigma|}{d_{K,\sigma}} \geq \frac{h_K^{d-2}}{\theta_\mathcal{D}}$ for all $\sigma \in \mathcal{E}_K$, so that

$$\delta_K \underline{w}_K \cdot \mathbb{A}_K \delta_K \underline{w}_K \geq \frac{\lambda_\flat \alpha_\flat}{\theta_\mathcal{D}} h_K^{d-2} |\delta_K \underline{w}_K|^2. \tag{B.1}$$

By the expression (3.38) of $\mathbb{D}(\underline{w}_\mathcal{D})$, and the local lower bound (B.1), we thus get

$$\begin{aligned}
\mathbb{D}(\underline{w}_\mathcal{D}) &= \sum_{K \in \mathcal{M}} r_K\big(\underline{u}_K^\infty \times \exp(\underline{w}_K)\big) \delta_K \underline{w}_K \cdot \mathbb{A}_K \delta_K \underline{w}_K \\
&\geq \frac{\lambda_\flat \alpha_\flat}{\theta_\mathcal{D}} \sum_{K \in \mathcal{M}} h_K^{d-2} r_K\big(\underline{u}_K^\infty \times \exp(\underline{w}_K)\big) |\delta_K \underline{w}_K|^2 \\
&= \frac{\lambda_\flat \alpha_\flat}{\theta_\mathcal{D}} \sum_{K \in \mathcal{M}} \sum_{\sigma \in \mathcal{E}_K} h_K^{d-2} r_K\big(\underline{u}_K^\infty \times \exp(\underline{w}_K)\big) (w_K - w_\sigma)^2.
\end{aligned}$$

Let $K \in \mathcal{M}$ and $\sigma \in \mathcal{E}_K$ be fixed. Using, successively, the definition (3.27) of $r_K$ combined with the definition (3.29) of $f_{|\mathcal{E}_K|}$, the combination of (3.34) with assumptions (3.28a) and (3.28c), and the assumptions (3.28b) and (3.28d) combined with the bound (2.2) on $|\mathcal{E}_K|$, we infer, for $w_\sigma \neq w_K$,

$$\begin{aligned}
r_K\big(\underline{u}_K^\infty \times \exp(\underline{w}_K)\big)(w_K - w_\sigma)^2 &\geq \frac{1}{|\mathcal{E}_K|} m\big(u_K^\infty \,\mathrm{e}^{w_K}, u_\sigma^\infty \,\mathrm{e}^{w_\sigma}\big)(w_K - w_\sigma)^2 \\
&\geq \frac{u_\flat^\infty}{|\mathcal{E}_K|} m\big(\,\mathrm{e}^{w_K}, \mathrm{e}^{w_\sigma}\big)(w_K - w_\sigma)^2 \\
&\geq \frac{u_\flat^\infty}{d\theta_\mathcal{D}^2} \left(\mathrm{e}^{w_K} - \mathrm{e}^{w_\sigma}\right)(w_K - w_\sigma) \geq 0,
\end{aligned}$$

and we verify that this inequality still holds when $w_\sigma = w_K$. Since $\mathbb{D}(\underline{w}_\mathcal{D}) \leq C_\sharp$ by (3.39), for all $K \in \mathcal{M}$, and all $\sigma \in \mathcal{E}_K$, we have

$$0 \leq \left(\mathrm{e}^{w_K} - \mathrm{e}^{w_\sigma}\right)(w_K - w_\sigma) \leq \zeta h_K^{2-d}, \tag{B.2}$$

with $\zeta := \frac{dC_\sharp \theta_\mathcal{D}^3}{\lambda_\flat \alpha_\flat u_\flat^\infty} \geq 0$ (recall that $\alpha_\flat$ depends on $\Omega$, $d$, and $\theta_\mathcal{D}$). Besides, since $\sum_{K \in \mathcal{M}} |K| u_K^\infty \,\mathrm{e}^{w_K} \leq M_\sharp$ again by (3.39), for all $K \in \mathcal{M}$, we have $|K| u_K^\infty \,\mathrm{e}^{w_K} \leq M_\sharp$ for all $K \in \mathcal{M}$. Similarly, since $\sum_{K \in \mathcal{M}} |K| u_K^\infty \,\mathrm{e}^{w_K} \geq M_\flat$, there exists $K_0 \in \mathcal{M}$ such that $|\Omega| u_{K_0}^\infty \,\mathrm{e}^{w_{K_0}} \geq M_\flat$. Combining these bounds, we infer that there exists $K_0 \in \mathcal{M}$ such that

$$\log\left(\frac{M_\flat}{|\Omega| u_\sharp^\infty}\right) \leq w_{K_0} \leq \log\left(\frac{M_\sharp}{|K_0| u_\flat^\infty}\right). \tag{B.3}$$

Now let us show that we can similarly frame all the other components of $\underline{w}_\mathcal{D}$.

For $a, x \in \mathbb{R}$, let us define $E(a,x) = \left(\mathrm{e}^x - \mathrm{e}^a\right)(x - a)$. Observe that $E(a, x+a)e^{-a} = (e^x - 1)x =: \xi(x)$ and that $\xi$ is strictly convex, minimal at $x = 0$ and $\xi(x) \to +\infty$ when $x \to \pm\infty$. Let $b, a_\sharp > 0$ and take $|a| < a_\sharp$. By the properties of $\xi$, if $E(a,x) \leq b$ then $|x| \leq \kappa_b(a_\sharp) := a_\sharp + \max\{|x| \text{ s.t. } 0 \leq \xi(x) \leq be^{a_\sharp}\}$. We can thus infer that if $(x_k)_{k=0\ldots m}$ is a finite sequence of real numbers such that $E(x_k, x_{k+1}) \leq b$ and $|x_0| < a_\sharp$, then $|x_m| \leq \kappa_b^{(m)}(a_\sharp)$ where $\kappa_b^{(m)}$ is $m$ compositions of $\kappa_b$. In particular the bound depends only on $a_\sharp$, $m$ and $b$.

Now we can conclude. Because of the connectivity of the mesh, for any cell $K$ (*resp.* face $\sigma$) there is a finite sequence of components of $\underline{w}_{\mathcal{D}}$, denoted $(x_k)_{k=0\dots m}$, starting at $x_0 = w_{K_0}$ and finishing at $x_m = w_K$ (*resp.* $x_m = w_\sigma$) such that, by (B.2), $E(x_k, x_{k+1}) \le b := \zeta h_{\mathcal{D}}^{2-d}$. The inequality (B.3) yields the initial bound on $|x_0|$, and one concludes by the above argument.

□

## B.2 A local comparison result

We prove a local comparison result between the matrices $\mathbb{A}_K$ and some (local) diagonal matrices. The proof relies on arguments that are similar to those advocated in [10] to analyse the VAG scheme.

**Lemma 3.** *For $K \in \mathcal{M}$, let $\mathbb{A}_K \in \mathbb{R}^{|\mathcal{E}_K|\times|\mathcal{E}_K|}$ be the matrix defined by (2.8). The matrices $\mathbb{A}_K$ are symmetric positive-definite, and there exists $C_A > 0$, only depending on $\Lambda$, $\Omega$, $d$, and $\theta_{\mathcal{D}}$ such that*

$$\forall K \in \mathcal{M}, \quad \mathrm{Cond}_2(\mathbb{A}_K) \le C_A,$$

*where $\mathrm{Cond}_2(\mathbb{A}_K) := \|\mathbb{A}_K^{-1}\|_2\|\mathbb{A}_K\|_2$ is the condition number of the matrix $\mathbb{A}_K$. Moreover, letting for $K \in \mathcal{M}$, $\mathbb{B}_K \in \mathbb{R}^{|\mathcal{E}_K|\times|\mathcal{E}_K|}$ be the diagonal matrix with entries*

$$B_K^{\sigma\sigma} := \sum_{\sigma'\in\mathcal{E}_K} |A_K^{\sigma\sigma'}| \qquad \text{for all } \sigma \in \mathcal{E}_K, \tag{B.4}$$

*there exists $C_B > 0$, only depending on $\Lambda$, $\Omega$, $d$, and $\theta_{\mathcal{D}}$ such that*

$$\forall K \in \mathcal{M}, \forall w \in \mathbb{R}^{|\mathcal{E}_K|}, \qquad w \cdot \mathbb{A}_K w \le w \cdot \mathbb{B}_K w \le C_B\, w \cdot \mathbb{A}_K w. \tag{B.5}$$

*Proof.* Let $K \in \mathcal{M}$ and $k := |\mathcal{E}_K|$. As direct consequence of its definition (2.8), the matrix $\mathbb{A}_K \in \mathbb{R}^{k\times k}$ is symmetric and positive semi-definite. Now, let $w := (w_\sigma)_{\sigma\in\mathcal{E}_K} \in \mathbb{R}^k$, and define $\underline{v}_K \in \underline{V}_K$ such that

$$v_K = 0 \qquad \text{and} \qquad v_\sigma = -w_\sigma \text{ for all } \sigma \in \mathcal{E}_K.$$

Then, $\delta_K \underline{v}_K = (v_K - v_\sigma)_{\sigma\in\mathcal{E}_K} = w$. By (B.1), we immediately get that

$$w \cdot \mathbb{A}_K w \ge \frac{\lambda_\flat \alpha_\flat}{\theta_{\mathcal{D}}} h_K^{d-2} |w|^2,$$

which implies, since $w \in \mathbb{R}^k$ is arbitrary, that $\mathbb{A}_K$ is invertible, and gives us a lower bound on its smallest eigenvalue. By the same arguments advocated to prove (B.1), noticing that $\frac{|\sigma|}{d_{K,\sigma}} \le \theta_{\mathcal{D}} h_K^{d-2}$ for all $\sigma \in \mathcal{E}_K$, we infer that

$$w \cdot \mathbb{A}_K w \le \lambda_\sharp \alpha_\sharp \theta_{\mathcal{D}} h_K^{d-2} |w|^2.$$

We eventually get, using the estimates on the eigenvalues of $\mathbb{A}_K$, that

$$\mathrm{Cond}_2(\mathbb{A}_K) \le \frac{\lambda_\sharp \alpha_\sharp}{\lambda_\flat \alpha_\flat} \theta_{\mathcal{D}}^2 =: C_A, \tag{B.6}$$

with $C_A > 0$ only depending on $\Lambda$, $\Omega$, $d$, and $\theta_{\mathcal{D}}$. Now, by (B.4), since $\mathbb{A}_K$ is symmetric, we have

$$w \cdot \mathbb{B}_K w = \sum_{\sigma\in\mathcal{E}_K} \sum_{\sigma'\in\mathcal{E}_K} |A_K^{\sigma\sigma'}| w_\sigma^2 = \sum_{\sigma\in\mathcal{E}_K} \sum_{\sigma'\in\mathcal{E}_K} |A_K^{\sigma\sigma'}| w_{\sigma'}^2,$$

and we can use the half-sum to get

$$w \cdot \mathbb{B}_K w = \sum_{\sigma\in\mathcal{E}_K} \sum_{\sigma'\in\mathcal{E}_K} |A_K^{\sigma\sigma'}| \frac{w_\sigma^2 + w_{\sigma'}^2}{2}.$$

Using Young's inequality, we infer

$$w\cdot\mathbb{A}_K w=\sum_{\sigma\in\mathcal{E}_K}\sum_{\sigma'\in\mathcal{E}_K}A_K^{\sigma\sigma'}w_\sigma w_{\sigma'}\le\sum_{\sigma\in\mathcal{E}_K}\sum_{\sigma'\in\mathcal{E}_K}|A_K^{\sigma\sigma'}||w_\sigma||w_{\sigma'}|$$
$$\le\sum_{\sigma\in\mathcal{E}_K}\sum_{\sigma'\in\mathcal{E}_K}|A_K^{\sigma\sigma'}|\frac{w_\sigma^2+w_{\sigma'}^2}{2}=w\cdot\mathbb{B}_K w.$$

For the second inequality, by symmetry of $\mathbb{A}_K$, we have

$$w\cdot\mathbb{B}_K w=\sum_{\sigma\in\mathcal{E}_K}B_K^{\sigma\sigma}w_\sigma^2\le\max_{\sigma\in\mathcal{E}_K}(B_K^{\sigma\sigma})\sum_{\sigma\in\mathcal{E}_K}w_\sigma^2=\max_{\sigma\in\mathcal{E}_K}\left(\sum_{\sigma'\in\mathcal{E}_K}|A_K^{\sigma'\sigma}|\right)|w|^2=\|\mathbb{A}_K\|_1|w|^2.$$

The space $\mathbb{R}^{k\times k}$ being of finite dimension, the norms $\|\cdot\|_1$ and $\|\cdot\|_2$ are equivalent, and there exists $\gamma_k>0$ such that $\|\cdot\|_1\le\gamma_k\|\cdot\|_2$. Moreover, since $\mathbb{A}_K$ is (symmetric) positive-definite, the following inequality holds:

$$w\cdot\mathbb{A}_K w\ge\frac{\|\mathbb{A}_K\|_2}{\mathrm{Cond}_2(\mathbb{A}_K)}|w|^2.$$

From the previous estimates and (B.6), we deduce that

$$w\cdot\mathbb{B}_K w\le\gamma_k\,\mathrm{Cond}_2(\mathbb{A}_K)\,w\cdot\mathbb{A}_K w\le\gamma_k\,C_A\,w\cdot\mathbb{A}_K w.$$

But, according to (2.2), we have $\max_{K\in\mathcal{M}}\gamma_k\le\max_{(d+1)\le l\le d\theta_{\mathcal{D}}^2}\gamma_l$, therefore

$$w\cdot\mathbb{B}_K w\le C_B\,w\cdot\mathbb{A}_K w,$$

where $C_B=C_A\max_{(d+1)\le l\le d\theta_{\mathcal{D}}^2}\gamma_l$ is a positive constant only depending on $\Lambda$, $\Omega$, $d$, and $\theta_{\mathcal{D}}$. This completes the proof of the comparison result (B.5). □

## Acknowledgements

The authors acknowledges support by the Labex CEMPI (ANR-11-LABX-0007-01). Claire Chainais-Hillairet acknowledges support by the ANR MOHYCON (ANR-17-CE40-0027-01).

## References

[1] A. Arnold, J. A. Carrillo, L. Desvillettes, J. Dolbeault, A. Jüngel, C. Lederman, P. A. Markowich, G. Toscani, and C. Villani, *Entropies and equilibria of many-particle systems: An essay on recent research*, Monatsh. Math., 142 (2004), pp. 35–43.

[2] C. Bataillon, F. Bouchon, C. Chainais-Hillairet, C. Desgranges, E. Hoarau, F. Martin, S. Perrin, M. Tupin, and J. Talandier, *Corrosion modelling of iron based alloy in nuclear waste repository*, Electrochimica Acta, 55 (2010), pp. 4451–4467.

[3] J. Bear, *Dynamics of Fluids in Porous Media*, Courier Corporation, 1988.

[4] L. Beirão da Veiga, J. Droniou, and G. Manzini, *A unified approach for handling convection terms in finite volumes and mimetic discretization methods for elliptic problems*, IMA J. Numer. Anal., 31 (2011), pp. 1357–1401.

[5] T. Bodineau, J. Lebowitz, C. Mouhot, and C. Villani, *Lyapunov functionals for boundary-driven nonlinear drift-diffusion equations*, Nonlinearity, 27 (2014), pp. 2111–2132.

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

[7] C. Buet and S. Dellacherie, *On the Chang and Cooper scheme applied to a linear Fokker-Planck equation*, Commun. Math. Sci., 8 (2010), pp. 1079–1090.

[8] C. Cancès, C. Chainais-Hillairet, M. Herda, and S. Krell, *Large time behavior of nonlinear finite volume schemes for convection-diffusion equations*, SIAM J. Numer. Anal., 58 (2020), pp. 2544–2571.

[9] C. Cancès, C. Chainais-Hillairet, and S. Krell, *Numerical analysis of a nonlinear free-energy diminishing Discrete Duality Finite Volume scheme for convection diffusion equations*, Comput. Methods Appl. Math., 18 (2018), pp. 407–432.

[10] C. Cancès and C. Guichard, *Numerical analysis of a robust free energy diminishing finite volume scheme for parabolic equations with gradient structure*, Found. Comput. Math., 17 (2017), pp. 1525–1584.

[11] J. A. Carrillo and G. Toscani, *Exponential convergence toward equilibrium for homogeneous Fokker-Planck-type equations*, Math. Methods Appl. Sci., 21 (1998), pp. 1269–1286.

[12] ——, *Asymptotic $L^1$-decay of solutions of the porous medium equation to self-similarity*, Indiana Univ. Math. J., 49 (2000), pp. 113–142.

[13] C. Chainais-Hillairet and M. Herda, *Large-time behaviour of a family of finite volume schemes for boundary-driven convection-diffusion equations*, IMA J. Numer. Anal., 40 (2020), pp. 2473–2504.

[14] J. Chang and G. Cooper, *A practical difference scheme for fokker-planck equations*, Journal of Computational Physics, 6 (1970), pp. 1–16.

[15] D. A. Di Pietro, J. Droniou, and A. Ern, *A discontinuous-skeletal method for advection-diffusion-reaction on general meshes*, SIAM J. Numer. Anal., 53 (2015), pp. 2135–2157.

[16] D. A. Di Pietro, A. Ern, and S. Lemaire, *An arbitrary-order and compact-stencil discretization of diffusion on general meshes based on local reconstruction operators*, Comput. Methods Appl. Math., 14 (2014), pp. 461–472.

[17] J. Droniou, *Non-coercive linear elliptic problems*, Potential Analysis, 17 (2002), pp. 181–203.

[18] ——, *Remarks on discretizations of convection terms in Hybrid Mimetic Mixed methods*, Networks & Heterogeneous Media, 5 (2010), pp. 545–563.

[19] ——, *Finite volume schemes for diffusion equations: Introduction to and review of modern methods*, Math. Models Methods Appl. Sci., 24 (2014), pp. 1575–1619.

[20] J. Droniou, R. Eymard, T. Gallouët, C. Guichard, and R. Herbin, *The Gradient Discretisation Method*, vol. 82 of Mathématiques & Applications, Springer International Publishing, Cham, Switzerland, 2018.

[21] J. Droniou, R. Eymard, T. Gallouët, and R. Herbin, *A unified approach to mimetic finite difference, hybrid finite volume and mixed finite volume methods*, Math. Models Methods Appl. Sci., 20 (2010), pp. 265–295.

[22] J. Droniou and J.-L. Vázquez, *Noncoercive convection-diffusion elliptic problems with Neumann boundary conditions*, Calc. Var., 34 (2009), pp. 413–434.

[23] L. C. Evans, *Partial Differential Equations: Second Edition*, vol. 19 of Graduate Studies in Mathematics, American Mathematical Society, Providence, R.I., 2010.

[24] R. Eymard, T. Gallouët, and R. Herbin, *Benchmark on anisotropic problems. SUSHI: a scheme using stabilization and hybrid interfaces for anisotropic heterogeneous diffusion problems*, in Finite Volumes for Complex Applications V - Problems & Perspectives, R. Eymard and J.-M. Hérard, eds., ISTE, London, 2008, pp. 801–814.

[25] ——, *Discretization of heterogeneous and anisotropic diffusion problems on general nonconforming meshes. SUSHI: a scheme using stabilization and hybrid interfaces*, IMA J. Numer. Anal., 30 (2010), pp. 1009–1043.

[26] F. Filbet and M. Herda, *A finite volume scheme for boundary-driven convection-diffusion equations with relative entropy structure*, Numer. Math., 137 (2017), pp. 535–577.

[27] R. Herbin and F. Hubert, *Benchmark on discretization schemes for anisotropic diffusion problems on general grids*, in Finite Volumes for Complex Applications V - Problems & Perspectives, R. Eymard and J.-M. Hérard, eds., ISTE, London, 2008, pp. 659–692.

[28] A. M. Il'in, *Differencing scheme for a differential equation with a small parameter affecting the highest derivative*, Math. Notes, 6 (1970), pp. 596–602.

[29] A. Jüngel, *Entropy Methods for Diffusive Partial Differential Equations*, SpringerBriefs in Mathematics, Springer International Publishing, Cham, Switzerland, 2016.

[30] C. Le Bris, F. Legoll, and F. Madiot, *Stabilisation de problèmes non coercifs via une méthode numérique utilisant la mesure invariante (Stabilization of non-coercive problems using the invariant measure)*, Comptes Rendus Mathématique, 354 (2016), pp. 799–803.

[31] L. Li and J.-G. Liu, *Large time behaviors of upwind schemes and B-schemes for Fokker-Planck equations on $\mathbb{R}$ by jump processes*, Mathematics of Computation, (2020).

[32] F. Madiot, *Multiscale finite element methods for advection-diffusion problems*, PhD thesis, Université Paris-Est, December 2016.

[33] P. A. Markowich, *The stationary semiconductor device equations*, Computational Microelectronics, Springer-Verlag, Vienna, 1986.

[34] P. A. Markowich, C. A. Ringhofer, and C. Schmeiser, *Semiconductor equations*, Springer-Verlag, Vienna, 1990.

[35] D. L. Scharfetter and H. K. Gummel, *Large-signal analysis of a silicon Read diode oscillator*, IEEE Transactions on Electron Devices, 16 (1969), pp. 64–77.

[36] G. Toscani, *Entropy production and the rate of convergence to equilibrium for the Fokker-Planck equation*, Quart. Appl. Math., 57 (1999), pp. 521–541.

[37] W. Van Roosbroeck, *Theory of the flow of electrons and holes in germanium and other semiconductors*, The Bell System Technical Journal, 29 (1950), pp. 560–607.