Discrete Gagliardo-Nirenberg inequality and application to the finite volume approximation of a convection-diffusion equation with a Joule effect term

Caterina Calgaro, Clément Cancès, Emmanuel Creusé

▶ **To cite this version:**

Caterina Calgaro, Clément Cancès, Emmanuel Creusé. Discrete Gagliardo-Nirenberg inequality and application to the finite volume approximation of a convection-diffusion equation with a Joule effect term. IMA Journal of Numerical Analysis, In press, 10.1093/imanum/drad063 . hal-03881410v2

HAL Id: hal-03881410
https://hal.science/hal-03881410v2

Submitted on 21 Dec 2023

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

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

Discrete Gagliardo-Nirenberg inequality and application to the finite volume approximation of a convection-diffusion equation with a Joule effect term

Caterina Calgaro[*†] Clément Cancès[‡] Emmanuel Creusé[§]

June 29, 2023

Abstract

A discrete order-two Gagliardo-Nirenberg inequality is established for piecewise constant functions defined on a two-dimensional structured mesh composed of rectangular cells. As in the continuous framework, this discrete Gagliardo-Nirenberg inequality allows to control in particular the L^4 norm of the discrete gradient of the numerical solution by the L^2 norm of its discrete Hessian times its L^∞ norm. This result is crucial for the convergence analysis of a finite volume method for the approximation of a convection-diffusion equation involving a Joule effect term on a uniform mesh in each direction. The convergence proof relies on compactness arguments and on *a priori* estimates under a smallness assumption on the data, which is essential also in the continuous framework.

AMS subject classification: 65M08 - 65M12

Keywords: discrete Gagliardo-Nirenberg inequality - Finite Volume scheme - Joule effect term.

1 Introduction

Variable density low-Mach models arise in a wide range of physical phenomena, in which the sound wave speed is much faster than the convective characteristic of the fluids. In the case of a calorically perfect gas, an asymptotic expansion of the variables with respect to the Mach number in the compressible Navier-Stokes equation leads to a low-Mach system [23]. In the particular configuration where the dynamic viscosity of the fluid can be explicitly given as a specific function of the temperature (see [5, 6]), a change of variables can be used in order to obtain a divergence-free velocity system. In that case, the mass conservation equation is reformulated in term of

[*]Univ. Lille, CNRS, Inria, UMR 8524 - Laboratoire Paul Painlevé, F-59000 Lille, France
[†]corresponding author, e-mail: caterina.calgaro@univ-lille.fr
[‡]Inria, Univ. Lille, CNRS, UMR 8524 - Laboratoire Paul Painlevé, F-59000 Lille, France
[§]Univ. Polytechnique Hauts-de-France, INSA Hauts-de-France, CERAMATHS, Laboratoire de Matériaux Céramiques et de Mathématiques, F-59313 Valenciennes, France

the temperature u instead of the density $1/u$, and contains a nonlinear term called the "Joule effect term" [8, Section 1] in analogy with systems accounting for energy dissipation stemming from electric currents, see for instance [20]. Results on local and global well-posedness of this system have been recently obtained under some smallness assumptions on the initial data, see e.g. [19, 8], still based on a formulation using the temperature as a primary variable.

From the numerical point of view, a combined finite volume - finite element scheme was proposed in [7] to simulate such a model in terms of temperature, velocity and pressure. The method is based on a time splitting, in which the first step consists in solving the mass conservation by an *ad-hoc* finite volume scheme. In dimensionless form, the mass conservation equation, set in a subdomain Ω of $\mathbb{R}^2$ and for positive times, writes

$$\partial_t u + \nabla \cdot (u\ \mathbf{v}) + \lambda\, |\nabla u|^2 - \lambda u\, \Delta u = 0, \tag{1}$$

where $\mathbf{v}$ is a given velocity field computed in the other step of the splitting algorithm. It is complemented by homogeneous Neumann boundary conditions and an initial datum u_0.

It has been proved in [7] that the scheme referred to as $\mathcal{SD}_{moy}\mathcal{J}_{up}$ therein preserves a discrete maximum principle property, as imposed by the physics of the problem. This paper is devoted to the numerical analysis of this finite volume numerical scheme for the approximation of the temperature, solution of (1).

Here, we rigorously establish the convergence of the finite volume solution towards the exact continuous one in the particular case of successively refined two-dimensional cartesian grids. As often, the method consists in deriving some *a priori* estimates on the numerical solution in order to establish the existence of discrete solutions to the scheme. Then, with the help of further estimates and of some compactness properties, the limits are proven to be the weak solutions of the equation. The originality of this contribution comes from the presence of the Joule effect term in the equation, leading to some specific difficulties.

Several times in our proof, we will make use of a discrete version of the inequality

$$\|\nabla u\|^2_{L^4(\Omega)^2} \leqslant C_{GN} \|\nabla^2 u\|_{L^2(\Omega)^{2\times 2}} \|u\|_{L^\infty(\Omega)} \tag{2}$$

which has not been established before up to our knowledge. At the continuous level, some classical Gagliardo-Nirenberg interpolation inequalities for intermediate derivatives in $\mathbb{R}^n$ have been established in the seminal papers of E. Gagliardo [15] and L. Nirenberg [24]. The following particular case of these results is stated in [14, Theorem 1.2] in the following form.

Theorem 1.1. *If $1 \leqslant q \leqslant \infty$, $1 \leqslant r < \infty$, $j, k \in \mathbb{N}$, $j < k$ and*

$$\frac{1}{p} = \frac{j}{kr} + \frac{k-j}{kq},$$

then there exists a constant C_{GN} independent of u such that

$$\|\nabla^j u\|^k_{L^p(\mathbb{R}^n)} \leqslant C_{GN} \|\nabla^k u\|^j_{L^r(\mathbb{R}^n)}\, \|u\|^{k-j}_{L^q(\mathbb{R}^n)} \quad \forall u \in L^q(\mathbb{R}^n) \cap W^{k,r}(\mathbb{R}^n).$$

This result holds in the particular case $j = 1$ and $k = 2$, so that if

$$\frac{2}{p} = \frac{1}{r} + \frac{1}{q},$$

we have:

$$\|\nabla u\|_{L^p(\mathbb{R}^n)} \leqslant C_{GN} \|\nabla^2 u\|^{1/2}_{L^r(\mathbb{R}^n)} \, \|u\|^{1/2}_{L^q(\mathbb{R}^n)} \quad \forall u \in L^q(\mathbb{R}^n) \cap W^{2,r}(\mathbb{R}^n). \tag{3}$$

Up to the fact that (2) holds on a bounded domain Ω (which does not yield particular difficulties), (2) is a particular case of (3). In what follows, we refer to (3) as a second order Gagliardo-Nirenberg inequality since the highest order of differentiation is $k = 2$. The first goal of this paper is to establish a discrete version of (3) in the particular case of some piecewise constant discrete functions defined on a cartesian grid. As far as we know, only first order discrete (Sobolev-)Gagliardo-Nirenberg inequalities are available in the literature so far, see [3, 2]. This new discrete estimate looks to us as a key element for new contributions in the field of the numerical analysis of partial differential equations.

The outline of the paper is the following. Section 2 introduces the discrete setting of the problem: the meshes, the associated discrete functional spaces and the discrete difference operators on these spaces. Section 3 is devoted to the second order discrete Gagliardo-Nirenberg inequality which is obtained following the lines of the continuous case detailed in [14], leading to Theorem 3.7. Then, Section 4 presents the finite volume scheme $\mathcal{SD}_{moy}\mathcal{J}_{up}$ previously introduced in [7] for the approximation of the convection-diffusion equation involving a Joule effect term, which is here formulated in the case of a two-dimensional Cartesian grid. We infer from the discrete maximum principle established in Section 5 that the scheme admits a unique solution, and further estimates of energy type are derived under some smallness assumption on the data. Section 6 then addresses the convergence of the finite volume solution towards the continuous one as the discretization parameters tend to 0, cf. Theorem 4.4. Some mostly elementary properties of the discrete operators are finally collected in Appendix for the ease of reading.

2 Discrete setting

In this section, we introduce the discrete framework such as the cartesian meshes, some discrete functional spaces as well as some differential operators needed for the following of the paper. We consider successively the 1D case and then the 2D one.

2.1 The case $d = 1$

Let $I =]\underline{x}, \overline{x}[$ be an open set of $\mathbb{R}$, consisting in a union of cells $\mathcal{M}$ defined by (see Figure 1):

$$\mathcal{M} = \{C_i =]x_{i-\frac{1}{2}}, x_{i+\frac{1}{2}}[, \ i \in [\![1, N]\!]\},$$

with $N \in \mathbb{N}^*$. We also define the set of shifted cells:

$$\widehat{\mathcal{M}} = \{C_{i+\frac{1}{2}} =]x_i, x_{i+1}[, \ i \in [\![0, N]\!]\}.$$

We denote $|I| = \overline{x} - \underline{x}$ the length of I and we define $x_0 = x_{\frac{1}{2}} = \underline{x}$ and $x_{N+\frac{1}{2}} = x_{N+1} = \overline{x}$. Let x_i be the center and $h_i = x_{i+\frac{1}{2}} - x_{i-\frac{1}{2}}$ the length of C_i for $i \in [\![1, N]\!]$. Let $x_{i+\frac{1}{2}}$ be the center and $h_{i+\frac{1}{2}} = x_{i+1} - x_i$ the length of $C_{i+\frac{1}{2}}$ for $i \in [\![0, N]\!]$.

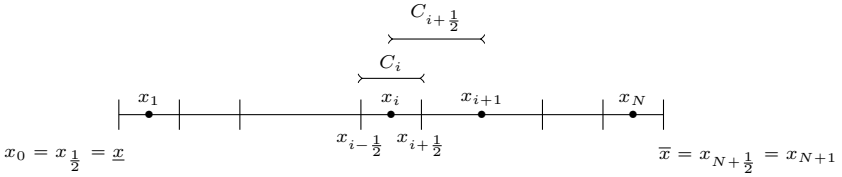

Figure 1: Notations for the grid cells of $\mathcal{M}$ and $\widehat{\mathcal{M}}$ in the case $d = 1$.

The spaces $H_{\mathcal{M}}(I)$, $\widehat{H}_{\mathcal{M}}(I)$ and $\widehat{H}^0_{\mathcal{M}}(I)$ are respectively defined by:

$$\begin{aligned} H_{\mathcal{M}}(I) &= \{v \in L^\infty(I) \mid v_{|C_i} = v_i \in \mathbb{R},\ i \in [\![1, N]\!]\}, \\ \widehat{H}_{\mathcal{M}}(I) &= \{v \in L^\infty(I) \mid v_{|C_{i+\frac{1}{2}}} = v_{i+\frac{1}{2}} \in \mathbb{R},\ i \in [\![0, N]\!]\}, \\ \widehat{H}^0_{\mathcal{M}}(I) &= \{v \in \widehat{H}_{\mathcal{M}}(I) \mid v_{\frac{1}{2}} = v_{N+\frac{1}{2}} = 0\}. \end{aligned}$$

The discrete operator δ_x is defined from $H_{\mathcal{M}}(I)$ in $\widehat{H}^0_{\mathcal{M}}(I)$ by:

$$\delta_x v(x)_{|C_{i+\frac{1}{2}}} = \begin{cases} \dfrac{v_{i+1} - v_i}{h_{i+\frac{1}{2}}} & \text{for } i \in [\![1, N-1]\!], \\ 0 & \text{for } i = 0 \text{ and } i = N. \end{cases}$$

Similarly, the discrete gradient operator δ_x^* is defined from $\widehat{H}_{\mathcal{M}}(I)$ in $H_{\mathcal{M}}(I)$ by:

$$\delta_x^* v(x)_{|C_i} = \frac{v_{i+\frac{1}{2}} - v_{i-\frac{1}{2}}}{h_i} \text{ for } i \in [\![1, N]\!].$$

The discrete second-order derivative operator δ_{xx} is defined from $H_{\mathcal{M}}(I)$ in $H_{\mathcal{M}}(I)$ by:

$$\delta_{xx} v = (\delta_x^* \circ \delta_x) v.$$

Now the interpolation operator π_x is defined from $H_{\mathcal{M}}(I)$ in $\widehat{H}_{\mathcal{M}}(I)$ by:

$$(\pi_x v)_{C_{i+\frac{1}{2}}} = \begin{cases} \dfrac{v_i + v_{i+1}}{2} & \text{for } i \in [\![1, N-1]\!], \\ v_1 & \text{for } i = 0, \\ v_N & \text{for } i = N, \end{cases}$$

and the interpolation operator π_x^* is defined from $\widehat{H}_{\mathcal{M}}(I)$ in $H_{\mathcal{M}}(I)$ by:

$$(\pi_x^* v)_{C_i} = \frac{v_{i-\frac{1}{2}} + v_{i+\frac{1}{2}}}{2} \text{ for } i \in [\![1, N]\!].$$

2.2 The case $d > 1$

2.2.1 Meshes and discrete functional spaces

We consider Ω a connected subset of $\mathbb{R}^d$ consisting in a union of rectangles ($d = 2$) or parallelepipeds ($d = 3$), possibly non-uniform. The edges (or faces) of these rectangles (or parallelepipeds) are assumed to be orthogonal to the canonical basis vectors. All

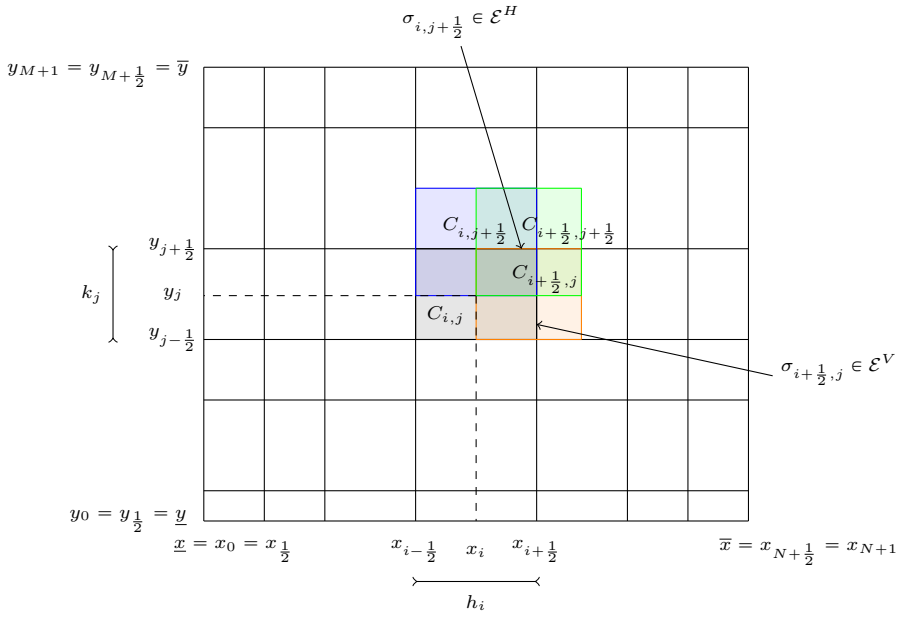

Figure 2: Notations for the grid cells of $\mathcal{M}, \widehat{\mathcal{M}}, \widetilde{\mathcal{M}}$ and $\overline{\mathcal{M}}$ in the case $d = 2$.

the notations are given in the case $d = 2$, but they can be generalized to the case $d = 3$.

Let $\Omega =]\underline{x}, \overline{x}[\times]\underline{y}, \overline{y}[\subset \mathbb{R}^2$ be the set of the grid cells $\mathcal{M}$ defined by (see Figure 2):

$$\mathcal{M} = \Big\{ C_{i,j} =]x_{i-\frac{1}{2}}, x_{i+\frac{1}{2}}[\times]y_{j-\frac{1}{2}}, y_{j+\frac{1}{2}}[,\ i \in [\![1, N]\!],\ j \in [\![1, M]\!] \Big\},$$

with $N, M \in \mathbb{N}^*$. We also define the set of shifted cells in the x-direction:

$$\widehat{\mathcal{M}} = \Big\{ C_{i+\frac{1}{2},j} =]x_i, x_{i+1}[\times]y_{j-\frac{1}{2}}, y_{j+\frac{1}{2}}[,\ i \in [\![0, N]\!],\ j \in [\![1, M]\!] \Big\},$$

the set of shifted cells in the y-direction:

$$\widetilde{\mathcal{M}} = \Big\{ C_{i,j+\frac{1}{2}} =]x_{i-\frac{1}{2}}, x_{i+\frac{1}{2}}[\times]y_j, y_{j+1}[,\ i \in [\![1, N]\!],\ j \in [\![0, M]\!] \Big\},$$

and the set of shifted cells in the x-y-directions:

$$\overline{\mathcal{M}} = \Big\{ C_{i+\frac{1}{2},j+\frac{1}{2}} =]x_i, x_{i+1}[\times]y_j, y_{j+1}[,\ i \in [\![0, N]\!],\ j \in [\![0, M]\!] \Big\}.$$

Similarly to the 1D case, we define:

$$h_{i+\frac{1}{2}} = x_{i+1} - x_i \text{ for } i \in [\![0, N]\!], \quad h_i = x_{i+\frac{1}{2}} - x_{i-\frac{1}{2}} \text{ for } i \in [\![1, N]\!],$$

$$k_{j+\frac{1}{2}} = y_{j+1} - y_j \text{ for } j \in [\![0, M]\!], \quad k_j = y_{j+\frac{1}{2}} - y_{j-\frac{1}{2}} \text{ for } j \in [\![1, M]\!].$$

We introduce the following functional spaces:

$$\begin{aligned}
H_{\mathcal{M}}(\Omega) &= \{v \in L^\infty(\Omega) \mid v_{|C_{i,j}} = v_{i,j} \in \mathbb{R},\ i \in [\![1,N]\!],\ j \in [\![1,M]\!]\},\\
\widehat{H}_{\mathcal{M}}(\Omega) &= \{v \in L^\infty(\Omega) \mid v_{|C_{i+\frac{1}{2},j}} = v_{i+\frac{1}{2},j} \in \mathbb{R},\ i \in [\![0,N]\!],\ j \in [\![1,M]\!]\},\\
\widehat{H}^0_{\mathcal{M}}(\Omega) &= \{v \in \widehat{H}_{\mathcal{M}}(\Omega) \mid v_{\frac{1}{2},j} = v_{N+\frac{1}{2},j} = 0,\ j \in [\![1,M]\!]\},\\
\widetilde{H}_{\mathcal{M}}(\Omega) &= \{v \in L^\infty(\Omega) \mid v_{|C_{i,j+\frac{1}{2}}} = v_{i,j+\frac{1}{2}} \in \mathbb{R},\ i \in [\![1,N]\!],\ j \in [\![0,M]\!]\},\\
\widetilde{H}^0_{\mathcal{M}}(\Omega) &= \{v \in \widetilde{H}_{\mathcal{M}}(\Omega) \mid v_{i,\frac{1}{2}} = v_{i,M+\frac{1}{2}} = 0,\ i \in [\![1,N]\!]\},
\end{aligned}$$

$$\begin{aligned}
\mathbb{H}_{\mathcal{M}}(\Omega) &= \widehat{H}_{\mathcal{M}}(\Omega) \times \widetilde{H}_{\mathcal{M}}(\Omega),\\
\mathbb{H}^0_{\mathcal{M}}(\Omega) &= \widehat{H}^0_{\mathcal{M}}(\Omega) \times \widetilde{H}^0_{\mathcal{M}}(\Omega),
\end{aligned}$$

$$\begin{aligned}
\bar{H}_{\mathcal{M}}(\Omega) &= \{v \in L^\infty(\Omega) \mid v_{|C_{i+\frac{1}{2},j+\frac{1}{2}}} = v_{i+\frac{1}{2},j+\frac{1}{2}} \in \mathbb{R},\ i \in [\![0,N]\!],\ j \in [\![0,M]\!]\},\\
\bar{H}^{0,x}_{\mathcal{M}}(\Omega) &= \{v \in \bar{H}_{\mathcal{M}}(\Omega),\ \mid v_{|C_{\frac{1}{2},j+\frac{1}{2}}} = v_{|C_{N+\frac{1}{2},j+\frac{1}{2}}} = 0,\ j \in [\![0,M]\!]\},\\
\bar{H}^{0,y}_{\mathcal{M}}(\Omega) &= \{v \in \bar{H}_{\mathcal{M}}(\Omega),\ \mid v_{|C_{i+\frac{1}{2},\frac{1}{2}}} = v_{|C_{i+\frac{1}{2},M+\frac{1}{2}}} = 0,\ i \in [\![0,N]\!]\},\\
\bar{H}^{0,0}_{\mathcal{M}}(\Omega) &= \bar{H}^{0,x}_{\mathcal{M}}(\Omega) \cap \bar{H}^{0,y}_{\mathcal{M}}(\Omega).
\end{aligned}$$

2.2.2 Discrete differential operators

The discrete operator δ_x is defined from $H_{\mathcal{M}}(\Omega)$ in $\widehat{H}^0_{\mathcal{M}}(\Omega)$ for $j \in [\![1,M]\!]$ (respectively from $\widetilde{H}_{\mathcal{M}}(\Omega)$ in $\bar{H}_{\mathcal{M}}(\Omega)$ for $j \in [\![0,M]\!] + \frac{1}{2}$) by:

$$\delta_x v(x,y)_{|C_{i+\frac{1}{2},j}} = \begin{cases} \dfrac{v_{i+1,j} - v_{i,j}}{h_{i+\frac{1}{2}}} & \text{for } i \in [\![1,N-1]\!],\\ 0 & \text{for } i = 0 \text{ and } i = N. \end{cases}$$

Similarly, the discrete operator δ_y is defined from $H_{\mathcal{M}}(\Omega)$ in $\widetilde{H}^0_{\mathcal{M}}(\Omega)$ for $i \in [\![1,N]\!]$ (respectively from $\widehat{H}_{\mathcal{M}}(\Omega)$ in $\bar{H}_{\mathcal{M}}(\Omega)$ for $i \in [\![0,N]\!] + \frac{1}{2}$) by :

$$\delta_y v(x,y)_{|C_{i,j+\frac{1}{2}}} = \begin{cases} \dfrac{v_{i,j+1} - v_{i,j}}{k_{j+\frac{1}{2}}} & \text{for } j \in [\![1,M-1]\!],\\ 0 & \text{for } j = 0 \text{ and } j = M. \end{cases}$$

The discrete operator δ_x^* is defined from $\widehat{H}_{\mathcal{M}}(\Omega)$ in $H_{\mathcal{M}}(\Omega)$ for $j \in [\![1,M]\!]$ (respectively from $\bar{H}_{\mathcal{M}}(\Omega)$ in $\widetilde{H}_{\mathcal{M}}(\Omega)$ for $j \in [\![0,M]\!] + \frac{1}{2}$) by:

$$\delta_x^* v(x,y)_{|C_{i,j}} = \frac{v_{i+\frac{1}{2},j} - v_{i-\frac{1}{2},j}}{h_i} \text{ for } i \in [\![1,N]\!].$$

Similarly, the discrete operator δ_y^* is defined from $\widetilde{H}_{\mathcal{M}}(\Omega)$ in $H_{\mathcal{M}}(\Omega)$ for $i \in [\![1,N]\!]$ (respectively from $\bar{H}_{\mathcal{M}}(\Omega)$ in $\widehat{H}_{\mathcal{M}}(\Omega)$ for $i \in [\![0,N]\!] + \frac{1}{2}$) by :

$$\delta_y^* v(x,y)_{|C_{i,j}} = \frac{v_{i,j+\frac{1}{2}} - v_{i,j-\frac{1}{2}}}{k_j} \text{ for } j \in [\![1,M]\!].$$

Then, the discrete gradient operator ∇_h is defined by:

$$\begin{aligned}
\nabla_h : \quad H_{\mathcal{M}}(\Omega) &\to \mathbb{H}^0_{\mathcal{M}}(\Omega)\\
v &\mapsto \nabla_h v = \begin{pmatrix} \delta_x v \\ \delta_y v \end{pmatrix}.
\end{aligned}$$

With a slight abuse of notation, we also denote in what follows by

$$\begin{array}{rccl} \nabla_h : & \mathbb{H}^0_{\mathcal{M}}(\Omega) & \to & H_{\mathcal{M}}(\Omega) \times \bar{H}_{\mathcal{M}}(\Omega) \times \bar{H}_{\mathcal{M}}(\Omega) \times H_{\mathcal{M}}(\Omega) \\ & \mathbf{v} = \begin{pmatrix} v_1 \\ v_2 \end{pmatrix} & \mapsto & \nabla_h \mathbf{v} = \begin{pmatrix} \delta_x^* v_1 & \delta_y v_1 \\ \delta_x v_2 & \delta_y^* v_2 \end{pmatrix}. \end{array} \tag{4}$$

The operators $\delta_{xx} = \delta_x^* \circ \delta_x$ and $\delta_{yy} = \delta_y^* \circ \delta_y$ are defined on $H_{\mathcal{M}}(\Omega)$, so that the discrete Laplace operator Δ_h is defined by:

$$\begin{array}{rccc} \Delta_h : & H_{\mathcal{M}}(\Omega) & \longrightarrow & H_{\mathcal{M}}(\Omega) \\ & v & \longrightarrow & \Delta_h v = (\delta_{xx} + \delta_{yy}) v. \end{array}$$

Finally, it remains to define the cross-derivatives by the operators δ_{yx} and δ_{xy} respectively defined from $H_{\mathcal{M}}(\Omega)$ in $\bar{H}^{0,0}_{\mathcal{M}}(\Omega)$ by:

$$\delta_{yx} v = \delta_{xy} v = (\delta_y \circ \delta_x) v = (\delta_x \circ \delta_y) v, \tag{5}$$

and the discrete Hessian matrix ∇_h^2 is defined for any $v \in H_{\mathcal{M}}(\Omega)$ by:

$$\nabla_h^2 v = \begin{pmatrix} \delta_{xx} v & \delta_{yx} v \\ \delta_{xy} v & \delta_{yy} v \end{pmatrix}.$$

2.2.3 Discrete interpolation operators

Some discrete interpolation operators are needed in order to pass from a given grid to another one, similarly to the ones given in subsection 2.1 for $d = 1$. The interpolation operator π_x is defined from $H_{\mathcal{M}}(\Omega)$ in $\widehat{H}_{\mathcal{M}}(\Omega)$ for $j \in [\![1, M]\!]$ (respectively from $\widetilde{H}_{\mathcal{M}}(\Omega)$ in $\bar{H}_{\mathcal{M}}(\Omega)$ for $j \in [\![0, M]\!] + \frac{1}{2}$) by:

$$(\pi_x v)_{C_{i+\frac{1}{2},j}} = \begin{cases} \dfrac{v_{i,j} + v_{i+1,j}}{2} & \text{for } i \in [\![1, N-1]\!], \\ v_{1,j} & \text{for } i = 0, \\ v_{N,j} & \text{for } i = N. \end{cases}$$

Similarly, the interpolation operator π_y is defined from $H_{\mathcal{M}}(\Omega)$ in $\widetilde{H}_{\mathcal{M}}(\Omega)$ for $i \in [\![1, N]\!]$ (respectively from $\widehat{H}_{\mathcal{M}}(\Omega)$ in $\bar{H}_{\mathcal{M}}(\Omega)$ for $i \in [\![0, N]\!] + \frac{1}{2}$) by:

$$(\pi_y v)_{C_{i,j+\frac{1}{2}}} = \begin{cases} \dfrac{v_{i,j} + v_{i,j+1}}{2} & \text{for } j \in [\![1, M-1]\!], \\ v_{i,1} & \text{for } j = 0, \\ v_{i,M} & \text{for } j = M. \end{cases}$$

The interpolation operator π_x^* is defined from $\widehat{H}_{\mathcal{M}}(\Omega)$ in $H_{\mathcal{M}}(\Omega)$ for $j \in [\![1, M]\!]$ (respectively from $\bar{H}_{\mathcal{M}}(\Omega)$ in $\widetilde{H}_{\mathcal{M}}(\Omega)$ for $j \in [\![0, M]\!] + \frac{1}{2}$) by:

$$(\pi_x^* v)_{C_{i,j}} = \frac{v_{i-\frac{1}{2},j} + v_{i+\frac{1}{2},j}}{2} \quad \text{for } i \in [\![1, N]\!].$$

Similarly, the interpolation operator π_y^* is defined from $\widetilde{H}_{\mathcal{M}}(\Omega)$ in $H_{\mathcal{M}}(\Omega)$ for $i \in [\![1, N]\!]$ (respectively from $\bar{H}_{\mathcal{M}}(\Omega)$ in $\widehat{H}_{\mathcal{M}}(\Omega)$ for $i \in [\![0, N]\!] + \frac{1}{2}$) by:

$$(\pi_y^* v)_{C_{i,j}} = \frac{v_{i,j-\frac{1}{2}} + v_{i,j+\frac{1}{2}}}{2} \quad \text{for } j \in [\![1, M]\!].$$

In the following of the paper, the discrete differential and interpolation operators fulfill some discrete properties, which are collected in Appendix A.

2.2.4 Norm definitions

Let $p \in \mathbb{R}$, $p \geqslant 1$. For any $v \in L^p(\Omega)$, the L^p norm is denoted:

$$\|v\|_{L^p(\Omega)} = \left(\int_\Omega |v|^p \, \mathrm{d}\mathbf{x} \right)^{1/p}. \tag{6}$$

The norm in the case $p = \infty$ means the essential supremum over Ω.
For any $\mathbf{v} = (v_i)_{1 \leqslant i \leqslant d} \in (L^p(\Omega))^d$, we define:

$$\|\mathbf{v}\|_{L^p(\Omega)} = \left(\sum_{i=1}^d \|v_i\|_{L^p(\Omega)}^p \right)^{1/p},$$

and for any $\underline{\mathbf{v}} = (v_{i,j})_{1 \leqslant i,j \leqslant d} \in (L^p(\Omega))^{d \times d}$, the L^p norm of $\underline{\mathbf{v}}$ is defined by:

$$\|\underline{\mathbf{v}}\|_{L^p(\Omega)}^p = \sum_{i,j=1}^d \|v_{i,j}\|_{L^p(\Omega)}^p. \tag{7}$$

3 The discrete Gagliardo-Nirenberg inequality

3.1 The 1D case

The goal of this subsection is to establish the discrete Gagliardo-Nirenberg inequality corresponding to the discrete 1D counterpart of (3):

Theorem 3.1. *Let $v \in H_{\mathcal{M}}(I)$, $1 \leqslant p, r < \infty$ and $1 \leqslant q \leqslant \infty$ such that*

$$\frac{2}{p} = \frac{1}{r} + \frac{1}{q}.$$

Then there exists C_{GN} independent of v such that:

$$\|\delta_x v\|_{L^p(I)} \leqslant C_{GN} \, \|\delta_{xx} v\|_{L^r(I)}^{1/2} \, \|v\|_{L^q(I)}^{1/2}. \tag{8}$$

We start to prove Theorem 3.1, following from the discrete point of view the work of [14] corresponding to the continuous case. We first establish some Lemma.

Lemma 3.2. *For any $v \in H_{\mathcal{M}}(I)$, any $J \subset I$ and for any $1 \leqslant p \leqslant \infty$ we have:*

$$\left\| v - \frac{1}{|J|} \int_J v(x)\,\mathrm{d}x \right\|_{L^p(J)} \leqslant 2 \inf_{c \in \mathbb{R}} \|v - c\|_{L^p(J)}.$$

Proof. The proof is exactly the same as the one of Lemma 3.1 of [14] since $H_{\mathcal{M}}(I) \subset L^p(J)$, $1 \leqslant p \leqslant \infty$. □

Lemma 3.3. *Let $r \in \mathbb{R}$, $r \geqslant 1$, $v \in H_{\mathcal{M}}(I)$ and $J \subset I$ such that:*

$$\int_J \delta_x v(x)\,\mathrm{d}x = 0. \tag{9}$$

Then for any $x \in J$ we have:

$$|\delta_x v(x)| \leqslant 2\, \|\delta_{xx} v\|_{L^r(J)}\, |J|^{\frac{r-1}{r}}. \tag{10}$$

Proof. We note $J =]\alpha, \beta[$, where $\alpha \in C_{i_\alpha + \frac{1}{2}}$ and $\beta \in C_{i_\beta + \frac{1}{2}}$ with $(i_\alpha, i_\beta) \in [\![0, N]\!]^2$, $i_\alpha \leqslant i_\beta$ (see Figure 3). Without loss of generality, we suppose $i_\alpha < i_\beta$.

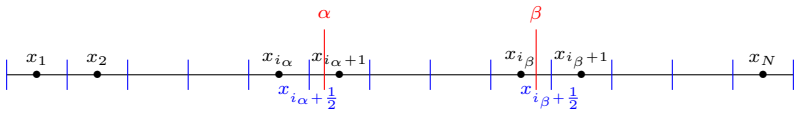

Figure 3: Configuration $J =]\alpha, \beta[\subset I$ with $\alpha > x_{i_\alpha + \frac{1}{2}}$ and $\beta < x_{i_\beta + \frac{1}{2}}$.

We define $\tilde{h}_{i_\alpha + \frac{1}{2}} = x_{i_\alpha + 1} - \alpha$, $\tilde{h}_{i_\beta + \frac{1}{2}} = \beta - x_{i_\beta}$ and $\tilde{h}_{k+\frac{1}{2}} = h_{k+\frac{1}{2}}$, $k \in [\![i_\alpha + 1, i_\beta - 1]\!]$. The property (9) can consequently be written as:

$$\sum_{k=i_\alpha}^{i_\beta} \tilde{h}_{k+\frac{1}{2}}\, (\delta_x v)_{k+\frac{1}{2}} = 0. \tag{11}$$

Let $x \in J$ and $j \in [\![i_\alpha, i_\beta]\!]$ such that $|(\delta_x v)_{j+\frac{1}{2}}| = \|\delta_x v\|_{L^\infty(J)}$. We have:

$$|\delta_x v(x)| \quad \leqslant \quad |(\delta_x v)_{j+\frac{1}{2}}|. \tag{12}$$

Considering now $i \in [\![i_\alpha, i_\beta]\!]$ such that:

$$|(\delta_x v)_{j+\frac{1}{2}} - (\delta_x v)_{i+\frac{1}{2}}| = \max_{k \in [\![i_\alpha, i_\beta]\!]} |(\delta_x v)_{j+\frac{1}{2}} - (\delta_x v)_{k+\frac{1}{2}}|,$$

we have from (11):

$$\begin{aligned} |(\delta_x v)_{j+\frac{1}{2}}| &\leqslant |(\delta_x v)_{j+\frac{1}{2}} - (\delta_x v)_{i+\frac{1}{2}}| \\ &\leqslant \sum_{k=i+1}^{j} |(\delta_x v)_{k+\frac{1}{2}} - (\delta_x v)_{k-\frac{1}{2}}| \\ &\leqslant \sum_{k=i_\alpha+1}^{i_\beta} |(\delta_x v)_{k+\frac{1}{2}} - (\delta_x v)_{k-\frac{1}{2}}|. \end{aligned} \tag{13}$$

We suppose first that $\alpha > x_{i_\alpha+\frac{1}{2}}$ and $\beta < x_{i_\beta+\frac{1}{2}}$, like in Figure 3. We introduce:

$$\begin{aligned}
\tilde{h}_{i_\alpha+1} &= \tilde{h}_{i_\alpha+\frac{1}{2}} + \frac{h_{i_\alpha+1}}{2}, \\
\tilde{h}_{i_\beta} &= \tilde{h}_{i_\beta+\frac{1}{2}} + \frac{h_{i_\beta}}{2}, \\
\tilde{h}_k &= h_k \quad \text{for } k \in [\![i_\alpha + 2, i_\beta - 1]\!].
\end{aligned}$$

We know that $h_{i_\alpha+1} \leqslant 2\,\tilde{h}_{i_\alpha+1}$ and $h_{i_\beta} \leqslant 2\,\tilde{h}_{i_\beta}$, and consequently:

$$\sum_{k=i_\alpha+1}^{i_\beta} \left| (\delta_x v)_{k+\frac{1}{2}} - (\delta_x v)_{k-\frac{1}{2}} \right| \leqslant 2 \sum_{k=i_\alpha+1}^{i_\beta} |((\delta_x^* \circ \delta_x)v)_k|\ \tilde{h}_k.$$

By the Hölder inequality we get:

$$\begin{aligned}
& \sum_{k=i_\alpha+1}^{i_\beta} \left| (\delta_x v)_{k+\frac{1}{2}} - (\delta_x v)_{k-\frac{1}{2}} \right| \\
\leqslant\ & 2 \left(\sum_{k=i_\alpha+1}^{i_\beta} |((\delta_x^* \circ \delta_x)v)_k|^r\ \tilde{h}_k \right)^{1/r} \left(\sum_{k=i_\alpha+1}^{i_\beta} \tilde{h}_k \right)^{\frac{r-1}{r}} \\
=\ & 2\, \|\delta_{xx} v\|_{L^r(J)}\, |J|^{\frac{r-1}{r}}. \qquad (14)
\end{aligned}$$

From (12), (13) and (14) we get (10) in the case $\alpha > x_{i_\alpha+\frac{1}{2}}$ and $\beta < x_{i_\beta+\frac{1}{2}}$. Then, in the case $\alpha < x_{i_\alpha+\frac{1}{2}}$ and $\beta > x_{i_\beta+\frac{1}{2}}$, we introduce:

$$\begin{aligned}
\tilde{h}_{i_\alpha} &= x_{i_\alpha+\frac{1}{2}} - \alpha, \\
\tilde{h}_{i_\beta+1} &= \beta - x_{i_\beta+\frac{1}{2}}, \\
\tilde{h}_k &= h_k \quad \text{for } k \in [\![i_\alpha + 1, i_\beta]\!].
\end{aligned}$$

This time, introducing $\overline{J} = [x_{i_\alpha+\frac{1}{2}}, x_{i_\beta+\frac{1}{2}}]$, we obtain by the Hölder inequality:

$$\begin{aligned}
& \sum_{k=i_\alpha+1}^{i_\beta} |(\delta_x v)_{k+\frac{1}{2}} - (\delta_x v)_{k-\frac{1}{2}}| \\
\leqslant\ & \left(\sum_{k=i_\alpha+1}^{i_\beta} |((\delta_x^* \circ \delta_x)v)_k|^r\ h_k \right)^{1/r} \left(\sum_{k=i_\alpha+1}^{i_\beta} h_k \right)^{\frac{r-1}{r}} \\
=\ & \|\delta_{xx} v\|_{L^r(\overline{J})}\, |\overline{J}|^{\frac{r-1}{r}} \\
\leqslant\ & \|\delta_{xx} v\|_{L^r(J)}\, |J|^{\frac{r-1}{r}}. \qquad (15)
\end{aligned}$$

From (12), (13) and (15) we get (10) in the case $\alpha < x_{i_\alpha+\frac{1}{2}}$ and $\beta > x_{i_\beta+\frac{1}{2}}$. Finally, in the two last cases (respectively $\alpha < x_{i_\alpha+\frac{1}{2}}$, $\beta < x_{i_\beta+\frac{1}{2}}$ and $\alpha > x_{i_\alpha+\frac{1}{2}}$, $\beta > x_{i_\beta+\frac{1}{2}}$), we proceed similarly and we obtain (10). The proof is complete. □

Lemma 3.4. *Let $r \in \mathbb{R}$, $r \geqslant 1$, $v \in H_{\mathcal{M}}(I)$ and $J \subset I$ such that*

$$\int_J v(x)\,\mathrm{d}x = 0. \tag{16}$$

Then for any $x \in J$ we have:

$$|v(x)| \leqslant 2\,\|\delta_x v\|_{L^r(J)}\,|J|^{\frac{r-1}{r}}. \tag{17}$$

Proof. The proof is very similar to the one of Lemma 3.3 and based on the same arguments. We note $J =]\alpha, \beta[$, where $\alpha \in C_{i_\alpha}$ and $\beta \in C_{i_\beta}$ with $(i_\alpha, i_\beta) \in [\![1, N]\!]^2, i_\alpha \leqslant i_\beta$. Without loss of generality, we suppose $i_\alpha < i_\beta$.
We define $\tilde{h}_{i_\alpha} = x_{i_\alpha+\frac{1}{2}} - \alpha$, $\tilde{h}_{i_\beta} = \beta - x_{i_\beta-\frac{1}{2}}$ and $\tilde{h}_k = h_k$, $k \in [\![i_\alpha + 1, i_\beta - 1]\!]$. The property (16) can be written as:

$$\sum_{k=i_\alpha}^{i_\beta} \tilde{h}_k\, v_k = 0. \tag{18}$$

Now, let $x \in J$ and $j \in [\![i_\alpha, i_\beta]\!]$ such that $|v_j| = \|v\|_{L^\infty(J)}$. We have:

$$|v(x)| \quad \leqslant \quad |v_j|. \tag{19}$$

Considering now $i \in [\![i_\alpha, i_\beta]\!]$ such that $|v_j - v_i| = \max\limits_{k \in [\![i_\alpha, i_\beta]\!]} |v_j - v_k|$, we have from (18):

$$|v_j| \leqslant |v_j - v_i| \leqslant \sum_{k=i+1}^{j} |v_k - v_{k-1}| \leqslant \sum_{k=i_\alpha+1}^{i_\beta} |v_k - v_{k-1}|. \tag{20}$$

We suppose first that $\alpha > x_{i_\alpha}$ and $\beta < x_{i_\beta}$. We introduce:

$$\begin{aligned}
\tilde{h}_{i_\alpha+\frac{1}{2}} &= \tilde{h}_{i_\alpha} + \frac{h_{i_\alpha+\frac{1}{2}}}{2}, \\
\tilde{h}_{i_\beta-\frac{1}{2}} &= \tilde{h}_{i_\beta} + \frac{h_{i_\beta-\frac{1}{2}}}{2}, \\
\tilde{h}_{k+\frac{1}{2}} &= h_{k+\frac{1}{2}} \quad \text{for } k \in [\![i_\alpha + 1, i_\beta - 2]\!].
\end{aligned}$$

We know that $h_{i_\alpha+\frac{1}{2}} \leqslant 2\,\tilde{h}_{i_\alpha+\frac{1}{2}}$ and $h_{i_\beta-\frac{1}{2}} \leqslant 2\,\tilde{h}_{i_\beta-\frac{1}{2}}$. Consequently:

$$\sum_{k=i_\alpha+1}^{i_\beta} |v_k - v_{k-1}| \quad \leqslant \quad 2 \sum_{k=i_\alpha+1}^{i_\beta} \left|(\delta_x v)_{k-\frac{1}{2}}\right|\, \tilde{h}_{k-\frac{1}{2}}.$$

By the Hölder inequality we get:

$$\begin{aligned}
\sum_{k=i_\alpha+1}^{i_\beta} |v_k - v_{k-1}| &\leqslant 2 \left(\sum_{k=i_\alpha+1}^{i_\beta} \left|(\delta_x v)_{k-\frac{1}{2}}\right|^r\, \tilde{h}_{k-\frac{1}{2}} \right)^{1/r} \left(\sum_{k=i_\alpha+1}^{i_\beta} \tilde{h}_{k-\frac{1}{2}} \right)^{\frac{r-1}{r}} \\
&= 2\,\|\delta_x v\|_{L^r(J)}\,|J|^{\frac{r-1}{r}}.
\end{aligned} \tag{21}$$

From (19), (20) and (21) we get (17) in the case $\alpha > x_{i_\alpha}$ and $\beta < x_{i_\beta}$.

For the three other cases ($\alpha < x_{i_\alpha}$ and $\beta < x_{i_\beta}$; $\alpha > x_{i_\alpha}$ and $\beta > x_{i_\beta}$; $\alpha < x_{i_\alpha}$ and $\beta > x_{i_\beta}$), we proceed in the same way, similarly to Lemma 3.3. □

Lemma 3.5. *Let $p, q, r \in \mathbb{R}$, $p \geqslant 1$, $q \geqslant 1$, $r \geqslant 1$, $J \subset I$. For any $v \in H_{\mathcal{M}}(I)$, there exists C independent of v such that*

$$\|\delta_x v\|_{L^p(J)} \leqslant C \left(|J|^{1+\frac{1}{p}-\frac{1}{r}} \|\delta_{xx} v\|_{L^r(J)} + |J|^{-1+\frac{1}{p}-\frac{1}{q}} \|v\|_{L^q(J)} \right).$$

Proof. The proof is similar to the one of Lemma 3.2 in [14], and we give it here for completeness. First we introduce

$$\overline{v} = \frac{1}{|J|} \int_J v(x)\,\mathrm{d}x.$$

From Lemma 3.2, we have:

$$\|v - \overline{v}\|_{L^q(J)} \approx \inf_{c \in \mathbb{R}} \|v - c\|_{L^q(J)},$$

so that we may assume $\overline{v} = 0$. We denote $d = \dfrac{1}{|J|} \displaystyle\int_J \delta_x v(x)\,\mathrm{d}x$ and X_0 the center of J, and we define

$$\widetilde{v}(x) = v(x) - d(x - X_0), \tag{22}$$

so that

$$\int_J \widetilde{v}(x)\,\mathrm{d}x = 0, \tag{23}$$

and

$$\int_J \delta_x \widetilde{v}(x)\,\mathrm{d}x = 0. \tag{24}$$

From (24) and Lemma 3.3, we get:

$$\|\delta_x \widetilde{v}\|_{L^p(J)} \leqslant 2\,|J|^{\frac{r-1}{r}+\frac{1}{p}} \|\delta_{xx} v\|_{L^r(J)}. \tag{25}$$

From (23), Lemma 3.4 and (25), we get:

$$\|\widetilde{v}\|_{L^q(J)} \leqslant 4\,|J|^{1+\frac{1}{q}+\frac{r-1}{r}} \|\delta_{xx} v\|_{L^r(J)}. \tag{26}$$

Finally we have:

$$\begin{aligned}
\|\delta_x v\|_{L^p(J)} &\overset{(22)}{\leqslant} \|\delta_x \widetilde{v}\|_{L^p(J)} + \|d\|_{L^p(J)} \\
&= \|\delta_x \widetilde{v}\|_{L^p(J)} + \frac{\|1\|_{L^p(J)} \cdot \|d(x - X_0)\|_{L^q(J)}}{\|x - X_0\|_{L^q(J)}} \\
&\overset{(25)}{\lesssim} |J|^{1+\frac{1}{p}-\frac{1}{r}} \|\delta_{xx} v\|_{L^r(J)} + |J|^{-1+\frac{1}{p}-\frac{1}{q}} \|d(x - X_0)\|_{L^q(J)} \\
&\overset{(22)}{\leqslant} |J|^{1+\frac{1}{p}-\frac{1}{r}} \|\delta_{xx} v\|_{L^r(J)} + |J|^{-1+\frac{1}{p}-\frac{1}{q}} (\|v\|_{L^q(J)} + \|\widetilde{v}\|_{L^q(J)}) \\
&\overset{(26)}{\lesssim} |J|^{1+\frac{1}{p}-\frac{1}{r}} \|\delta_{xx} v\|_{L^r(J)} + |J|^{-1+\frac{1}{p}-\frac{1}{q}} \|v\|_{L^q(J)}.
\end{aligned}$$

The proof is complete. □

Lemma 3.6. *Let $v \in H_{\mathcal{M}}(I)$ and $1 \leqslant p, r < \infty$, $1 \leqslant q \leqslant \infty$ such that*

$$\frac{2}{p} = \frac{1}{r} + \frac{1}{q}.$$

Then, there exists a sequence of open intervals (I_k), which covers $\overline{I}$, such that:

$$\begin{aligned} |I_k|^{1+\frac{1}{p}-\frac{1}{r}} \|\delta_{xx} v\|_{L^r(I_k)} &= |I_k|^{-1+\frac{1}{p}-\frac{1}{q}} \|v\|_{L^q(I_k)}, \\ \sum_k \chi_{I_k} &\leqslant 4. \end{aligned}$$

Proof. The proof is exactly the same as the one of Lemma 3.3 in [14]. We only mention that we need the values of r and q not to be equal to $+\infty$, so that the functions ω_x and α_x remain continuous, since in the proof we have to replace $\mathcal{C}_c^\infty(\mathbb{R})$ by $H_{\mathcal{M}}(I)$. □

Proof. (of Theorem 3.1)

First, we consider $1 \leqslant p_n, q_n, r_n < \infty$ such that $\dfrac{2}{p_n} = \dfrac{1}{r_n} + \dfrac{1}{q_n}$. Following the proof of Lemma 3.4 in [14] and using previous Lemma 3.5 and 3.6 (which respectively correspond to the discrete versions of Lemma 3.2 and 3.3 in [14]), we obtain:

$$\|\delta_x v\|_{L^{p_n}(I)}^{p_n} \lesssim \|\delta_{xx} v\|_{L^{r_n}(I)}^{\frac{p_n}{2}} \|v\|_{L^{q_n}(I)}^{\frac{p_n}{2}}, \tag{27}$$

so that (8) holds. Then, we can write:

$$\|v\|_{L^{q_n}(I)} \leqslant \|v\|_{L^\infty(I)}^{\frac{q_n-1}{q_n}} \|v\|_{L^1(I)}^{\frac{1}{q_n}},$$

and thanks to (27) we get:

$$\|\delta_x v\|_{L^{p_n}(I)}^{p_n} \lesssim \|\delta_{xx} v\|_{L^{r_n}(I)}^{\frac{p_n}{2}} \|v\|_{L^\infty(I)}^{\frac{p_n}{2}\left(\frac{q_n-1}{q_n}\right)} \|v\|_{L^1(I)}^{\frac{p_n}{2q_n}}.$$

Now it remains to make q_n tend towards $+\infty$ to obtain (8) in the case $q = \infty$. □

3.2 The 2D case

The goal of this subsection is now to establish the discrete Gagliardo-Nirenberg inequality corresponding to the discrete 2D counterpart of (3):

Theorem 3.7. *Let $\Omega =]\underline{x}, \overline{x}[\times]\underline{y}, \overline{y}[$ be an open set of $\mathbb{R}^2$, $v \in H_{\mathcal{M}}(\Omega)$, $1 \leqslant p, r < \infty$ and $1 \leqslant q \leqslant \infty$ such that*

$$\frac{2}{p} = \frac{1}{r} + \frac{1}{q}.$$

Then, we have:

$$\|\nabla_h v\|_{L^p(\Omega)} \leqslant 2^{-\frac{1}{2q}} C_{GN} \|\nabla_h^2 v\|_{L^r(\Omega)}^{1/2} \|v\|_{L^q(\Omega)}^{1/2}, \tag{28}$$

where C_{GN} is the constant arising in Theorem 3.1.

Proof. First of all, from $v \in H_{\mathcal{M}}(\Omega)$ we define some discrete one-variable functions: For $j \in [\![1, M]\!]$, $v^{(j)} \in H_{\mathcal{M}}(\mathcal{I}_x)$ with $\mathcal{I}_x =]\underline{x}, \overline{x}[$ and

$$\|\delta_x v^{(j)}\|_{L^p(\mathcal{I}_x)}^p = \int_{\mathcal{I}_x} |\delta_x v^{(j)}|^p \, \mathrm{d}x.$$

For $i \in [\![1, N]\!]$, $v^{(i)} \in H_{\mathcal{M}}(\mathcal{I}_y)$ with $\mathcal{I}_y =]\underline{y}, \overline{y}[$ and

$$\|\delta_y v^{(i)}\|_{L^p(\mathcal{I}_y)}^p = \int_{\mathcal{I}_y} |\delta_y v^{(i)}|^p \, \mathrm{d}y.$$

We have:

$$\begin{aligned}\|\nabla_h v\|_{L^p(\Omega)}^p &= \int_\Omega |\delta_x v|^p \, \mathrm{d}\mathbf{x} + \int_\Omega |\delta_y v|^p \, \mathrm{d}\mathbf{x} \\ &= \sum_{j=1}^M k_j \, \|\delta_x v^{(j)}\|_{L^p(\mathcal{I}_x)}^p + \sum_{i=1}^N h_i \, \|\delta_y v^{(i)}\|_{L^p(\mathcal{I}_y)}^p.\end{aligned}$$

From Theorem 3.1 and Hölder inequality, we get:

$$\begin{aligned}\|\nabla_h v\|_{L^p(\Omega)}^p &\leqslant C_{GN}^p \left(\sum_{j=1}^M k_j \, \|\delta_{xx} v^{(j)}\|_{L^r(\mathcal{I}_x)}^{\frac{p}{2}} \, \|v^{(j)}\|_{L^q(\mathcal{I}_x)}^{\frac{p}{2}} + \sum_{i=1}^N h_i \, \|\delta_{yy} v^{(i)}\|_{L^r(\mathcal{I}_y)}^{\frac{p}{2}} \|v^{(i)}\|_{L^q(\mathcal{I}_y)}^{\frac{p}{2}} \right) \\ &\leqslant C_{GN}^p \left(\|\delta_{xx} v\|_{L^r(\Omega)}^{\frac{p}{2}} \; + \; \|\delta_{yy} v\|_{L^r(\Omega)}^{\frac{p}{2}} \right) \, \|v\|_{L^q(\Omega)}^{\frac{p}{2}} \\ &\leqslant 2^{\frac{p}{2r}-1} \, C_{GN}^p \, \|\nabla_h^2 v\|_{L^r(\Omega)}^{\frac{p}{2}} \, \|v\|_{L^q(\Omega)}^{\frac{p}{2}},\end{aligned}$$

so that (28) holds. □

Remark 3.8. *In the following section, we will focus on the case $p = 4, r = 2, q = \infty$:*

$$\|\nabla_h v\|_{L^4(\Omega)} \leqslant C_{GN} \, \|\nabla_h^2 \, v\|_{L^2(\Omega)}^{1/2} \, \|v_h\|_{L^\infty(\Omega)}^{1/2}. \tag{29}$$

Remark 3.9. *Up to a slight modification of the prefactor $2^{-\frac{1}{2q}}$ in $2^{-\frac{1}{q}}$, the discrete Gagliardo-Nirenberg inequality* (28) *also holds when the domain Ω is a subset of $\mathbb{R}^3$. Moreover, the equality* (86) *also holds true in the three-dimensional context. Indeed, the definitions of the discrete operators, scalar product and norms can be done in a similar way to the two-dimensional case. An induction argument is used to conclude the proof of Theorem 3.7, and a term-by-term identification can be done to obtain* (86).

4 Finite Volume scheme and a priori estimates

4.1 Model and continuous results

In this section, we are interested in a convection-diffusion equation involving a Joule effect term, given by:

$$\begin{aligned}\partial_t u + \nabla \cdot (u \, \mathbf{v}) + \lambda \, |\nabla u|^2 - \lambda u \, \Delta u &= 0, \quad \forall \, \mathbf{x} \in \Omega, \; \forall \, t \in]0, T], && \text{(30a)} \\ \nabla u(\mathbf{x}, t) \cdot \mathbf{n} &= 0, \quad \forall \, \mathbf{x} \in \partial\Omega, \; \forall \, t \in]0, T], && \text{(30b)} \\ u(\mathbf{x}, 0) &= u_0(\mathbf{x}), \quad \text{in } \Omega, && \text{(30c)}\end{aligned}$$

where $\Omega =]\underline{x}, \overline{x}[\times]\underline{y}, \overline{y}[\subset \mathbb{R}^2$, $\mathbf{n}$ is the outward unit normal vector to $\partial\Omega$, $T > 0$ is an arbitrary finite time horizon, $\lambda > 0$ is a fixed parameter, and the vector field $\mathbf{v} : Q_T = \Omega \times [0, T] \to \mathbb{R}^2$ is divergence free and satisfies a no-slip boundary condition, i.e. $\mathbf{v}(\mathbf{x}, t) = 0$ for all $\mathbf{x} \in \partial\Omega$ and $t \in [0, T]$. The system (30) can be seen as a particular case of a global low-Mach model with temperature dependent viscosity, in the case where $\mathbf{v}$ is a given datum of the problem (see e.g. [19, 8]). A local well-posedness result for strong solutions to (30) has been established in [8, Theorem 1]. More precisely, assuming that

$$u_0 \in H^2_N(\Omega) = \{w \in H^2(\Omega) \text{ s.t. } \nabla w(\mathbf{x}) \cdot \mathbf{n}(\mathbf{x}) = 0 \text{ for a.e. } \mathbf{x} \in \partial\Omega\},$$

that the convective velocity satisfies

$$\mathbf{v} \in L^2(0, T; V_0(\Omega) \cap (H^2(\Omega))^2) \quad \text{with } V_0(\Omega) = \{\mathbf{w} \in (H^1_0(\Omega))^2;\ \nabla \cdot \mathbf{w} = 0,\ \text{in } \Omega\}, \tag{31}$$

and if there exist two real numbers $u_\flat$ and $u^\sharp$ such that

$$0 < u_\flat \leqslant u_0(\mathbf{x}) \leqslant u^\sharp, \quad \forall \mathbf{x} \in \Omega, \tag{32}$$

then there exists $T > 0$ such that the system (30) admits a unique solution with the following regularity:

$$u \in L^2(0, T; H^3(\Omega)) \cap L^\infty(0, T; H^2_N(\Omega)), \quad \partial_t u \in L^2(0, T; H^1(\Omega)) \tag{33}$$

with

$$0 < u_\flat \leqslant u \leqslant u^\sharp \quad \text{a.e. } Q_T. \tag{34}$$

In this paper, we rather work with a weaker notion of solutions demanding for less regularity than (33).

Definition 4.1. *A function u is said to be a global in time* weak solution *to Problem* (30) *if $u \in L^\infty(Q_T; [u_\flat, u^\sharp]) \cap L^\infty((0, T); H^1(\Omega))$ with $\partial_t u$ and $\nabla^2 u \in L^2(Q_T)$, if $\nabla u \cdot \mathbf{n} = 0$ on $\partial\Omega \times (0, T)$, and if* (30a) *holds (with each term belonging to $L^2(Q_T)$).*

With such a lower regularity requirement, we are able to prove the existence of a global-in-time weak solution.

Theorem 4.2. *Suppose $u_0 \in H^1(\Omega)$ and that the assumptions* (31)-(32) *are satisfied. If*

$$u^\sharp - u_\flat \leqslant \delta, \tag{35}$$

for some $\delta > 0$ small enough (with a condition similar to the one of Theorem 4.4 below), then there exists a weak solution u satisfying

$$\|u - \overline{u}\|_{L^\infty(\Omega \times \mathbb{R}_+)} \leqslant \delta,$$

with $\overline{u} = \dfrac{u_\flat + u^\sharp}{2} > 0$. Moreover there exists $C \geqslant 0$ such that for all $t > 0$:

$$\|u(t) - \overline{u}\|^2_{H^1(\Omega)} + \int_0^t \left(\|\nabla u(s)\|^2_{L^2(\Omega)} + \|\Delta u(s)\|^2_{L^2(\Omega)} \right) ds \leqslant C \left(\|u_0 - \overline{u}\|^2_{H^1(\Omega)} + \int_0^t \|\nabla \mathbf{v}(s)\|^2_{L^2(\Omega)} ds \right). \tag{36}$$

The existence of such a global in time weak solution is a by-product of the Theorem 4.4 on the convergence of the finite volume scheme to be introduced in the next section. Note also that the assumption (35) is necessary to prove that the system (30) admits a unique global-in-time strong solution (33) with (34) (see [19]).

4.2 The Finite Volume scheme

We notice that $\nabla \cdot (u\nabla u) = |\nabla u|^2 + u\Delta u$. Then, the way to discretize the Joule effect term $|\nabla u|^2$ arising in (30a) must be consistent with the non-linear diffusion one. This is important in order to ensure some properties on the numerical solution, such as some maximum principles which hold at the continuous level. Moreover, the non-conservative way to write the diffusion term is consistent with the analysis that we will do, which mimics the continuous one. A rather similar Finite Volume (FV) scheme was initially introduced in [7].

In addition to the notations of subsection 2.2.1, we denote $\mathcal{E} = \mathcal{E}^H \cup \mathcal{E}^V$ the set of the horizontal and vertical edges of the mesh, i.e.

$$\mathcal{E}^H = \Big\{\sigma_{i,j+\frac{1}{2}} =]x_{i-\frac{1}{2}}, x_{i+\frac{1}{2}}[\times\{y_{j+\frac{1}{2}}\},\ i \in [\![1, N]\!],\ j \in [\![0, M]\!]\Big\},$$

$$\mathcal{E}^V = \Big\{\sigma_{i+\frac{1}{2},j} = \{x_{i+\frac{1}{2}}\}\times]y_{j-\frac{1}{2}}, y_{j+\frac{1}{2}}[,\ i \in [\![0, N]\!],\ j \in [\![1, M]\!]\Big\}.$$

Now we introduce the definition of a uniform mesh in each direction.

Definition 4.3. *A mesh $\mathcal{M}$ is said uniform in each direction if $h_i \equiv h_x$ for $i \in [\![1, N]\!]$ and $k_j \equiv h_y$ for $j \in [\![1, M]\!]$.*

From now on and for the sake of simplicity, we assume that the mesh $\mathcal{M}$ is uniform in each direction. As usual in the finite volume context, the size of the mesh is then defined as the diameter of the cells, i.e.

$$h = \sqrt{h_x^2 + h_y^2}.$$

We also introduce the transmissibility coefficient, given by

$$a_\sigma = \frac{h_y}{h_x} \text{ for } \sigma \in \mathcal{E}^V \quad \text{ and } \quad a_\sigma = \frac{h_x}{h_y} \text{ for } \sigma \in \mathcal{E}^H.$$

Let us introduce the space

$$V_{\mathcal{E},0}(\Omega) = \left\{\mathbf{v}_h = (v_{1,h}, v_{2,h}) \in \mathbb{H}^0_{\mathcal{M}}(\Omega) \mid \mathrm{div}_h \mathbf{v}_h = 0\right\}, \tag{37}$$

where the operator div_h is defined from $\mathbb{H}_{\mathcal{M}}(\Omega)$ in $H_{\mathcal{M}}(\Omega)$ by

$$\mathrm{div}_h \mathbf{v}_h = \delta_x^* v_{1,h} + \delta_y^* v_{2,h}.$$

Let $\mathcal{E}_{i,j}$ be the boundary of the control volume $C_{i,j}$ ($i \in [\![1, N]\!], j \in [\![1, M]\!]$). For $\sigma \in \mathcal{E}_{i,j}$, we denote by $\mathbf{n}_{i,j,\sigma}$ the exterior unit normal vector to σ. Given a fixed but arbitrary finite time horizon $T > 0$, we split the time interval $[0, T]$ in a uniform partition of time step $\tau = T/N_T$ for some $N_T \in \mathbb{N}_{>0}$, we define $t^n = n\tau$ ($0 \leqslant n \leqslant$

N_T) (so that $[0,T] = \bigcup_{0\leqslant n<N_T}[t^n, t^{n+1}]$). For any velocity field $\mathbf{v} = (v_1, v_2) \in L^2(0,T;V_0(\Omega))$, we define $\mathbf{v}_{h\tau} = (v_{1,h\tau}, v_{2,h\tau}) \in L^2(0,T;V_{\mathcal{E},0}(\Omega))$ by setting

$$v_{1,h\tau}(\mathbf{x},t) = v^n_{i+1/2,j} \ \text{ if } (\mathbf{x},t) \in C_{i+1/2,j} \times (t^n, t^{n+1}), \tag{38a}$$

$$v_{2,h\tau}(\mathbf{x},t) = v^n_{i,j+1/2} \ \text{ if } (\mathbf{x},t) \in C_{i,j+1/2} \times (t^n, t^{n+1}), \tag{38b}$$

with

$$v^n_{i+\frac{1}{2},j} = \frac{1}{\tau}\int_{t^n}^{t^{n+1}} \frac{1}{h_y} \int_{\sigma_{i+\frac{1}{2},j}} v_1(\mathbf{x},s)\, d\sigma(\mathbf{x})\, ds, \quad i \in [\![1, N-1]\!],\ j \in [\![1, M]\!], \tag{38c}$$

$$v^n_{i,j+\frac{1}{2}} = \frac{1}{\tau}\int_{t^n}^{t^{n+1}} \frac{1}{h_x} \int_{\sigma_{i,j+\frac{1}{2}}} v_2(\mathbf{x},s)\, d\sigma(\mathbf{x})\, ds, \quad i \in [\![1, N]\!],\ j \in [\![1, M-1]\!]. \tag{38d}$$

The integrands in the above formulas have to be understood as the traces of $(v_1, v_2) \in L^2(0,T;(H^2(\Omega))^2)$ on the edges. Moreover, since $H^2(\Omega)$ embeds in $L^\infty(\Omega)$ (this also holds true in the three-dimensional setting), then

$$|v^n_{i,j+\frac{1}{2}}| \leqslant \frac{1}{\tau}\int_{t^n}^{t^{n+1}} \|\mathbf{v}(\cdot,t)\|_\infty dt$$

We define moreover $\mathbf{v}^n_h \in V_{\mathcal{E},0}(\Omega)$ by:

$$\mathbf{v}^n_h(\mathbf{x}) = (v^n_{1,h}(\mathbf{x}), v^n_{2,h}(\mathbf{x})) = \frac{1}{\tau}\int_{t^n}^{t^{n+1}} \mathbf{v}_{h\tau}(\mathbf{x},s)\, ds \ \ \forall\, \mathbf{x} \in \Omega.$$

We infer from Jensen's inequality that

$$\|\mathbf{v}_{h\tau}\|_{L^2(0,T;L^\infty(\Omega))} \leqslant \|\mathbf{v}\|_{L^2(0,T;L^\infty(\Omega))} \leqslant C_\Omega \|\mathbf{v}\|_{L^2(0,T:H^2(\Omega))} \tag{39}$$

with C_Ω being the continuity constant for the injection of $H^2(\Omega)$ into $L^\infty(\Omega)$.

The initial data u_0 is discretized into

$$u^0_{i,j} = \frac{1}{h_x h_y}\int_{C_{i,j}} u_0(\mathbf{x})\mathrm{d}\mathbf{x}, \qquad i \in [\![1, N]\!], j \in [\![1, M]\!]. \tag{40}$$

Assuming that $u^n_h \in H_{\mathcal{M}}(\Omega)$ is a known approximation of $u(\cdot, t^n)$, we are looking for an approximation $u^{n+1}_h \in H_{\mathcal{M}}(\Omega)$ of $u(\cdot, t^{n+1})$, with

$$u^n_h(\mathbf{x}) = u^n_{i,j} \quad \text{if } \mathbf{x} \in C_{i,j}, \quad n \geqslant 0.$$

The space-time approximate solution $u_{h\tau} \in L^\infty(0,T;H_{\mathcal{M}}(\Omega))$ is then defined almost everywhere by

$$u_{h\tau}(\mathbf{x},t) = u^{n+1}_h(\mathbf{x}) \quad \text{if } t \in (t^n, t^{n+1}].$$

The scheme is obtained by integrating (30a) on each $C_{i,j} \in \mathcal{M}$, leading to

$$h_x h_y \frac{u^{n+1}_{i,j} - u^n_{i,j}}{\tau} + \sum_{\sigma\in\mathcal{E}_{i,j}} v^n_{i,j,\sigma}\, u^{n+1}_{i,j,\sigma,+} + \lambda\, h_x h_y \left(\mathcal{J}_{i,j}(u^{n+1}_h) - u^{n+1}_{i,j}(\Delta_h u_h)^{n+1}_{i,j}\right) = 0. \tag{41}$$

In the above equation (41), we defined $v^n_{i,j,\sigma}$ by

$$v^n_{i,j,\sigma} = \begin{cases} \pm h_y \, v^n_{i\pm\frac{1}{2},j} & \text{if } \sigma = \sigma_{i\pm\frac{1}{2},j} \in \mathcal{E}^V, \\ \pm h_x \, v^n_{i,j\pm\frac{1}{2}} & \text{if } \sigma = \sigma_{i,j\pm\frac{1}{2}} \in \mathcal{E}^H, \end{cases}$$

and by $u^{n+1}_{i,j,\sigma,+}$ the upstream choice for the convection term defined for $\sigma \in \mathcal{E}_{i,j}$:

$$u^{n+1}_{i,j,\sigma,+} = \begin{cases} u^{n+1}_{i,j} & \text{if } v^n_{i,j,\sigma} \geqslant 0, \\ u^{n+1}_{i,j,\sigma} & \text{otherwise,} \end{cases} \quad \text{with} \quad u^{n+1}_{i,j,\sigma} = \begin{cases} u^{n+1}_{i\pm1,j} & \text{if } \sigma = \sigma_{i\pm\frac{1}{2},j} \in \mathcal{E}^V, \\ u^{n+1}_{i,j\pm1} & \text{if } \sigma = \sigma_{i,j\pm\frac{1}{2}} \in \mathcal{E}^H, \\ u^{n+1}_{i,j} & \text{if } \sigma \subset \partial\Omega. \end{cases}$$

The discretization of the Joule effect term is more original as we set

$$\mathcal{J}_{i,j}(u^{n+1}_h) = \frac{1}{h_x h_y} \sum_{\sigma \in \mathcal{E}_{i,j}} a_\sigma \left((u^{n+1}_{i,j} - u^{n+1}_{i,j,\sigma})^+ \right)^2, \tag{42}$$

where $a^+ = \max(0, a)$. We also denote by

$$\mathcal{J}_h(u^{n+1}_h)(\mathbf{x}) = \mathcal{J}_{i,j}(u^{n+1}_h) \quad \text{if } \mathbf{x} \in C_{i,j}. \tag{43}$$

This discretization of the Joule effect term $|\nabla u|^2$ can be thought as some dual counterpart of the upstream convection, see [9]. It enjoys the following key property for the preservation of the maximum principle:

$$u^{n+1}_{i,j} \leqslant u^{n+1}_{i,j,\sigma} \text{ for all } \sigma \in \mathcal{E}_{i,j} \implies \mathcal{J}_{i,j}(u^{n+1}_h) = 0,$$

transposing to the discrete setting the fact that $|\nabla u|^2$ vanishes at the minima of u.

Besides the second order discrete Gagliardo-Nirenberg inequality stated in Section 3, the main result of the paper can be gathered in the following statement.

Theorem 4.4. *Let $u_0 \in H^1(\Omega)$ be such that $u_\flat \leqslant u_0 \leqslant u^\sharp$ for some (strictly) positive constants $u_\flat, u^\sharp$, then the numerical scheme (40)–(41)–(42) admits a unique iterated in time solution $u_{h\tau}$ with $u_\flat \leqslant u_{h\tau} \leqslant u^\sharp$ a.e. in Q_T. Moreover, if $u^\sharp - u_\flat < \delta$ with $0 < \delta < \frac{2}{(C_{GN})^2}(\sqrt{1 + \frac{u_\flat}{2}} - 1)$, then, up to a subsequence,*

$$u_{h\tau} \underset{h,\tau \to 0}{\longrightarrow} u \quad \text{a.e. in } Q_T$$

where u is a weak solution to the continuous problem in the sense of Definition 4.1.

Remark 4.5. *The constraint on δ might look restrictive but it is imposed by the global well-posedness of the continuous problem. We emphasize that the proposed convergence result applies to (30), but it is also motivated by a practical application on a ghost effect system (see [22, 19]). Ghost effect systems are formally derived to describe regimes in which the compressible Navier-Stokes system is incomplete, in particular when the classical heat-conduction equation fails to correctly describe the temperature field of the gas. In such a physical context, the parameter δ is expected to be small. The analysis done in this work can be considered as part of the analysis of a numerical scheme for a ghost effect system or a low Mach model expressed in velocity, pressure and temperature variables, as proposed in [7] where some numerical tests are also presented.*

The two next sections are devoted to the proof of Theorem 4.4. Moreover, finer convergence properties will be derived along the proof, especially in Section 6.1.

5 Numerical analysis at fixed grid

The goal of this section is to prove the well-posedness of the numerical scheme as well as estimates which are uniform with respect to the grid. Those estimates will serve as cornerstones for the convergence proof reported in Section 6.

5.1 Maximum principle and existence of a discrete solution

We first establish a uniform L^∞ a priori estimate on the discrete solution, on which we will rely to show the existence of a discrete solution u_h to the scheme (41).

Proposition 5.1. *Assume that there exist two positive constants $u_\flat, u^\sharp$ such that*

$$0 < u_\flat \leqslant u_0 \leqslant u^\sharp. \tag{44}$$

Then for all $n \in [\![1, N_T]\!]$ the finite volume scheme (40)–(41) admits a unique solution $u_h^n \in H_{\mathcal{M}}(\Omega)$ which satisfies

$$0 < u_\flat \leqslant u_{i,j}^n \leqslant u^\sharp \quad \forall\, i \in [\![1, N]\!], j \in [\![1, M]\!],\ n \in [\![1, N_T]\!]. \tag{45}$$

Proof. The proof is done by induction over n. The initialization for $n = 0$ is straightforward in view of (44) and the definition (40) of the initial discrete solution. We perform a harmless modification of the scheme, which now writes

$$h_x h_y \frac{u_{i,j}^{n+1} - u_{i,j}^n}{\tau} + \sum_{\sigma \in \mathcal{E}_{i,j}} v_{i,j,\sigma}^n\, u_{i,j,\sigma,+}^{n+1} + \lambda h_x h_y \left(\mathcal{J}_{i,j}(u_h^{n+1}) - (u_{i,j}^{n+1})^+ (\Delta_h u_h)_{i,j}^{n+1} \right) = 0, \tag{46}$$

instead of (41). Of course, once (45) is established, we get that solutions to (46) are also solutions to (41).

The modified scheme (46) can be rewritten in the compact form

$$\mathcal{F}_{i,j}(u_h^{n+1}) = u_{i,j}^n, \qquad i \in [\![1, N]\!], j \in [\![1, M]\!],$$

where $\mathcal{F}_h = (\mathcal{F}_{i,j})_{i,j} : H_{\mathcal{M}}(\Omega) \to H_{\mathcal{M}}(\Omega)$ is increasing w.r.t. $u_{i,j}^{n+1}$ and non-increasing w.r.t. $u_{k,\ell}^{n+1}$ as soon as $(k, \ell) \neq (i, j)$.Moreover, since $\mathbf{v}_h^n$ is discrete divergence free, one has for all $\kappa \in \mathbb{R}$ that

$$\mathcal{F}_{i,j}(\kappa_h) = \kappa, \qquad i \in [\![1, N]\!], j \in [\![1, M]\!], \tag{47}$$

where κ_h is the element of $H_{\mathcal{M}}(\Omega)$ which is constant equal to κ. The Jacobian matrix $\mathbb{J}(u_h^{n+1}) = \left(\frac{\partial \mathcal{F}_{i,j}}{\partial u_{k,\ell}}(u_h^{n+1}) \right)_{(i,j),(k,\ell)}$ is a M-matrix in the sense of [18, Definition 4.8].

Let $\check{u}_h^{n+1}$ be another solution to (46) corresponding to some previous step value $\check{u}_h^n$, then

$$\mathcal{F}_h(u_h^{n+1}) - \mathcal{F}_h(\check{u}_h^{n+1}) = \overline{\mathbb{J}}(u_h^{n+1}, \check{u}_h^{n+1})(u_h^{n+1} - \check{u}_h^{n+1}) = u_h^n - \check{u}_h^n,$$

where

$$\overline{\mathbb{J}}(u_h^{n+1}, \check{u}_h^{n+1}) = \int_0^1 \mathbb{J}(\check{u}_h^{n+1} + t(u_h^{n+1} - \check{u}_h^{n+1}))\mathrm{d}t$$

is also a M matrix. It is in particular invertible with $\overline{\mathbb{J}}(u_h^{n+1}, \check{u}_h^{n+1})^{-1} \geqslant 0$ componentwise. Therefore,

$$u_h^n \geqslant \check{u}_h^n \quad \Longrightarrow \; u_h^{n+1} \geqslant \check{u}_h^{n+1}.$$

This yields in particular the uniqueness of the solution to (46), as well as the maximum principle (45) thanks to (47) if one chooses $\check{u}_h^n = \check{u}_h^{n+1}$ constant equal to $u_\flat$ or $u^\sharp$.

Finally, the existence of a solution to the modified scheme (46), and thus to the original one (41) is obtained thanks to some classical topological degree argument. We refer to [21, 10] for a general presentation of the topological gradient theory, and to [11] for its first (up to our knowledge) use in the context of finite volumes. □

5.2 Further estimates

The goal of this section is to establish the next estimates required to establish the convergence of the scheme. The main and next one is a $L^\infty_{\mathrm{loc}}(H^1) \cap L^2_{\mathrm{loc}}(H^2_N)$ estimate obtained under some smallness assumption on the data.

Proposition 5.2. *Assume that* (44) *holds. Then, there exists* $\delta > 0$ *such that if*

$$u^\sharp - u_\flat \leqslant \delta, \tag{48}$$

then there exists $c > 0$ *only depending on* Ω, u_0, $\mathbf{v}$, λ, δ *and* T *such that the solution* $u_h^n \in H_{\mathcal{M}}(\Omega)$ *of the scheme* (41) *built at Proposition 5.1 satisfies the following estimates:*

$$\|\nabla_h u_h^n\|_{L^2(\Omega)} \leqslant c, \quad \forall n \in [\![1, N_T]\!], \tag{49}$$

$$\sum_{n=0}^{N_T-1} \tau \left(\|\Delta_h u_h^{n+1}\|^2_{L^2(\Omega)} + \|\mathcal{J}_h(u_h^{n+1})\|^2_{L^2(\Omega)} \right) \leqslant c. \tag{50}$$

Proof. Before addressing the properties of u_h^n, $n \geqslant 1$, induced by the scheme, let us first remark that

$$\|\nabla_h u_h^0\|_{L^2(\Omega)} \leqslant \|\nabla u_0\|_{L^2(\Omega)} \tag{51}$$

thanks to the definition (40) of u_h^0 and to successive uses of Jensen's inequality and Fubini's theorem. We refer for instance to [12, Lemma 9.4] for an extension of (51) to the more complex case of non-structured grids.

Given $n \in [\![0, N_T - 1]\!]$, we multiply (41) by $(-\Delta_h u_h)_{i,j}^{n+1}$ and we sum for $i = 1, \ldots, N$ and $j = 1, \ldots, M$:

$$\begin{aligned} -\int_\Omega \frac{u_h^{n+1} - u_h^n}{\tau} \Delta_h u_h^{n+1} \,\mathrm{d}\mathbf{x} + \lambda \int_\Omega u_h^{n+1} (\Delta_h u_h^{n+1})^2 \,\mathrm{d}\mathbf{x} \\ = \sum_{C_{i,j} \in \mathcal{M}} \sum_{\sigma \in \mathcal{E}_{i,j}} v_{i,j,\sigma}^n \, u_{i,j,\sigma,+}^{n+1} (\Delta_h u_h)_{i,j}^{n+1} + \lambda \int_\Omega \mathcal{J}(u_h^{n+1}) \Delta_h u_h^{n+1} \mathrm{d}\mathbf{x}. \end{aligned} \tag{52}$$

Owing to (85) and to the convexity inequality $(b-a)b \geqslant \dfrac{b^2}{2} - \dfrac{a^2}{2}$, the first term in the left-hand side can be underestimated by

$$-\int_\Omega \frac{u_h^{n+1}-u_h^n}{\tau}\Delta_h u_h^{n+1} \geqslant \frac{1}{2\tau}\|\nabla_h u_h^{n+1}\|^2_{L^2(\Omega)} - \frac{1}{2\tau}\|\nabla_h u_h^n\|^2_{L^2(\Omega)}. \tag{53}$$

For the second term of the left-hand side, the maximum principle (see Proposition 5.1) implies that

$$\lambda\int_\Omega u_h^{n+1}(\Delta_h u_h^{n+1})^2 \mathrm{d}\mathbf{x} \geqslant \lambda\, u_\flat\, \|\Delta_h u_h^{n+1}\|^2_{L^2(\Omega)}. \tag{54}$$

For the convection term, we have two contributions corresponding respectively to the centered approximation for the convection and to the numerical diffusion stemming from the upwinding, see (98). We recall that the properties of the discrete interpolation operators are collected in Appendix A. Concerning the centered part, we deduce from (97), (99) and Lemma B.1 that

$$\sum_{C_{i,j}\in\mathcal{M}}\sum_{\sigma\in\mathcal{E}_{i,j}} v^n_{i,j,\sigma}\, u^{n+1}_{i,j,\sigma,c}(\Delta_h u_h)^{n+1}_{i,j} = -T_1^{(1,2)} - T_1^{(2,2)} - T_2^{(1,2)} - T_2^{(2,2)}, \tag{55}$$

with:

$$\begin{aligned}
T_1^{(1,2)} &= \int_\Omega \pi_x(\delta_x^* v^n_{1,h}\cdot \pi_x^*(\delta_x u_h^{n+1}))\cdot\delta_x u_h^{n+1}\, d\mathbf{x},\\
T_1^{(2,2)} &= \int_\Omega \pi_y^*(\delta_x v^n_{2,h}\cdot \pi_x(\delta_y u_h^{n+1}))\cdot\delta_x u_h^{n+1}\, d\mathbf{x},\\
T_2^{(1,2)} &= \int_\Omega \pi_x^*(\delta_y v^n_{1,h}\cdot \pi_y(\delta_x u_h^{n+1}))\cdot\delta_y u_h^{n+1}\, d\mathbf{x},
\end{aligned}$$

and

$$T_2^{(2,2)} = \int_\Omega \pi_y(\delta_y^* v^n_{2,h}\cdot \pi_y^*(\delta_y u_h^{n+1}))\cdot\delta_y u_h^{n+1}\, d\mathbf{x}.$$

The combination of Proposition A.5, standard Hölder inequalities together with Proposition A.6 yields

$$\begin{aligned}
|T_1^{(1,2)}| &\leqslant \|\delta_x^* v^n_{1,h}\|_{L^2(\Omega)}\, \|\pi_x^*(\delta_x u_h^{n+1})\|^2_{L^4(\Omega)}\\
&\leqslant \|\delta_x^* v^n_{1,h}\|_{L^2(\Omega)}\, \|\delta_x u_h^{n+1}\|^2_{L^4(\Omega)},\\
|T_1^{(2,2)}| &\leqslant \|\delta_x v^n_{2,h}\|_{L^2(\Omega)}\, \|\pi_x(\delta_y u_h^{n+1})\|_{L^4(\Omega)}\, \|\pi_y(\delta_x u_h^{n+1})\|_{L^4(\Omega)}\\
&\leqslant \|\delta_x v^n_{2,h}\|_{L^2(\Omega)}\, \|\delta_y u_h^{n+1}\|_{L^4(\Omega)}\, \|\delta_x u_h^{n+1}\|_{L^4(\Omega)},\\
|T_2^{(1,2)}| &\leqslant \|\delta_y v^n_{1,h}\|_{L^2(\Omega)}\, \|\pi_y(\delta_x u_h^{n+1})\|_{L^4(\Omega)}\, \|\pi_x(\delta_y u_h^{n+1})\|_{L^4(\Omega)}\\
&\leqslant \|\delta_y v^n_{1,h}\|_{L^2(\Omega)}\, \|\delta_x u_h^{n+1}\|_{L^4(\Omega)}\, \|\delta_y u_h^{n+1}\|_{L^4(\Omega)},\\
|T_2^{(2,2)}| &\leqslant \|\delta_y^* v^n_{2,h}\|_{L^2(\Omega)}\, \|\pi_y^*(\delta_y u_h^{n+1})\|^2_{L^4(\Omega)}\\
&\leqslant \|\delta_y^* v^n_{2,h}\|_{L^2(\Omega)}\, \|\delta_y u_h^{n+1}\|^2_{L^4(\Omega)},
\end{aligned}$$

whence the estimate

$$\begin{aligned}\sum_{C_{i,j}\in\mathcal{M}}\sum_{\sigma\in\mathcal{E}_{i,j}} & v^n_{i,j,\sigma}\, u^{n+1}_{i,j,\sigma,c}(\Delta_h u_h)^{n+1}_{i,j}\\ &\leqslant \frac{1}{2\,\lambda}\|\nabla_h \mathbf{v}^n_h\|^2_{L^2(\Omega)} + \frac{\lambda}{2}\Big(\|\delta_x u^{n+1}_h\|^2_{L^4(\Omega)} + \|\delta_y u^{n+1}_h\|^2_{L^4(\Omega)}\Big)^2\\ &\leqslant \frac{1}{2\,\lambda}\|\nabla_h \mathbf{v}^n_h\|^2_{L^2(\Omega)} + \lambda\,(C_{GN})^4\frac{(u^\sharp-u_\flat)^2}{8}\|\Delta_h\, u^{n+1}_h\|^2_{L^2(\Omega)}. \quad (56)\end{aligned}$$

For the last inequality we used the Gagliardo-Nirenberg inequality (29) applied to $u^{n+1}_h - \dfrac{u_\flat+u^\sharp}{2} \in H_{\mathcal{M}}(\Omega)$) combined with identity (86). Let us now focus on the numerical diffusion part corresponding to the second term in (98). Since

$$\big|u^{n+1}_{i,j,\sigma,+} - u^{n+1}_{i,j,\sigma,c}\big| = \frac{1}{2}|u^{n+1}_{i,j} - u^{n+1}_{i,j,\sigma}|, \qquad \forall\,\sigma\in\mathcal{E}_{i,j},$$

one can rewrite

$$\begin{aligned}A := \Big|h_x h_y \sum_{C_{i,j}\in\mathcal{M}}\sum_{\sigma\in\mathcal{E}_{i,j}} & v^n_{i,j,\sigma}\big(u^{n+1}_{i,j,\sigma,+} - u^{n+1}_{i,j,\sigma,c}\big)(\Delta_h u_h)^{n+1}_{i,j}\Big|\\ &\leqslant \frac{h_x h_y}{2}\sum_{C_{i,j}\in\mathcal{M}}\big|(\Delta_h u_h)^{n+1}_{i,j}\big|\sum_{\sigma\in\mathcal{E}_{i,j}}|v^n_{i,j,\sigma}||u^{n+1}_{i,j} - u^{n+1}_{i,j,\sigma}|\end{aligned}$$

which together with Young's inequality leads to

$$A \leqslant B_\epsilon + \frac{\epsilon}{4}\|\Delta_h u^{n+1}_h\|^2_{L^2(\Omega)},$$

with

$$B_\epsilon = \frac{1}{4\epsilon}h_x h_y \sum_{C_{i,j}\in\mathcal{M}}\left(\sum_{\sigma\in\mathcal{E}_{i,j}}|v^n_{i,j,\sigma}||u^{n+1}_{i,j} - u^{n+1}_{i,j,\sigma}|\right)^2,$$

and where $\epsilon>0$ will be fixed later on. Using now the elementary $(a+b+c+d)^2 \leqslant 4(a^2+b^2+c^2+d^2)$ and Young's inequality, we get that

$$\begin{aligned}B_\epsilon &\leqslant \frac{1}{\epsilon}h_x h_y \sum_{C_{i,j}\in\mathcal{M}}\sum_{\sigma\in\mathcal{E}_{i,j}}|v^n_{i,j,\sigma}|^2|u^{n+1}_{i,j} - u^{n+1}_{i,j,\sigma}|^2\\ &\leqslant \frac{1}{4\alpha\epsilon}h_x h_y \sum_{C_{i,j}\in\mathcal{M}}\sum_{\sigma\in\mathcal{E}_{i,j}}|v^n_{i,j,\sigma}|^4 + \frac{\alpha}{\epsilon}h_x h_y \sum_{C_{i,j}\in\mathcal{M}}\sum_{\sigma\in\mathcal{E}_{i,j}}|u^{n+1}_{i,j} - u^{n+1}_{i,j,\sigma}|^4\\ &\leqslant \frac{1}{2\alpha\epsilon}\|\mathbf{v}^n_h\|^4_{L^4(\Omega)} + \frac{2\alpha}{\epsilon}h^4\|\nabla_h u^{n+1}_h\|^4_{L^4(\Omega)}\\ &\leqslant \frac{(C_S)^4}{2\alpha\epsilon}\|\nabla_h \mathbf{v}^n_h\|^4_{L^2(\Omega)} + \frac{\alpha}{2\epsilon}(C_{GN})^4 h^4 (u^\sharp-u_\flat)^2\|\Delta_h u^{n+1}_h\|^2_{L^2(\Omega)},\end{aligned}$$

where the last inequality is a consequence of the discrete Sobolev inequality (see for instance [13])

$$\|\mathbf{v}^n_h\|_{L^4(\Omega)} \leqslant C_S\|\nabla_h \mathbf{v}^n_h\|_{L^2(\Omega)}$$

and of the discrete Gagliardo-Nirenberg inequality (29) combined with identity (86). For the parameter $\alpha > 0$, we choose $\alpha = \lambda\epsilon/(4h^4)$, so that

$$B_\epsilon \leqslant \frac{2h^4(C_S)^4}{\lambda\epsilon^2}\|\nabla_h \mathbf{v}_h^n\|_{L^2(\Omega)}^4 + \frac{\lambda}{8}(C_{GN})^4(u^\sharp - u_\flat)^2\|\Delta_h u_h^{n+1}\|_{L^2(\Omega)}^2.$$

Setting $\epsilon = \lambda(C_{GN})^4(u^\sharp - u_\flat)^2$, one gets that

$$A \leqslant \lambda\left(\frac{hC_S}{\lambda(C_{GN})^2(u^\sharp - u_\flat)}\right)^4\|\nabla_h \mathbf{v}_h^n\|_{L^2(\Omega)}^4 + \frac{3\lambda}{8}(C_{GN})^4(u^\sharp - u_\flat)^2\|\Delta_h u_h^{n+1}\|_{L^2(\Omega)}^2. \tag{57}$$

Finally, applying Cauchy-Schwartz inequality to the last term in (52) leads to

$$\lambda\int_\Omega \mathcal{J}_h(u_h^{n+1})\Delta_h u_h^{n+1}\mathrm{d}\mathbf{x} \leqslant \lambda\|\Delta_h u_h^{n+1}\|_{L^2(\Omega)}\|\mathcal{J}_h(u_h^{n+1})\|_{L^2(\Omega)}.$$

The definition (42) and (43) of $\mathcal{J}_h$ is so that

$$\begin{aligned}
\|\mathcal{J}_h(u_h^{n+1})\|_{L^2(\Omega)}^2 &= \frac{1}{h_x h_y}\sum_{C_{i,j}\in\mathcal{M}}\left(\sum_{\sigma\in\mathcal{E}_{i,j}} a_\sigma\left((u_{i,j}^{n+1} - u_{i,j,\sigma}^{n+1})^+\right)^2\right)^2 \\
&\leqslant \frac{4}{h_x h_y}\sum_{C_{i,j}\in\mathcal{M}}\sum_{\sigma\in\mathcal{E}_{i,j}} a_\sigma^2\left((u_{i,j}^{n+1} - u_{i,j,\sigma}^{n+1})^+\right)^4 \\
&\leqslant 4h_x h_y\sum_{C_{i,j}\in\mathcal{M}}\sum_{\sigma\in\mathcal{E}_{i,j}}\left(\frac{(u_{i,j}^{n+1} - u_{i,j,\sigma}^{n+1})^+}{h_\sigma}\right)^4,
\end{aligned}$$

with $h_\sigma = h_x$ if $\sigma \in \mathcal{E}^V$ and $h_\sigma = h_y$ if $\sigma \in \mathcal{E}^H$ is the distance between the cell centers $|\mathbf{x}_{i,j} - \mathbf{x}_{i,j,\sigma}|$. Due to the positive part, each edge σ is counted once in the last sum, and we deduce that

$$\|\mathcal{J}_h(u_h^{n+1})\|_{L^2(\Omega)} \leqslant 2\left(h_x h_y\sum_{\sigma\in\mathcal{E}}\left(\frac{u_{i,j}^{n+1} - u_{i,j,\sigma}^{n+1}}{h_\sigma}\right)^4\right)^{1/2} = 2\|\nabla_h u_h^{n+1}\|_{L^4(\Omega)}^2.$$

Applying again (29) and (86), one gets that

$$\|\mathcal{J}_h(u_h^{n+1})\|_{L^2(\Omega)} \leqslant 2(C_{GN})^2(u^\sharp - u_\flat)\|\Delta_h u_h^{n+1}\|_{L^2(\Omega)}, \tag{58}$$

and then that

$$\lambda\int_\Omega \mathcal{J}_h(u_h^{n+1})\Delta_h u_h^{n+1}\mathrm{d}\mathbf{x} \leqslant 2\lambda(C_{GN})^2(u^\sharp - u_\flat)\|\Delta_h u_h^{n+1}\|_{L^2(\Omega)}^2. \tag{59}$$

Eventually, combining (53)–(59) in (52) leads to

$$\begin{aligned}
&\frac{1}{2\tau}\left(\|\nabla_h u_h^{n+1}\|_{L^2(\Omega)}^2 - \|\nabla_h u_h^n\|_{L^2(\Omega)}^2\right) \\
&+\lambda\|\Delta_h u_h^{n+1}\|_{L^2(\Omega)}^2\left(u_\flat - \frac{1}{2}(C_{GN})^4\delta^2 - 2(C_{GN})^2\delta\right) \leqslant \left(\frac{1}{2\lambda} + C_1\left(\frac{h}{\delta}\right)^4\right)\|\nabla_h \mathbf{v}_h^n\|_{L^2(\Omega)}^4
\end{aligned}$$

with C_1 depending only on Ω (via C_{GN} and C_S) and on λ. For $\delta < \frac{2}{(C_{GN})^2}(\sqrt{1+\frac{u_\flat}{2}}-1)$, the term in front of $\|\Delta_h u_h^{n+1}\|^2$ is positive. Owing to Lemma B.2, the above right-hand side is bounded by some quantity not depending on the mesh.

□

From the above estimate, we deduce a uniform $L^2(Q_T)$ estimate on $\delta_\tau u_{h\tau} \in L^\infty(0,T;H_{\mathcal{M}})$ defined by

$$\delta_\tau u_{h\tau}(\cdot,t) = \frac{u_h^{n+1}-u_h^n}{\tau} \quad \text{if } t \in [t^n, t^{n+1}). \tag{60}$$

Then the following estimate directly follows from the use of the estimates of Proposition 5.2 in the scheme (41).

Corollary 5.3. *Under the assumptions of Proposition 5.2, there exists $C \geqslant 0$ depending only on Ω, u_0, $\mathbf{v}$, λ, δ and T such that*

$$\iint_{Q_T} |\delta_\tau u_{h\tau}|^2 \,\mathrm{d}\mathbf{x}\mathrm{d}t \leqslant C.$$

6 Convergence of the finite volume scheme

The purpose of this section is to establish the convergence of the scheme thanks to compactness arguments. Given $(\mathcal{M}_m)_{m\geqslant 0}$ a sequence of admissible meshes with size h_m tending to 0 as m tends to $+\infty$, and given $(\tau_m)_{m\geqslant 0}$ be a sequence of positive time steps tending to 0, then denoting by $(u_{h_m\tau_m})_{m\geqslant 0}$ the corresponding sequence of approximate solution provided by Proposition 5.1, then one aims to show that, up to the extraction of a subsequence, $u_{h_m\tau_m}$ tends to a weak solution u to (30) in the sense of Definition 4.1. Our proof is based on compactness arguments. We start in Section 6.1 to establish some compactness properties on the approximate solutions $(u_{h_m\tau_m})_m$, then the limit value will be identified as a weak solution to the problem in Section 6.2.

6.1 Some compactness properties

First, it follows from Proposition 5.1 that there exists $u \in L^\infty(Q_T)$ with $u_\flat \leqslant u \leqslant u^\sharp$ such that, up to a subsequence, there holds

$$u_{h_m\tau_m} \underset{m\to+\infty}{\longrightarrow} u \quad \text{in the } L^\infty(Q_T)\text{-weak-}\star \text{ sense.} \tag{61}$$

Moreover, thanks to the (uniform w.r.t. m) $L^2(Q_T)$ bounds on $\nabla_{h_m} u_{h_m\tau_m}$ and $\delta_{\tau_m} u_{h_m\tau_m}$ respectively established in Proposition 5.2 and Corollary 5.3, we can mimic the technics detailed in [12] to get estimates on the space-and-time translates

$$\int_0^{T-\zeta}\int_{\Omega_{\boldsymbol{\xi}}} |u_{h_m\tau_m}(\mathbf{x}+\boldsymbol{\xi},t+\zeta)-u_{h_m\tau_m}(\mathbf{x},t)|^2 \mathrm{d}\mathbf{x}\mathrm{d}t \leqslant C\left(\zeta^2+|\boldsymbol{\xi}|^2\right), \quad \zeta\in(0,T),\ \boldsymbol{\xi}\in\mathbb{R}^2,$$

with C not depending on m, and with $\Omega_{\boldsymbol{\xi}} = \{\mathbf{x} \in \Omega \mid \mathbf{x} + \boldsymbol{\xi} \in \Omega\}$. This in particular yields the relative compactness of the sequence $(u_{h_m\tau_m})_m$ in $L^2(Q_T)$ thanks to Kolmogorov's compactness criterion. Therefore, up to the extraction of a subsequence, we get that

$$u_{h_m\tau_m} \underset{m\to+\infty}{\longrightarrow} u \quad \text{a.e. in } Q_T. \tag{62}$$

Besides, we deduce from Estimate (50) that, still up to a subsequence, there holds

$$\Delta_{h_m} u_{h_m\tau_m} \underset{m\to+\infty}{\longrightarrow} \Delta u \quad \text{weakly in } L^2(Q_T). \tag{63}$$

Indeed, the (uniform w.r.t. m) $L^2(Q_T)$ bound on $\Delta_{h_m} u_{h_m\tau_m}$ ensures the existence of some weak limit $\mathfrak{d} \in L^2(Q_T)$. Then following the program of [12], the identification of $\mathfrak{d} = \Delta u$ is then obtained in the distributional sense. Similarly, we deduce from Corollary 5.3 that

$$\delta_{\tau_m} u_{h_m\tau_m} \underset{m\to+\infty}{\longrightarrow} \partial_t u \quad \text{weakly in } L^2(Q_T). \tag{64}$$

Concerning the sequence $(\nabla_{h_m} u_{h_m\tau_m})_m$, we have the uniform $L^\infty(0,T;L^2(\Omega))^2$ estimate (49) as well as a $L^4(Q_T)^2$ estimate stemming from the combination of (29), (45) and (50). After identifying the weak limit in the distributional sense once again, one gets that

$$\nabla_{h_m} u_{h_m\tau_m} \underset{m\to+\infty}{\longrightarrow} \nabla u \quad \text{in the } L^4(Q_T)\text{-weak and } L^\infty(0,T;L^2(\Omega))^2\text{-weak-}\star \text{ senses.} \tag{65}$$

Further compactness is required to pass in the limit in the Joule effect term $\mathcal{J}_{h_m}(u_{h_m\tau_m})$, whence next lemma.

Lemma 6.1. *Up to extraction of a subsequence, the following convergence holds:*

$$\nabla_{h_m} u_{h_m\tau_m} \underset{m\to+\infty}{\longrightarrow} \nabla u \quad \textit{a.e. in } Q_T.$$

Proof. The proof relies on some discrete Aubin-Lions-Simon lemma. In the proof, we make use of the result presented in [1] but we stress that a proof building on [17, 16] is also possible.

We proceed direction-wise, proving that $\delta_x u_{h_m\tau_m} \in \widehat{H}^0_{\mathcal{M}_m}(\Omega)$ converges pointwise towards $\partial_x u$. Of course, proving the convergence of $\delta_y u_{h_m\tau_m} \in \widetilde{H}^0_{\mathcal{M}_m}(\Omega)$ towards $\partial_y u$ is similar.

The combination of Estimate (50) with identity (86) provides that

$$\|\nabla_{h_m} \delta_x u_{h_m\tau_m}\|_{L^2(Q_T)^2} \leqslant C$$

for some C not depending on m, providing some compactness with respect to the space variable on $(\delta_x u_{h_m\tau_m})_m$. On the other hand, given $\varphi \in C^\infty_c(Q_T)$, and denoting by $\widetilde{\varphi}_{h_m\tau_m}$ the piecewise constant in time and space function defined by

$$\widetilde{\varphi}_{h_m\tau_m}(\mathbf{x},t) = \frac{1}{\tau_m h_{x,m} h_{y,m}} \int_{n\tau_m}^{(n+1)\tau_m} \int_{C_{i+1/2,j}} \varphi \,\mathrm{d}\mathbf{x}\mathrm{d}t,$$

then

$$\iint_{Q_T} \delta_x \delta_{\tau_m} u_{h_m\tau_m} \varphi \,\mathrm{d}\mathbf{x}\mathrm{d}t = \iint_{Q_T} \delta_x \delta_{\tau_m} u_{h_m\tau_m} \widetilde{\varphi}_{h_m\tau_m} \,\mathrm{d}\mathbf{x}\mathrm{d}t = -\iint_{Q_T} \delta_{\tau_m} u_{h_m\tau_m} \delta_x^* \widetilde{\varphi}_{h_m\tau_m} \,\mathrm{d}\mathbf{x}\mathrm{d}t.$$

Applying Cauchy-Schwarz inequality and using Corollary 5.3 and [12, Lemma 9.4] yields

$$\iint_{Q_T} \delta_x \delta_{\tau_m} u_{h_m\tau_m} \varphi \,\mathrm{d}\mathbf{x}\mathrm{d}t \leqslant C \|\nabla_{h_m} \widetilde{\varphi}_{h_m\tau_m}\|_{L^2(Q_T)^2} \leqslant C \|\nabla\varphi\|_{L^2(Q_T)^2}.$$

We can thus apply Theorem 3.9 of [1] which ensures that $\delta_x u_{h_m\tau_m}$ converges pointwise. Because of (65), the limit is $\partial_x u$. □

Notice that thanks to (65) and Lemma 6.1, we can apply Vitali's convergence theorem and claim that

$$\nabla_{h_m} u_{h_m\tau_m} \underset{m\to+\infty}{\longrightarrow} \nabla u \quad \text{in } L^2(Q_T). \tag{66}$$

With the above compactness properties, we have enough material to pass to the limit in the scheme. This is the purpose of next section.

6.2 Limits are weak solutions

As a preliminary to the identification of any limit value u of $u_{h_m\tau_m}$ as in Section 6.1 as a weak solution, let us first show the consistency of the discretization (42) and (43) of the Joule term.

Lemma 6.2. *Up to a subsequence there holds*

$$\mathcal{J}_{h_m}(u_{h_m\tau_m}) \underset{m\to+\infty}{\longrightarrow} |\nabla u|^2 \quad \textit{weakly in } L^2(Q_T).$$

Proof. First, it follows from (66) that

$$|\nabla_{h_m} u_{h_m\tau_m}|^2 \underset{m\to+\infty}{\longrightarrow} |\nabla u|^2 \quad \text{in } L^1(Q_T). \tag{67}$$

On the other hand, we deduce from (50) that there exists some $\mathfrak{J} \in L^2(Q_T)$ such that

$$\mathcal{J}_{h_m}(u_{h_m\tau_m}) \underset{m\to+\infty}{\longrightarrow} \mathfrak{J} \quad \text{weakly in } L^2(Q_T).$$

Let us now identify $\mathfrak{J}$ as $|\nabla u|^2$ in the distributional sense. Let $\varphi \in C_c^\infty(Q_T)$, and define $\varphi_{h_m\tau_m}$ by setting $\varphi_{i,j}^n = \varphi(\mathbf{x}_{i,j}, t^n)$ for all $C_{i,j} \in \mathcal{M}_m$ and all $n \in [\![1, N_{T,m}]\!]$, then

$$\iint_{Q_T} \left(\mathcal{J}_{h_m}(u_{h_m\tau_m}) - |\nabla_{h_m} u_{h_m\tau_m}|^2\right) \varphi \,\mathrm{d}\mathbf{x}\mathrm{d}t \leqslant \mathcal{R}_m(\varphi) + \mathcal{S}_m(\varphi), \tag{68}$$

with

$$\mathcal{R}_m(\varphi) = \iint_{Q_T} \left(\mathcal{J}_{h_m}(u_{h_m\tau_m}) + |\nabla_{h_m} u_{h_m\tau_m}|^2\right) |\varphi_{h_m\tau_m} - \varphi| \,\mathrm{d}\mathbf{x}\mathrm{d}t$$

and

$$\mathcal{S}_m(\varphi) = \iint_{Q_T} \left(\mathcal{J}_{h_m}(u_{h_m\tau_m}) - |\nabla_{h_m} u_{h_m\tau_m}|^2 \right) \varphi_{h_m\tau_m} \, \mathrm{d}\mathbf{x}\mathrm{d}t$$

Due to the regularity of φ and to the boundedness in $L^1(Q_T)$ of $\mathcal{J}_{h_m}(u_{h_m\tau_m})$ and $|\nabla_{h_m} u_{h_m\tau_m}|^2$, we infer that

$$|\mathcal{R}_m(\varphi)| \leqslant C(h_m + \tau_m) \underset{m\to+\infty}{\longrightarrow} 0. \tag{69}$$

Since

$$\int_{C_{i,j}} |\nabla_{h_m} u^n_{h_m}|^2 \mathrm{d}\mathbf{x} = \frac{1}{2} \sum_{\sigma\in\mathcal{E}_{i,j}} a_\sigma \left(u^n_{i,j} - u^n_{i,j,\sigma} \right)^2,$$

one deduces from the definition (43) of $\mathcal{J}_{h_m}(u_{h_m\tau_m})$ that

$$\begin{aligned}
|\mathcal{S}_m(\varphi)| &= \left| \sum_{n=1}^{N_{T,m}} \tau_m \sum_{C_{i,j}\in\mathcal{M}_m} \varphi^n_{i,j} \sum_{\sigma\in\mathcal{E}_{i,j}} a_\sigma \left[\left(\left(u^n_{i,j,\sigma} - u^n_{i,j} \right)^+ \right)^2 - \frac{1}{2} \left(u^n_{i,j,\sigma} - u^n_{i,j} \right)^2 \right] \right| \\
&\leqslant \frac{1}{2} \sum_{n=1}^{N_{T,m}} \tau_m \sum_{C_{i,j}\in\mathcal{M}_m} \sum_{\sigma\in\mathcal{E}_m} a_\sigma \left(u^n_{i,j,\sigma} - u^n_{i,j} \right)^2 \left| \varphi^n_{i,j} - \varphi^n_{i,j,\sigma} \right| \\
&\leqslant Ch_m \sum_{n=1}^{N_{T,m}} \tau_m \sum_{C_{i,j}\in\mathcal{M}_m} \sum_{\sigma\in\mathcal{E}_m} a_\sigma \left(u^n_{i,j,\sigma} - u^n_{i,j} \right)^2 \leqslant Ch_m \underset{m\to+\infty}{\longrightarrow} 0,
\end{aligned} \tag{70}$$

the last inequality being a consequence of (49). Then we deduce from (67)–(70) that $\mathfrak{J} = |\nabla u|^2$, concluding the proof of Lemma 6.2. □

Our next lemma is about the boundary conditions for the limit u, which is shown to belong to $L^2(0,T;H^2_N(\Omega))$.

Lemma 6.3. *Let u be a limit of $(u_{h_m\tau_m})_m$ as in Section 6.1, then $\nabla u \cdot \mathbf{n} = 0$ on $\partial\Omega \times (0,T)$.*

Proof. First note that since Δu belongs to $L^2(Q_T)$, cf. (63) and since Ω is convex, then u belongs to $L^2(0,T;H^2(\Omega))$ and $\nabla u \in L^2(0,T;H^1(\Omega))^2$ admits strong traces in $L^2(0,T;H^{1/2}(\partial\Omega))^2 \subset L^2(\partial\Omega\times(0,T))^2$. Let us show that $-\partial_x u = \nabla u \cdot \mathbf{n} = 0$ on $\left(\{\underline{x}\} \times (\underline{y},\overline{y})\right) \times (0,T)$, the treatment of the other parts of the boundary being similar.

We proceed as in [4, Section 4.2]. Fix $\epsilon > 0$, then the triangle inequality gives

$$\int_0^T \int_{\underline{y}}^{\overline{y}} |\partial_x u((\underline{x},y),t)| \mathrm{d}y\mathrm{d}t = \int_0^T \frac{1}{\epsilon} \int_0^\epsilon \int_{\underline{y}}^{\overline{y}} |\partial_x u((\underline{x},y),t)| \mathrm{d}y\mathrm{d}s\mathrm{d}t \leqslant A^\epsilon + B^\epsilon_m + C^\epsilon_m, \tag{71}$$

with

$$\begin{aligned}
A^\epsilon &= \int_0^T \frac{1}{\epsilon} \int_0^\epsilon \int_{\underline{y}}^{\overline{y}} |\partial_x u((\underline{x},y),t) - \partial_x u((\underline{x}+s,y),t)| \mathrm{d}y\mathrm{d}s\mathrm{d}t, \\
B^\epsilon_m &= \int_0^T \frac{1}{\epsilon} \int_0^\epsilon \int_{\underline{y}}^{\overline{y}} |\partial_x u((\underline{x}+s,y),t) - \delta_x u_{h_m\tau_m}((\underline{x}+s,y),t)| \mathrm{d}y\mathrm{d}s\mathrm{d}t, \\
C^\epsilon_m &= \int_0^T \frac{1}{\epsilon} \int_0^\epsilon \int_{\underline{y}}^{\overline{y}} |\delta_x u_{h_m\tau_m}((\underline{x}+s,y),t)| \mathrm{d}y\mathrm{d}s\mathrm{d}t.
\end{aligned}$$

Applying Cauchy-Schwarz inequality, one gets that

$$\begin{aligned} A^\epsilon &\leqslant \int_0^T \frac{1}{\epsilon}\int_0^\epsilon \int_{\underline{y}}^{\overline{y}}\int_0^s |\partial_{xx}u((\underline{x}+r,y),t)|\,\mathrm{d}r\mathrm{d}y\mathrm{d}s\mathrm{d}t \\ &\leqslant \left(\int_0^T \frac{1}{\epsilon}\int_0^\epsilon \int_{\underline{y}}^{\overline{y}}\int_0^s |\partial_{xx}u((\underline{x}+r,y),t)|^2\,\mathrm{d}r\mathrm{d}y\mathrm{d}s\mathrm{d}t\right)^{1/2}\sqrt{\frac{T(\overline{y}-\underline{y})\epsilon}{2}}. \end{aligned}$$

We then infer from the lower semi-continuity of the norm for the weak convergence (63) that

$$A^\epsilon \leqslant C\sqrt{\epsilon}, \qquad \forall\,\epsilon>0. \tag{72}$$

For the second term B_m^ϵ, the strong convergence of $\delta_x u_{h_m\tau_m}$ towards $\partial_x u$ in $L^2(Q_T)$ (thus also in $L^1(Q_T)$) stated in (66) implies that

$$\lim_{m\to+\infty} B_m^\epsilon = 0, \qquad \forall\,\epsilon>0. \tag{73}$$

For the third term C_m^ϵ, we use the fact that $u_{1,j}^n = u_{0,j}^n$ to write that

$$\begin{aligned} C_m^\epsilon &\leqslant \frac{1}{\epsilon}\sum_{n=1}^{N_{T,m}}\tau_m\sum_{j=1}^{M_m}h_{y,m}\sum_{i=0}^{\lceil\frac{\epsilon}{h_{x,m}}\rceil}|u_{i+1,j}^n-u_{i,j}^n| \\ &\leqslant \frac{1}{\epsilon}\sum_{n=1}^{N_{T,m}}\tau_m\sum_{j=1}^{M_m}h_{y,m}\sum_{i=0}^{\lceil\frac{\epsilon}{h_{x,m}}\rceil}\left|\sum_{\ell=1}^{i}(u_{\ell+1,j}^n-2u_{\ell,j}^n+u_{\ell-1,j}^n)\right| \\ &\leqslant \frac{1}{\epsilon}\sum_{n=1}^{N_{T,m}}\tau_m\sum_{j=1}^{M_m}h_{y,m}\sum_{i=0}^{\lceil\frac{\epsilon}{h_{x,m}}\rceil}\sum_{\ell=1}^{i}\left|u_{\ell+1,j}^n-2u_{\ell,j}^n+u_{\ell-1,j}^n\right|. \end{aligned}$$

Then Cauchy-Schwarz inequality provides that

$$\begin{aligned} (C_m^\epsilon)^2 \leqslant &\left(\frac{1}{\epsilon}\sum_{n=1}^{N_{T,m}}\tau_m\sum_{j=1}^{M_m}h_{y,m}h_{x,m}^2\sum_{i=0}^{\lceil\frac{\epsilon}{h_{x,m}}\rceil}\sum_{\ell=1}^{i}\left(\frac{u_{\ell+1,j}^n-2u_{\ell,j}^n+u_{\ell-1,j}^n}{h_{x,m}^2}\right)^2\right) \\ &\times\left(\frac{1}{\epsilon}\sum_{n=1}^{N_{T,m}}\tau_m\sum_{j=1}^{M_m}h_{y,m}\sum_{i=0}^{\lceil\frac{\epsilon}{h_{x,m}}\rceil}\sum_{\ell=1}^{i}h_{x,m}^2\right) =: C_m^{\epsilon,(1)}\times C_m^{\epsilon,(2)}. \end{aligned} \tag{74}$$

The first term in the above right-hand side can be overestimated by

$$C_m^{\epsilon,(1)} \leqslant \frac{\epsilon+h_{x,m}}{\epsilon}\iint_{Q_T}|\delta_x^*\delta_x u_{h_m\tau_m}|^2\,\mathrm{d}\mathbf{x}\mathrm{d}t \leqslant \left(1+\frac{h_{x,m}}{\epsilon}\right)\|\Delta_h u_{h_m\tau_m}\|_{L^2(Q_T)}^2.$$

Hence we infer from (50) that

$$\limsup_{m\to+\infty} C_m^{\epsilon,(1)} < +\infty. \tag{75}$$

Concerning $C_m^{\epsilon,(2)}$, we first use the elementary inequality $\lceil \frac{a}{b} \rceil b \leqslant a + b$ to write

$$\sum_{i=0}^{\lceil \frac{\epsilon}{h_{x,m}} \rceil} \sum_{\ell=1}^{i} h_{x,m}^2 \leqslant \frac{1}{2}(\epsilon + h_{x,m})(\epsilon + 2h_{x,m}).$$

Therefore, we obtain that

$$C_m^{\epsilon,(2)} \leqslant \epsilon T(\overline{y} - \underline{y}) \left(1 + \frac{h_{x,m}}{\epsilon}\right) \left(\frac{1}{2} + \frac{h_{x,m}}{\epsilon}\right),$$

and thus that

$$\limsup_{m \to +\infty} C_m^{\epsilon,(2)} \leqslant C\epsilon. \tag{76}$$

Combining (72)–(76) in (71) yields

$$\int_0^T \int_{\underline{y}}^{\overline{y}} |\partial_x u((\underline{x}, y), t)| \mathrm{d}y \mathrm{d}t \leqslant C\sqrt{\epsilon}$$

for any $\epsilon > 0$, whence the desired result. □

It only remains to prove the following proposition to conclude the proof of Theorem 4.4.

Proposition 6.4. *Let u be a limit value of $(u_{h_m \tau_m})_m$ as in Section 6.1, then u is a weak solution to* (30) *in the sense of Definition 4.1.*

Proof. With Lemma 6.3 in addition to the compactness result stated in Section 6.1, we know that the limit u belongs to the right space to be a weak solution. We only have to check that (30a) is fulfilled by the limit u. To this end, denote by

$$\mathcal{W}_{i,j}^{n+1} = \frac{1}{h_{x,m} h_{y,m}} \sum_{\sigma \in \mathcal{E}_{i,j}} v_{i,j,\sigma}^n \, u_{i,j,\sigma,+}^{n+1}$$

for $i \in [\![1, N_m]\!]$, $j \in [\![1, M_m]\!]$, and $n \in [\![1, N_{T,m}]\!]$, then define $\mathcal{W}_{h_m \tau_m}$ by

$$\mathcal{W}_{h_m \tau_m}(\mathbf{x}, t) = \mathcal{W}_{i,j}^n \quad \text{if } (\mathbf{x}, t) \in C_{i,j} \times (t^{n-1}, t^n].$$

The scheme (41) then rewrites under the compact form

$$\delta_{\tau_m} u_{h_m \tau_m} + \mathcal{W}_{h_m \tau_m} + \lambda \left(\mathcal{J}_{h_m}(u_{h_m \tau_m}) - u_{h_m \tau_m} \Delta_{h_m} u_{h_m \tau_m} \right) = 0. \tag{77}$$

Due to (61), (62) and (63), we can easily check that

$$u_{h_m \tau_m} \Delta_{h_m} u_{h_m \tau_m} \underset{m \to +\infty}{\longrightarrow} u \Delta u \quad \text{weakly in } L^2(Q_T).$$

Bearing in mind (64) and Lemma 6.2, it only remains to prove the

$$\mathcal{W}_{h_m \tau_m} \underset{m \to +\infty}{\longrightarrow} \nabla \cdot (u\mathbf{v}) = \mathbf{v} \cdot \nabla u \quad \text{weakly in } L^2(Q_T), \tag{78}$$

to pass to the limit in (77) to recover (30a).

Since $\mathbf{v}_h^n$ is divergence free, one has $\sum_{\sigma\in\mathcal{E}_{i,j}} v_{i,j,\sigma}^n = 0$, so that $\mathcal{W}_{i,j}^{n+1}$ rewrites

$$\mathcal{W}_{i,j}^{n+1} = \frac{1}{h_{x,m}h_{y,m}} \sum_{\sigma\in\mathcal{E}_{i,j}} v_{i,j,\sigma}^n \, (u_{i,j,\sigma,+}^{n+1} - u_{i,j}^{n+1}).$$

Then it follows from the definition of $u_{i,j,\sigma,+}^{n+1}$ that

$$v_{i,j,\sigma}^n \, (u_{i,j,\sigma,+}^{n+1} - u_{i,j}^{n+1}) = (v_{i,j,\sigma}^n)^- (u_{i,j}^{n+1} - u_{i,j,\sigma}^{n+1}),$$

where $(v_{i,j,\sigma}^n)^- = \max(0, -v_{i,j,\sigma}^n)$ denotes the negative part of $v_{i,j,\sigma}^n$.

$$\begin{aligned}\iint_{Q_T} |\mathcal{W}_{h_m\tau_m}|^2 \mathrm{d}\mathbf{x}\mathrm{d}t &= \sum_{n=1}^{N_{T,m}} \tau_m \sum_{C_{i,j}\in\mathcal{M}_m} \frac{1}{h_{x,m}h_{y,m}} \left(\sum_{\sigma\in\mathcal{E}_{i,j}} (v_{i,j,\sigma}^{n-1})^- \, (u_{i,j}^n - u_{i,j,\sigma}^n) \right)^2 \\ &\leqslant 4 \sum_{n=1}^{N_{T,m}} \tau_m \sum_{C_{i,j}\in\mathcal{M}_m} \frac{1}{h_{x,m}h_{y,m}} \sum_{\sigma\in\mathcal{E}_{i,j}} \left((v_{i,j,\sigma}^{n-1})^-\right)^2 \left((u_{i,j}^n - u_{i,j,\sigma}^n)\right)^2.\end{aligned}$$

Bearing in mind the definition of $v_{i,j,\sigma}^n$, we get that

$$\iint_{Q_T} |\mathcal{W}_{h_m\tau_m}|^2 \mathrm{d}\mathbf{x}\mathrm{d}t \leqslant 2 \sum_{n=1}^{N_{T,m}} \tau_m \|\mathbf{v}_h^n\|_\infty^2 \sum_{C_{i,j}\in\mathcal{M}_m} \sum_{\sigma\in\mathcal{E}_{i,j}} a_\sigma (u_{i,j}^n - u_{i,j,\sigma}^n)^2.$$

Making use of (49) and (39), one eventually gets that

$$\iint_{Q_T} |\mathcal{W}_{h_m\tau_m}|^2 \mathrm{d}\mathbf{x}\mathrm{d}t \leqslant C.$$

In particular, there exists some $\mathcal{W}^* \in L^2(Q_T)$ such that

$$\mathcal{W}_{h_m\tau_m} \underset{m\to+\infty}{\longrightarrow} \mathcal{W}^* \quad \text{weakly in } L^2(Q_T). \tag{79}$$

Let us identify $\mathcal{W}^*$ as $\mathbf{v}\cdot\nabla u$ in the distributional sense. Let $\varphi \in C_c^\infty(Q_T)$, then defining $\varphi_{h_m\tau_m}$ as the piecewise constant in time and space function built from the cell values $\varphi_{i,j}^n = \varphi(\mathbf{x}_{i,j}, t^n)$, then

$$\left| \iint_{Q_T} \mathcal{W}_{h_m\tau_m} (\varphi - \varphi_{h_m\tau_m}) \, \mathrm{d}\mathbf{x}\mathrm{d}t \right| \leqslant \iint_{Q_T} |\mathcal{W}_{h_m\tau_m}| \, |\varphi - \varphi_{h_m\tau_m}| \mathrm{d}\mathbf{x}\mathrm{d}t. \tag{80}$$

Due to the regularity of φ and the boundedness in $L^2(Q_T)$ of $\mathcal{W}_{h_m\tau_m}$, the above right-hand side tends to 0 as m tends to $+\infty$. We write

$$\iint_{Q_T} \mathcal{W}_{h_m\tau_m} \varphi_{h_m\tau_m} \mathrm{d}\mathbf{x}\mathrm{d}t = \mathcal{T}_m^{(1)}(\varphi) + \mathcal{T}_m^{(2)}(\varphi)$$

with

$$\begin{aligned}\mathcal{T}_m^{(1)}(\varphi) &= \frac{1}{2}\sum_{n=1}^{N_{T,m}}\tau_m\sum_{C_{i,j}\in\mathcal{M}_m}\varphi_{i,j}^n\sum_{\sigma\in\mathcal{E}_{i,j}}v_{i,j,\sigma}^{n-1}\left(u_{i,j,\sigma}^n-u_{i,j}^n\right)\\ &= \iint_{Q_T}\mathbf{v}_{h_m\tau_m}\cdot\nabla_{h_m}u_{h_m\tau_m}\,\varphi_{h_m\tau_m}\mathrm{d}\mathbf{x}\mathrm{d}t,\end{aligned}$$

and

$$\mathcal{T}_m^{(2)}(\varphi)=\frac{1}{2}\sum_{n=1}^{N_{T,m}}\tau_m\sum_{\sigma\in\mathcal{E}_m}v_{i,j,\sigma}^{n-1}\left(u_{i,j}^n-u_{i,j,\sigma}^n\right)\left(\varphi_{i,j}^n-\varphi_{i,j,\sigma}^n\right).$$

In particular, by similar arguments to those which lead to (79) combined with the regularity of φ, one gets that

$$\left|\mathcal{T}_m^{(2)}(\varphi)\right|\leqslant C(h_m+\tau_m). \tag{81}$$

Besides, mimicking the calculations in the proof of Lemma B.2, one readily shows that $\mathbf{v}_{h_m\tau_m}$ tends to $\mathbf{v}$ in $L^2(Q_T)$ as m goes to $+\infty$. Together with (65) and with the uniform convergence of $\varphi_{h_m\tau_m}$ towards φ stemming from its regularity, this allows to pass to the limit in $\mathcal{T}_m^{(1)}(\varphi)$:

$$\mathcal{T}_m^{(1)}(\varphi)\underset{m\to+\infty}{\longrightarrow}\iint_{Q_T}\mathbf{v}\cdot\nabla u\,\varphi\,\mathrm{d}\mathbf{x}\mathrm{d}t,$$

concluding the proof of (78) and thus of Proposition 6.4. □

Acknowledgments

This work was supported in part by the Labex CEMPI (ANR-11-LABX-0007-01).

A Properties of discrete differential and interpolation operators for the 2D case

We give some properties related to the differential and interpolation operators.

Proposition A.1. *We have the following properties:*

$$\begin{aligned}&\text{For any } v\in\widehat{H}_{\mathcal{M}}(\Omega),\quad(\delta_y\circ\delta_x^*)\,v=(\delta_x^*\circ\delta_y)\,v.\\ &\text{For any } v\in\widetilde{H}_{\mathcal{M}}(\Omega),\quad(\delta_x\circ\delta_y^*)\,v=(\delta_y^*\circ\delta_x)\,v.\end{aligned} \tag{82}$$

Proof. The proof is direct from the above definitions of spaces and operators. □

Proposition A.2. *The following discrete Stokes formula hold:*

$$\begin{aligned}
\text{For any } (v,w) \in H_{\mathcal{M}}(\Omega) \times \widehat{H}^0_{\mathcal{M}}(\Omega), \quad \int_\Omega v\ \delta_x^* w &= -\int_\Omega \delta_x v\ w. \\
\text{For any } (v,w) \in H_{\mathcal{M}}(\Omega) \times \widetilde{H}^0_{\mathcal{M}}(\Omega), \quad \int_\Omega v\ \delta_y^* w &= -\int_\Omega \delta_y v\ w. \qquad (83) \\
\text{For any } (v,w) \in \bar{H}^{x,0}_{\mathcal{M}}(\Omega) \times \widetilde{H}_{\mathcal{M}}(\Omega), \quad \int_\Omega \delta_x^* v\ w &= -\int_\Omega v\ \delta_x w. \qquad (84) \\
\text{For any } (v,w) \in \bar{H}^{y,0}_{\mathcal{M}}(\Omega) \times \widehat{H}_{\mathcal{M}}(\Omega), \quad \int_\Omega \delta_y^* v\ w &= -\int_\Omega v\ \delta_y w. \\
\text{For any } (v,w) \in (H_{\mathcal{M}}(\Omega))^2, \quad \int_\Omega v\ \Delta_h w &= -\int_\Omega \nabla_h v\ \cdot\ \nabla_h w. \qquad (85)
\end{aligned}$$

Proof. The proof is direct from the above definitions of spaces and operators. □

Consequently, we can also verify the following norms equality:

Proposition A.3. *For any $v \in H_{\mathcal{M}}(\Omega)$, we have:*

$$\|\nabla_h^2 v\|_{L^2(\Omega)} = \|\Delta_h v\|_{L^2(\Omega)}. \qquad (86)$$

Proof. Let us consider $v \in H_{\mathcal{M}}(\Omega)$. We write:

$$\begin{aligned}
\int_\Omega \delta_{xx} v\ \delta_{yy} v\, \mathrm{d}\mathbf{x} &= \int_\Omega (\delta_x^* \circ \delta_x) v\ (\delta_y^* \circ \delta_y) v\, \mathrm{d}\mathbf{x} \\
&\overset{(83)}{=} -\int_\Omega (\delta_y \circ \delta_x^* \circ \delta_x) v\ \delta_y v\, \mathrm{d}\mathbf{x} \\
&\overset{(82)}{=} -\int_\Omega (\delta_x^* \circ \delta_y \circ \delta_x) v\ \delta_y v\, \mathrm{d}\mathbf{x} \\
&\overset{(84)}{=} \int_\Omega (\delta_y \circ \delta_x) v\ (\delta_x \circ \delta_y) v\, \mathrm{d}\mathbf{x} \\
&\overset{(5)}{=} \|\delta_{xy} v\|^2_{L^2(\Omega)} \overset{(5)}{=} \|\delta_{yx} v\|^2_{L^2(\Omega)}.
\end{aligned}$$

Consequently, from the definitions (6) and (7), (86) holds. □

Proposition A.4. *Let assume that the mesh $\mathcal{M}$ is uniform in each direction in the sense of Definition 4.3. Then:*

$$\begin{aligned}
\text{For any } v \in \widehat{H}_{\mathcal{M}}(\Omega), \quad (\delta_x \circ \pi_x^*) v &= (\pi_x \circ \delta_x^*) v, \qquad (87) \\
\text{For any } v \in \widetilde{H}_{\mathcal{M}}(\Omega), \quad (\delta_x \circ \pi_y^*) v &= (\pi_y^* \circ \delta_x) v, \qquad (88) \\
\text{For any } v \in \widehat{H}_{\mathcal{M}}(\Omega), \quad (\delta_x^* \circ \pi_y) v &= (\pi_y \circ \delta_x^*) v, \\
\text{For any } v \in \widetilde{H}_{\mathcal{M}}(\Omega), \quad (\delta_y^* \circ \pi_x) v &= (\pi_x \circ \delta_y^*) v, \qquad (89) \\
\text{For any } (v,w) \in (\widehat{H}_{\mathcal{M}}(\Omega))^2, \quad \delta_x^*(vw) &= \pi_x^* v\, \delta_x^* w\ +\ \delta_x^* v\, \pi_x^* w \qquad (90) \\
\text{For any } (v,w) \in (\widetilde{H}_{\mathcal{M}}(\Omega))^2, \quad \delta_x(vw) &= \pi_x v\, \delta_x w\ +\ \delta_x v\, \pi_x w \qquad (91)
\end{aligned}$$

Proof. The proof is direct from the above definitions of spaces and operators. □

Proposition A.5. *Let assume that the mesh $\mathcal{M}$ is uniform in each direction in the sense of Definition 4.3. Then:*

$$\text{For any } (v,w) \in H_{\mathcal{M}}(\Omega) \times \widehat{H}_{\mathcal{M}}(\Omega), \quad \int_\Omega \pi_x(v)\, w\, d\mathbf{x} = \int_\Omega v\, \pi_x^*(w)\, d\mathbf{x} \tag{92}$$

$$\text{For any } (v,w) \in H_{\mathcal{M}}(\Omega) \times \widetilde{H}_{\mathcal{M}}(\Omega), \quad \int_\Omega \pi_y(v)\, w\, d\mathbf{x} = \int_\Omega v\, \pi_y^*(w)\, d\mathbf{x} \tag{93}$$

$$\text{For any } (v,w) \in \bar{H}_{\mathcal{M}}(\Omega) \times \widehat{H}_{\mathcal{M}}(\Omega), \quad \int_\Omega \pi_y^*(v)\, w\, d\mathbf{x} = \int_\Omega v\, \pi_y(w)\, d\mathbf{x} \tag{94}$$

$$\text{For any } (v,w) \in \bar{H}_{\mathcal{M}}(\Omega) \times \widetilde{H}_{\mathcal{M}}(\Omega), \quad \int_\Omega \pi_x^*(v)\, w\, d\mathbf{x} = \int_\Omega v\, \pi_x(w)\, d\mathbf{x} \tag{95}$$

Proof. The proof is direct from the above definitions of spaces and operators. □

Proposition A.6. *Let assume that the mesh $\mathcal{M}$ is uniform in each direction in the sense of Definition 4.3. Then:*

$$\begin{aligned}
&\text{For any } v \in \widehat{H}_{\mathcal{M}}(\Omega), \quad \|\pi_x^*(v)\|_{L^4(\Omega)} \leqslant \|v\|_{L^4(\Omega)},\\
&\text{For any } v \in \widetilde{H}_{\mathcal{M}}(\Omega), \quad \|\pi_y^*(v)\|_{L^4(\Omega)} \leqslant \|v\|_{L^4(\Omega)},\\
&\text{For any } v \in \widetilde{H}_{\mathcal{M}}(\Omega), \quad \|\pi_x(v)\|_{L^4(\Omega)} \leqslant \|v\|_{L^4(\Omega)},\\
&\text{For any } v \in \widehat{H}_{\mathcal{M}}(\Omega), \quad \|\pi_y(v)\|_{L^4(\Omega)} \leqslant \|v\|_{L^4(\Omega)}.
\end{aligned}$$

Proof. The proof is direct from the above definitions of spaces, operators and the Young inequality. □

B Some technical lemmas

Prior to the statement of Lemma B.1 to which this section is devoted, one needs to define some quantities. For the convection term, we define $\mathbf{v}_h \cdot \nabla_h u_h \in H_{\mathcal{M}}(\Omega)$ by setting

$$\mathbf{v}_h \cdot \nabla_h u_h = \pi_x^*(v_{1,h}\, \delta_x u_h) + \pi_y^*(v_{2,h}\, \delta_y u_h). \tag{96}$$

A simple calculation gives

$$(\mathbf{v}_h \cdot \nabla_h u_h)_{|C_{i,j}} = \frac{1}{h_x h_y} \sum_{\sigma \in \mathcal{E}_{i,j}} v_{i,j,\sigma}\, u_{i,j,\sigma,c}, \tag{97}$$

with the centered choice

$$u_{i,j,\sigma,c} = \begin{cases} \dfrac{u_{i,j} + u_{i\pm1,j}}{2} & \text{if } \sigma = \sigma_{i\pm1/2,j},\\[2ex] \dfrac{u_{i,j} + u_{i,j\pm1}}{2} & \text{if } \sigma = \sigma_{i,j\pm1/2}. \end{cases}$$

Finally, for the upwind choice used in the scheme (41), we have:

$$\sum_{\sigma \in \mathcal{E}_{i,j}} v_{i,j,\sigma}\, u_{i,j,\sigma,+} = h_x h_y\, (\mathbf{v}_h \cdot \nabla_h u_h)_{|C_{i,j}} + \sum_{\sigma \in \mathcal{E}_{i,j}} v_{i,j,\sigma}(u_{i,j,\sigma,+} - u_{i,j,\sigma,c}). \tag{98}$$

We then deduce from (85) that

$$\int_\Omega (\mathbf{v}_h \cdot \nabla_h u_h)\ \Delta_h u_h \,\mathrm{d}\mathbf{x} = -\int_\Omega \nabla_h(\mathbf{v}_h \cdot \nabla_h u_h) \cdot \nabla_h u_h \,\mathrm{d}\mathbf{x}. \tag{99}$$

We are now in position to state the following technical lemma to be used for the numerical analysis of the scheme (41).

Lemma B.1. *For any $u_h \in H_{\mathcal{M}}(\Omega)$ and $\mathbf{v}_h = (v_{1,h}, v_{2,h}) \in V_{\mathcal{E},0}(\Omega)$ we have:*

$$\int_\Omega \nabla_h(\mathbf{v}_h \cdot \nabla_h u_h) \cdot \nabla_h u_h \ d\mathbf{x} = T_1^{(1,2)} + T_1^{(2,2)} + T_2^{(1,2)} + T_2^{(2,2)}, \tag{100}$$

with:

$$T_1^{(1,2)} = \int_\Omega \pi_x(\delta_x^* v_{1,h} \cdot \pi_x^*(\delta_x u_h)) \cdot \delta_x u_h \ d\mathbf{x},$$

$$T_1^{(2,2)} = \int_\Omega \pi_y^*(\delta_x v_{2,h} \cdot \pi_x(\delta_y u_h)) \cdot \delta_x u_h \ d\mathbf{x},$$

$$T_2^{(1,2)} = \int_\Omega \pi_x^*(\delta_y v_{1,h} \cdot \pi_y(\delta_x u_h)) \cdot \delta_y u_h \ d\mathbf{x},$$

and

$$T_2^{(2,2)} = \int_\Omega \pi_y(\delta_y^* v_{2,h} \cdot \pi_y^*(\delta_y u_h)) \cdot \delta_y u_h \ d\mathbf{x}.$$

Proof. From the definition of the discrete gradient ∇_h, we have

$$\int_\Omega \nabla_h(\mathbf{v}_h \cdot \nabla_h u_h) \cdot \nabla_h u_h \,\mathrm{d}\mathbf{x} = T_1 + T_2$$

with

$$T_1 = \int_\Omega \delta_x(\mathbf{v}_h \cdot \nabla_h u_h) \cdot \delta_x u_h \,\mathrm{d}\mathbf{x} \quad \text{and} \quad T_2 = \int_\Omega \delta_y(\mathbf{v}_h \cdot \nabla_h u_h) \cdot \delta_y u_h \,\mathrm{d}\mathbf{x}$$

Moreover from (96), (87) and (88), the term $T_1 = T_1^{(1)} + T_1^{(2)}$ splits into

$$T_1^{(1)} = \int_\Omega \pi_x(\delta_x^*(v_{1,h}\delta_x u_h)) \cdot \delta_x u_h \mathrm{d}\mathbf{x} \quad \text{and} \quad T_1^{(2)} = \int_\Omega \pi_y^*(\delta_x(v_{2,h}\delta_y u_h)) \cdot \delta_x u_h \mathrm{d}\mathbf{x}.$$

Now from (90) and (91), these expressions further decompose as

$$T_1^{(1)} = T_1^{(1,1)} + T_1^{(1,2)} \quad \text{and} \quad T_1^{(2)} = T_1^{(2,1)} + T_1^{(2,2)}$$

with $T_1^{(1,2)}$ and $T_1^{(2,2)}$ as in the statement of Lemma B.1 above, and

$$\begin{aligned} T_1^{(1,1)} &= \int_\Omega \pi_x(\pi_x^*(v_{1,h})\, \delta_{xx} u_h) \cdot \delta_x u_h \,\mathrm{d}\mathbf{x}, \\ T_1^{(2,1)} &= \int_\Omega \pi_y^*(\pi_x(v_{2,h})\, (\delta_x \circ \delta_y) u_h) \cdot \delta_x u_h \,\mathrm{d}\mathbf{x}. \end{aligned}$$

Since $(\delta_x u_h)_{\frac{1}{2},j} = (\delta_x u_h)_{N+\frac{1}{2},j} = 0$, we can write:

$$\begin{aligned} T_1^{(1,1)} &= \frac{h_x h_y}{2} \sum_{i=2}^{N} \sum_{j=1}^{M} \left[(\pi_x^*(v_{1,h}) \delta_{xx} u_h)_{i,j} + (\pi_x^*(v_{1,h}) \delta_{xx} u_h)_{i-1,j} \right] (\delta_x u_h)_{i-\frac{1}{2},j} \\ &= \frac{h_y}{2} \sum_{i=2}^{N} \sum_{j=1}^{M} (\pi_x^*(v_{1,h}))_{i,j} ((\delta_x u_h)_{i+\frac{1}{2},j} - (\delta_x u_h)_{i-\frac{1}{2},j}) (\delta_x u_h)_{i-\frac{1}{2},j} \\ &\quad + \frac{h_y}{2} \sum_{i=2}^{N} \sum_{j=1}^{M} (\pi_x^*(v_{1,h}))_{i-1,j} ((\delta_x u_h)_{i-\frac{1}{2},j} - (\delta_x u_h)_{i-\frac{3}{2},j}) (\delta_x u_h)_{i-\frac{1}{2},j} \\ &= -\frac{1}{2} \int_\Omega \delta_x(\pi_x^* v_{1,h}) (\delta_x u_h)^2 \, \mathrm{d}\mathbf{x}. \end{aligned} \tag{101}$$

As highlighted in (5), δ_x and δ_y commute, so that

$$T_1^{(2,1)} = \int_\Omega \pi_y^*(\pi_x(v_{2,h}) \, (\delta_y \circ \delta_x) u_h) \cdot \delta_x u_h \, \mathrm{d}\mathbf{x}.$$

Since $(\delta_y u_h)_{i,\frac{1}{2}} = (\delta_y u_h)_{i,M+\frac{1}{2}} = 0$, one can proceed as for $T_1^{(1,1)}$ to get that

$$T_1^{(2,1)} = -\frac{1}{2} \int_\Omega \delta_y^*(\pi_x v_{2,h}) (\delta_x u_h)^2 \, \mathrm{d}\mathbf{x}. \tag{102}$$

Finally, summing up (101)–(102) and using (87) and (89) yields

$$\begin{aligned} T_1^{(1,1)} + T_1^{(2,1)} &= -\frac{1}{2} \int_\Omega [\delta_x(\pi_x^* v_{1,h}) + \delta_y^*(\pi_x v_{2,h})] (\delta_x u_h)^2 \, \mathrm{d}\mathbf{x} \\ &= -\frac{1}{2} \int_\Omega [\pi_x(\delta_x^* v_{1,h} + \delta_y^* v_{2,h})] (\delta_x u_h)^2 \, \mathrm{d}\mathbf{x} = 0 \end{aligned}$$

since $\mathbf{v}_h \in V_{\mathcal{E},0}(\Omega)$. Therefore $T_1 = T_1^{(1,2)} + T_1^{(2,2)}$, as well as $T_2 = T_2^{(1,2)} + T_2^{(2,2)}$ thanks to similar calculations. Consequently (100) holds. □

Lemma B.2. *Let $\mathbf{v} \in V_0(\Omega) \cap H^2(\Omega)^2$, and let $\mathbf{v}_h \in V_{\mathcal{E},0}(\Omega)$ be defined by* (38)*, then*

$$\|\mathbf{v}_h\|_{L^2(\Omega)^2} \leqslant \|\mathbf{v}\|_{L^2(\Omega)^2} + \frac{h^2}{2} \left(\|\partial_{xx} v_1\|_{L^2(\Omega)^2}^2 + \|\partial_{yy} v_2\|_{L^2(\Omega)^2}^2 \right)^{1/2} \tag{103}$$

Moreover, we have the following estimate on $\nabla_h \mathbf{v}_h$ defined by (4)*:*

$$\|\nabla_h \mathbf{v}_h\|_{L^2(\Omega)^{2\times 2}} \leqslant \|\nabla \mathbf{v}\|_{L^2(\Omega)^{2\times 2}} + h \|\partial_x \partial_y \mathbf{v}\|_{L^2(\Omega)^2}. \tag{104}$$

Proof. Let us start by establishing (103). Denote by

$$\overline{v}_{i+1/2,j} = \frac{1}{h_x h_y} \iint_{C_{i+1/2,j}} v_1 \mathrm{d}\mathbf{x}, \qquad \overline{v}_{i,j+1/2} = \frac{1}{h_x h_y} \iint_{C_{i,j+1/2}} v_2 \mathrm{d}\mathbf{x},$$

and by $\overline{v}_1, \overline{v}_2$ and $\overline{\mathbf{v}}_h$ the corresponding elements in $\widehat{H}_{\mathcal{M}}(\Omega)$, $\widetilde{H}_{\mathcal{M}}(\Omega)$ and $\mathbb{H}_{\mathcal{M}}(\Omega)$. Then Jensen's inequality gives that

$$\|\overline{\mathbf{v}}_h\|_{L^2(\Omega)^2} \leqslant \|\mathbf{v}\|_{L^2(\Omega)^2}.$$

So (103) holds true provided that

$$\|\mathbf{v}_h - \overline{\mathbf{v}}_h\|_{L^2(\Omega)^2} \leqslant \frac{h^2}{2}\left(\|\partial_{xx} v_1\|^2_{L^2(\Omega)^2} + \|\partial_{yy} v_2\|^2_{L^2(\Omega)^2}\right)^{1/2}. \tag{105}$$

To establish (105), let us remark that

$$\begin{aligned}
\overline{v}_{i+1/2,j} - v_{i+1/2,j} &= \frac{1}{h_x h_y} \iint\limits_{C_{i+1/2,j}} (v_1(x,y) - v_1(x_{i+1/2}, y)) \mathrm{d}x\mathrm{d}y \\
&= \frac{1}{h_x h_y} \iint\limits_{C_{i+1/2,j}} \int_0^{h_x/2} \int_{-\xi}^{\xi} \partial_{xx} v_1(x_{i+1/2} + s, y) \mathrm{d}s\mathrm{d}\xi\mathrm{d}x\mathrm{d}y.
\end{aligned}$$

As a consequence

$$|\overline{v}_{i+1/2,j} - v_{i+1/2,j}| \leqslant \frac{h_x}{2h_y} \iint\limits_{C_{i+1/2,j}} |\partial_{xx} v_1(\mathbf{x})| \, \mathrm{d}\mathbf{x},$$

hence the Cauchy-Schwarz inequality provides

$$|\overline{v}_{i+1/2,j} - v_{i+1/2,j}|^2 \leqslant \frac{h_x^3}{4h_y} \iint\limits_{C_{i+1/2,j}} |\partial_{xx} v_1(\mathbf{x})|^2 \, \mathrm{d}\mathbf{x}.$$

Finally, summing up over $C_{i+1/2,j} \in \widehat{\mathcal{M}}$ leads to

$$\|\overline{v}_{1,h} - v_{1,h}\|^2_{L^2(\Omega)} \leqslant \frac{h_x^4}{4} \|\partial_{xx} v_1\|^2_{L^2(\Omega)},$$

whereas similar computations yield

$$\|\overline{v}_{2,h} - v_{2,h}\|^2_{L^2(\Omega)} \leqslant \frac{h_y^4}{4} \|\partial_{yy} v_2\|^2_{L^2(\Omega)}.$$

Therefore, (105) holds true, and thus (103) too.

We now focus on inequality (104). Let us first show some control on the first diagonal term of $\nabla_h \mathbf{v}_h$, the second being similar. Let $C_{i,j} \in \mathcal{M}$, then

$$\delta_x^* v_{1,h|_{C_{i,j}}} = \frac{1}{h_x h_y} \int_{y_{j-1/2}}^{y_{j+1/2}} \big(v_1(x_{i+1/2}, y) - v_1(x_{i-1/2}, y)\big) \, \mathrm{d}y = \frac{1}{h_x h_y} \iint\limits_{C_{i,j}} \partial_x v_1(x,y) \mathrm{d}\mathbf{x}.$$

Then we deduce from Jensen's inequality that

$$\left|\delta_x^* v_{1,h|_{C_{i,j}}}\right|^2 \leqslant \frac{1}{h_x h_y} \iint\limits_{C_{i,j}} |\partial_x v_1(x,y)|^2 \, \mathrm{d}\mathbf{x},$$

and thus, after summing up over $C_{i,j} \in \mathcal{M}$, that

$$\|\delta_x^* v_{1,h}\|_{L^2(\Omega)} \leqslant \|\partial_x v_1\|_{L^2(\Omega)}, \qquad \|\delta_y^* v_{2,h}\|_{L^2(\Omega)} \leqslant \|\partial_y v_2\|_{L^2(\Omega)}. \tag{106}$$

Let us focus now on the extra-diagonal terms, and particularly on $\delta_y v_{1,h}$, the case of $\delta_x v_{2,h}$ being similar. We denote by $\overline{\delta_y v_{1,h}}$ the element of $\overline{H}_{\mathcal{M}}(\Omega)$ defined by

$$\overline{\delta_y v_{1,h}}_{|C_{i+1/2,j+1/2}} = \frac{1}{h_x h_y} \iint_{C_{i+1/2,j+1/2}} \partial_y v_1(\mathbf{x}) \mathrm{d}\mathbf{x}, \qquad \forall\, C_{i+1/2,j+1/2} \in \overline{\mathcal{M}},$$

then owing to Jensen's inequality, we get that

$$\|\overline{\delta_y v_{1,h}}\|_{L^2(\Omega)} \leqslant \|\partial_y v_1\|_{L^2(\Omega)}. \tag{107}$$

On the other hand,

$$\begin{aligned}
&\big(\overline{\delta_y v_{1,h}} - \delta_y v_{1,h}\big)_{|C_{i+1/2,j+1/2}} \\
&\quad = \frac{1}{h_x h_y^2} \int_{-h_y/2}^{h_y/2} \int_{y_j}^{y_{j+1}} \int_{x_i}^{x_{i+1}} \partial_y (v_1(x_{i+1/2}, y+s) - v_1(x, y+s)) \mathrm{d}x\mathrm{d}y\mathrm{d}s \\
&\quad = \frac{1}{h_x h_y^2} \int_{-h_y/2}^{h_y/2} \int_{y_j}^{y_{j+1}} \int_{x_i}^{x_{i+1}} \int_{x_{i+1/2}}^{x} \partial_x \partial_y v_1(\xi, y+s) \mathrm{d}\xi\mathrm{d}x\mathrm{d}y\mathrm{d}s,
\end{aligned}$$

whence

$$\big|\overline{\delta_y v_{1,h}} - \delta_y v_{1,h}\big|_{|C_{i+1/2,j+1/2}} \leqslant \frac{1}{h_y} \iint_{C_{i+1/2,j+1/2}} |\partial_x \partial_y v_1(\mathbf{x})| \, \mathrm{d}\mathbf{x}.$$

Applying Jensen's inequality, we obtain that

$$\big|\overline{\delta_y v_{1,h}} - \delta_y v_{1,h}\big|^2_{|C_{i+1/2,j+1/2}} \leqslant \frac{h_x}{h_y} \iint_{C_{i+1/2,j+1/2}} |\partial_x \partial_y v_1(\mathbf{x})|^2 \, \mathrm{d}\mathbf{x},$$

leading to

$$\|\overline{\delta_y v_{1,h}} - \delta_y v_{1,h}\|_{L^2(\Omega)} \leqslant h_x \|\partial_x \partial_y v_1\|_{L^2(\Omega)} \tag{108}$$

after summation over $i \in [\![1, N-1]\!]$ and $j \in [\![1, M-1]\!]$. Similarly, we get that

$$\|\overline{\delta_x v_{2,h}}\|_{L^2(\Omega)} \leqslant \|\partial_x v_2\|_{L^2(\Omega)} \quad \text{and} \quad \|\overline{\delta_x v_{2,h}} - \delta_x v_{2,h}\|_{L^2(\Omega)} \leqslant h_x \|\partial_x \partial_y v_2\|_{L^2(\Omega)}. \tag{109}$$

The combination of (107), (108) and (109) gives (104). □

References

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

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

[3] F. Bouchut, R. Eymard, and A. Prignet. Finite volume schemes for the approximation via characteristics of linear convection equations with irregular data. *J. Evol. Equ.*, 11(3):687–724, 2011.

[4] K. Brenner, C. Cancès, and D. Hilhorst. Finite volume approximation for an immiscible two-phase flow in porous media with discontinuous capillary pressure. *Comput. Geosci.*, 17(3):573–597, 2013.

[5] D. Bresch, E. H. Essoufi, and M. Sy. Effect of density dependent viscosities on multiphasic incompressible fluid models. *J. Math. Fluid Mech.*, 9(3):377–397, 2007.

[6] D. Bresch, V. Giovangigli, and E. Zatorska. Two-velocity hydrodynamics in fluid mechanics: Part I. Well posedness for zero Mach number systems. *J. Math. Pures Appl. (9)*, 104(4):762–800, 2015.

[7] C. Calgaro, C. Colin, and E. Creusé. A combined finite volume-finite element scheme for a low-Mach system involving a Joule term. *AIMS Math.*, 5(1):311–331, 2020.

[8] C. Calgaro, C. Colin, E. Creusé, and E. Zahrouni. Approximation by an iterative method of a low-Mach model with temperature dependent viscosity. *Math. Methods Appl. Sci.*, 42(1):250–271, 2019.

[9] C. Cancès, T. O. Gallouët, and Gabriele Todeschi. A variational finite volume scheme for Wasserstein gradient flows. *Numer. Math.*, 146(3):437–480, 2020.

[10] K. Deimling. *Nonlinear functional analysis*. Springer-Verlag, Berlin, 1985.

[11] R. Eymard, T. Gallouët, M. Ghilani, and R. Herbin. Error estimates for the approximate solutions of a nonlinear hyperbolic equation given by finite volume schemes. *IMA J. Numer. Anal.*, 18(4):563–594, 1998.

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

[13] R. Eymard, T. Gallouët, and R. Herbin. Discretization of heterogeneous and anisotropic diffusion problems on general nonconforming meshes sushi: a scheme using stabilization and hybrid interfaces. *IMA J. Numer. Anal.*, 30(4):1009–1043, 2010.

[14] A. Fiorenza, M. R. Formica, T. G. Roskovec, and F. Soudský. Detailed proof of classical Gagliardo-Nirenberg interpolation inequality with historical remarks. *Z. Anal. Anwend.*, 40(2):217–236, 2021.

[15] E. Gagliardo. Ulteriori proprietà di alcune classi di funzioni in più variabili. *Ricerche Mat.*, 8:24–51, 1959.

[16] T. Gallouët. Discrete functional analysis tools for some evolution equations. *Comput. Methods Appl. Math.*, 18(3):477–493, 2018.

[17] T. Gallouët and J.-C. Latché. Compactness of discrete approximate solutions to parabolic PDEs—application to a turbulence model. *Commun. Pure Appl. Anal.*, 11(6):2371–2391, 2012.

[18] W. Hackbusch. *Elliptic Differential Equations: Theory and Numerical Treatment*. Springer Series in Computational Mathematics 18. Springer-Verlag Berlin Heidelberg, 2nd edition, 2017.

[19] F. Huang and W. Tan. On the strong solution of the ghost effect system. *SIAM J. Math. Anal.*, 49(5):3496–3526, 2017.

[20] A. Jüngel. *Quasi-hydrodynamic semiconductor equations*. Number 41 in Progress in Nonlinear Differential Equations and their applications. Birkhäuser, Basel, 2001.

[21] J. Leray and J. Schauder. Topologie et équations fonctionnelles. *Ann. Sci. École Norm. Sup.*, 51((3)):45–78, 1934.

[22] C.D. Levermore, W. Sun, and K. Trivisa. Local well-posedness of a ghost system effect. *Indiana Univ. Math. J.*, 60:517–576, 2011.

[23] A. Majda and J. Sethian. The derivation and numerical solution of the equations for zero Mach number combustion. *Combustion Science and Technology*, 42:185–205, 1985.

[24] L. Nirenberg. On elliptic partial differential equations. *Ann. Scuola Norm. Sup. Pisa Cl. Sci. (3)*, 13:115–162, 1959.