



HAL
open science

Discretization of the coupled heat and electrical diffusion problems by the finite element and the finite volume methods

Abdallah Bradji, Raphaelae Herbin

► **To cite this version:**

Abdallah Bradji, Raphaelae Herbin. Discretization of the coupled heat and electrical diffusion problems by the finite element and the finite volume methods. IMA Journal of Numerical Analysis, 2008, 28 (3), pp.469-495. <10.1093/imanum/drm030>. <hal-00123247>

HAL Id: hal-00123247

<https://hal.science/hal-00123247v1>

Submitted on 8 Jan 2007

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.



HAL Authorization

DISCRETIZATION OF THE COUPLED HEAT AND ELECTRICAL DIFFUSION PROBLEMS BY THE FINITE ELEMENT AND THE FINITE VOLUME METHODS

ABDALLAH BRADJI† AND RAPHAÈLE HERBIN*

ABSTRACT. The modelling of the heat diffusion coupled with electrical diffusion yields a nonlinear system of elliptic equations. The ohmic losses which appear as a source term in the heat diffusion equation is a nonlinear term which lies in L^1 . A finite element scheme and a finite volume scheme are considered for the discretization of the system; in both cases, we show that the approximate solution obtained with the scheme converges, up to a subsequence, to a solution of the coupled elliptic system.

1. INTRODUCTION

Let Ω be a bounded domain of \mathbb{R}^d , $d = 2$ or 3 , made up of a thermally and electrically conducting material. It is well known that the diffusion of electricity in a resistive medium induces some heating, known as ohmic losses. Such a situation arises for instance in the modelling of fuel cells, see e.g. [22, 23] and references therein. Let ϕ denote the electrical potential, and κ the electrical conductivity; then the ohmic losses may be written as $\kappa \nabla \phi \cdot \nabla \phi$. Since ϕ is the solution of a diffusion equation, it is reasonable to seek ϕ in the space $H^1(\Omega)$, so that $\nabla \phi \cdot \nabla \phi \in L^1$. Hence the heat diffusion equation has a right hand-side in L^1 , and its analysis falls out of the usual variational framework. Our aim in this paper is to study the convergence of approximate solutions to the resulting coupled problem obtained with both a linear finite element method and a cell centred finite volume scheme.

The theory of elliptic and parabolic equations with irregular right-hand-side goes back to the pioneering work of G. Stampacchia [31], where solutions to the linear problem are defined by duality. Later on, L. Boccardo, T. Gallouët and co-authors [4, 5, 6, 7] introduced the tools and setting in which one may define solutions to such problems: these so-called entropy solutions [2] were found to be equivalent to the so-called renormalized solutions of P.-L. Lions and F. Murat [29], as well as to the solutions obtained by approximation, as defined by [14]. In the linear case, all these solutions are also equivalent to those of Stampacchia.

Other solutions obtained by approximation were defined thanks to numerical schemes. They also lead to the existence of a solution, but more importantly, they yield a constructive way to compute approximate solutions of the problem. Convergence of the finite volume scheme was proven in [26] for the Laplace equation with right-hand-side measure; the proof was generalized in [15] to noncoercive convection diffusion problems. Convergence of the finite element scheme, with irregular data, on bi-dimensional polygonal domains was proven for Delaunay triangular meshes in [24] and in [9] for three-dimensional tetrahedral meshes under geometrical conditions. Error estimates may be obtained using some “suitable” negative Sobolev spaces as in [30] or interpolation error, under regularity assumptions on the solution, as in [12, 9].

Key words and phrases. Nonlinear elliptic system, Diffusion equation, Finite element scheme, Finite volume scheme, L^1 -data, Ohmic losses.

In the present work, we use some of the techniques introduced in the above references to prove the convergence of both the finite volume and finite element methods for the approximation of the above mentioned heat and electricity diffusion problem. The considered system of semilinear elliptic partial differential equations is such that the right-hand-side of the second equation depends on the solution of the first one and is in L^1 . The paper is organized as follows: in section 2, we present the continuous problem, its weak form and the known result about existence [25]. In section 3, we describe the finite volume and finite element methods for the approximation of the system and prove the existence of a solution to the resulting discrete system for both cases. The convergence of the finite element scheme is proven in section 4, and the convergence of the finite volume scheme in section 5. In both cases, the proof of convergence is based on a priori estimates, compactness result and a passage to the limit in the scheme. Some conclusions and perspectives are drawn in the last section.

2. THE CONTINUOUS PROBLEM

We wish to find some numerical approximation of solutions to the following nonlinear coupled elliptic system, which models the thermal and electrical diffusion in a material subject to ohmic losses:

$$-\nabla \cdot (\kappa(x, u(x)) \nabla \phi(x)) = f(x, u(x)), \quad x \in \Omega, \quad (1)$$

$$\phi(x) = 0, \quad x \in \partial\Omega, \quad (2)$$

$$-\nabla \cdot (\lambda(x, u(x)) \nabla u(x)) = \kappa(x, u(x)) |\nabla \phi|^2(x), \quad x \in \Omega, \quad (3)$$

$$u(x) = 0, \quad x \in \partial\Omega, \quad (4)$$

where Ω is a convex polygonal open subset of \mathbb{R}^d , $d = 2$ or 3 , with boundary $\partial\Omega$, ϕ denotes the electrical potential and u the temperature; the electrical conductivity κ , the thermal conductivity λ and the source term f are functions from $\Omega \times \mathbb{R}$ to \mathbb{R} satisfying the following Assumptions:

Assumption 1. *The functions κ, λ and f , defined from $\Omega \times \mathbb{R}$ to \mathbb{R} , are bounded and continuous with respect to $y \in \mathbb{R}$ for a.e. $x \in \Omega$, and measurable with respect to $x \in \Omega$ for any $y \in \mathbb{R}$, and such that:*

$$\exists \alpha > 0; \quad \alpha \leq \kappa(x, y) \text{ and } \alpha \leq \lambda(x, y), \quad \forall y \in \mathbb{R}, \text{ for a.e. } x \in \Omega. \quad (5)$$

The following existence result was proven in [25]:

Theorem 2.1. *Under Assumption 1, there exists a solution to the following weak form of Problem (1)–(4):*

$$\left\{ \begin{array}{l} (\phi, u) \in H_0^1(\Omega) \times \cap_{p < \frac{d}{d-1}} W_0^{1,p}(\Omega), \\ \int_{\Omega} \kappa(\cdot, u) \nabla \phi \cdot \nabla \psi \, dx = \int_{\Omega} f(\cdot, u) \psi \, dx, \quad \forall \psi \in H_0^1(\Omega) \\ \int_{\Omega} \lambda(\cdot, u) \nabla u \cdot \nabla v \, dx = \int_{\Omega} \kappa(\cdot, u) |\nabla \phi|^2 v \, dx, \quad \forall v \in \cup_{r > d} W_0^{1,r}(\Omega). \end{array} \right. \quad (6)$$

Note that the exponents $\frac{d}{d-1}$ and d are conjugate, and that, for $r > d$ the space $W_0^{1,r}(\Omega)$ is continuously imbedded in the space $\mathcal{C}(\overline{\Omega}, \mathbb{R})$; therefore all terms in (6) make sense. In the case $d = 2$, we have $u \in W_0^{1,p}(\Omega)$ for all $p < 2$, but in general, $u \notin H_0^1(\Omega)$. Similarly, if $d = 3$, $u \in W_0^{1,p}(\Omega)$ for all $p < \frac{3}{2}$.

The proof of this theorem relies mainly on the analysis tools which were developed for the analysis of elliptic equations with irregular right-hand-side, see for instance [4] or [6]. We shall not need to assume this existence result for our present analysis. Indeed, the existence of a solution to (1)–(4) is obtained as a by-product of the convergence of the scheme. Nevertheless, a large part

of the convergence analysis of the schemes is inspired from the ideas developed in [25] for the existence result, and we shall again use the ideas of [4] and [7] in our proofs.

3. THE DISCRETIZATION SCHEMES

In [23], the numerical simulation of solid oxide fuel cells led to a mathematical model involving a set of semilinear partial differential equations, the unknowns of which were the temperature, the electrical potential and the concentrations of various chemical species in the porous media of the cell. System (1)–(4) is a sub–problem of this latter model, obtained by leaving out the chemical species diffusion equations. In [23], three different discretization schemes were implemented and compared, namely the linear finite element method, the mixed finite element method, and the cell centred finite volume method. Because of interface conditions involving the electrical current [28], a precise approximation of the electrical flux is needed at the interfaces, the linear finite element method was found to be less adapted than the two latter methods, so that finally the mixed finite element method and the cell centred finite volume method were numerically compared. The cell centred finite volume method was found to be easier to implement and comparable to the mixed finite element method as to the ratio precision *vs.* computing time, so that it was finally chosen for the simulations of different geometries of fuel cells [22]. Here we shall give a theoretical justification of the convergence of both the linear finite element method and the cell centred finite volume method for the discretization of system (1)–(4). Let us first start by introducing the finite element scheme.

3.1. The finite element scheme. Let \mathcal{M} denote a finite element mesh of Ω , consisting of simplices and satisfying the usual conditions, see e.g. [10, p. 61], that is:

Definition 1 (Finite element mesh). *Let \mathcal{M} be a set of open triangular (in two space dimensions) or tetrahedral (in three space dimensions) subsets of Ω such that:*

- $\bar{\Omega} = \bigcup_{T \in \mathcal{M}} \bar{T}$.
- For any $(T, T') \in \mathcal{M}^2$, $T \neq T' \implies T \cap T' = \emptyset$.
- For any $(T, T') \in \mathcal{M}^2$, $\bar{T} \cap \bar{T}' = \emptyset$ or $\bar{T} \cap \bar{T}'$ is an edge (or a face in three space dimensions) of T and T' .

We define the mesh size of \mathcal{M} by $h_{\mathcal{M}} = \sup\{\text{diam}(T), T \in \mathcal{M}\}$, where $\text{diam}(T)$ denotes the diameter of T .

The set of vertices x_i of the finite element mesh is indexed by $\mathcal{V} = \mathcal{I} \cup \mathcal{B}$, where \mathcal{I} (resp. \mathcal{B}) refers to the interior (resp. boundary) vertices, namely the vertices laying in Ω (resp. on $\partial\Omega$). For any $i \in \mathcal{V}$ let ξ_i be the basis function associated with the vertex x_i , defined by:

$$\begin{cases} \xi_i \in \mathcal{C}(\bar{\Omega}), \xi_i|_T \in \mathbb{P}_1 \text{ for all } T \in \mathcal{M}, \\ \xi_i(x_i) = 1, \xi_i(x_j) = 0, \forall j \in \mathcal{V} \text{ such that } j \neq i, \end{cases}$$

where \mathbb{P}_1 is the set of affine functions. Let us then consider the linear finite element space spanned by the basis functions $(\xi_i)_{i \in \mathcal{I}}$:

$$V_0^{\mathcal{M}} = \{u \in \mathcal{C}(\bar{\Omega}); u|_T \in \mathbb{P}_1 \text{ for all } T \in \mathcal{M} \text{ and } u = 0 \text{ on } \partial\Omega\}, \quad (7)$$

A finite element approximation of (1)-(4) may then be given by:

$$\left\{ \begin{array}{l} \text{Find } (u_{\mathcal{M}}, \phi_{\mathcal{M}}) \in V_0^{\mathcal{M}} \times V_0^{\mathcal{M}} \text{ such that :} \\ \int_{\Omega} \kappa(\cdot, u_{\mathcal{M}}) \nabla \phi_{\mathcal{M}} \cdot \nabla \psi \, dx = \int_{\Omega} f(\cdot, u_{\mathcal{M}}) \psi \, dx, \forall \psi \in V_0^{\mathcal{M}}, \\ \int_{\Omega} \lambda(\cdot, u_{\mathcal{M}}) \nabla u_{\mathcal{M}} \cdot \nabla v \, dx = \int_{\Omega} \kappa(\cdot, u_{\mathcal{M}}) |\nabla \phi_{\mathcal{M}}|^2 v \, dx, \forall v \in V_0^{\mathcal{M}}. \end{array} \right. \quad (8)$$

To prove the convergence of the finite element scheme (8), we need the following assumption on the mesh \mathcal{M} , which are also required for the discrete maximum principle to hold, see e.g. [10, p. 148].

Assumption 2. *Let \mathcal{M} be a simplicial finite element mesh in the sense of Definition 1. We assume that for all $u_{\mathcal{M}} \in V_0^{\mathcal{M}}$,*

$$\theta_{i,j}^{\lambda}(u_{\mathcal{M}}) = - \int_{\Omega} \lambda(\cdot, u_{\mathcal{M}}) \nabla \xi_i \cdot \nabla \xi_j \, dx \geq 0, \quad \forall (i, j) \in \mathcal{I} \times \mathcal{V} \text{ such that } i \neq j, \quad (9)$$

In the case of the Laplace operator (i.e. $\lambda(\cdot, u_{\mathcal{M}}) \equiv 1$), it is well known (see e.g. [9]) that in two space dimensions, Assumption 2 is equivalent to the fact that \mathcal{M} is Delaunay, i.e. for every edge $[x_i x_j]$ of the triangulation such that $[x_i x_j] \not\subset \partial\Omega$, the sum of the two opposite angles facing $[x_i x_j]$ is less or equal π . In three space dimensions, this condition holds if every inner dihedral angle of every tetrahedron is acute; however, there is to our knowledge no constructive way yet known to build such meshes [3, 16].

Remark 3.1. *Note that the condition (9) of Assumption 2 maybe replaced by condition (1.14) of [9], which we recall (abbreviating $\theta_{i,j}^{\lambda}(u_{\mathcal{M}})$ to $\theta_{i,j}$ for notational convenience):*

$$\forall i \in \mathcal{I}, \quad \theta_{i,i} + \sum_{\substack{j \in \mathcal{I} \\ j \neq i}} |\theta_{i,j}| \leq 0. \quad (10)$$

It is easily seen that (9) implies (10). Indeed, let $i \in \mathcal{I}$, thanks to the fact that $\sum_{j \in \mathcal{V}} \xi_j = 1$, we get that $\theta_{i,i} = - \sum_{j \in \mathcal{V}, j \neq i} \theta_{i,j}$; using assumption (9), we then get that $\theta_{i,i} \leq - \sum_{j \in \mathcal{V}, j \neq i} |\theta_{i,j}|$ which implies (10).

Now let \mathcal{I}_B denote the set of interior nodes which are neighbours to the boundary nodes; let us show that if condition (10) is satisfied, then condition (9) holds for any $(i, j) \in \mathcal{V} \times \mathcal{V} \setminus \mathcal{I}_B \times \mathcal{I}_B$. Indeed, if i (resp. j) $\in \mathcal{V} \setminus \mathcal{I}_B$, then $\theta_{i,k}$ (resp. $\theta_{j,k}$) = 0, and therefore:

$$\sum_{\substack{j \in \mathcal{I} \\ j \neq i}} |\theta_{i,j}| = \sum_{\substack{j \in \mathcal{V} \\ j \neq i}} |\theta_{i,j}| \geq \sum_{\substack{j \in \mathcal{V} \\ j \neq i}} \theta_{i,j} = -\theta_{i,i}.$$

Hence we get that

$$\theta_{i,i} + \sum_{\substack{j \in \mathcal{I} \\ j \neq i}} |\theta_{i,j}| \geq 0,$$

with equality if and only if $\theta_{i,j} \geq 0$ for all $j \in \mathcal{I}$, which shows that this latter condition must hold in order for (10) to hold.

Hence condition (9) (which is the usual condition for the so called discrete maximum principle, see e.g. [11]) is slightly stronger than (10). Nevertheless, we prefer to use (9) for which some constructive characterizations are known.

3.2. A cell centred finite volume scheme. To define a finite volume approximation, we introduce an admissible mesh \mathcal{T} in the sense of [17, Definition 9.1 page 762], which we recall here for the sake of completeness:

Definition 2 (Admissible meshes). *Let Ω be an open bounded polygonal subset of \mathbb{R}^d , $d = 2$ or 3 . An admissible finite volume mesh of Ω , denoted by \mathcal{T} , is given by a family of “control volumes”, which are open polygonal convex subsets of Ω , a family of subsets of $\overline{\Omega}$ contained in hyperplanes of \mathbb{R}^d , denoted by \mathcal{E} (these are the edges in two space dimensions, or faces in three space dimensions, of the control volumes), with strictly positive $(d - 1)$ -dimensional measure, and a family $(x_K)_{K \in \mathcal{T}}$ of points of Ω satisfying the following properties:*

- (i) *The closure of the union of all the control volumes is $\overline{\Omega}$.*
- (ii) *For any $K \in \mathcal{T}$, there exists a subset \mathcal{E}_K of \mathcal{E} such that $\partial K = \overline{K} \setminus K = \cup_{\sigma \in \mathcal{E}_K} \overline{\sigma}$. Furthermore, $\mathcal{E} = \cup_{K \in \mathcal{T}} \mathcal{E}_K$.*
- (iii) *For any $(K, L) \in \mathcal{T}^2$ with $K \neq L$, either the $(d - 1)$ -dimensional Lebesgue measure of $\overline{K} \cap \overline{L}$ is 0 or $\overline{K} \cap \overline{L} = \overline{\sigma}$ for some $\sigma \in \mathcal{E}$, which will then be denoted by $K|L$.*
- (iv) *The family of points $(x_K)_{K \in \mathcal{T}}$ is such that $x_K \in \overline{K}$ (for all $K \in \mathcal{T}$) and, if $\sigma = K|L$, it is assumed that $x_K \neq x_L$, and that the straight line going through x_K and x_L is orthogonal to $K|L$.*

An example of two cells of such a mesh is given in Figure 1, along with some notations.

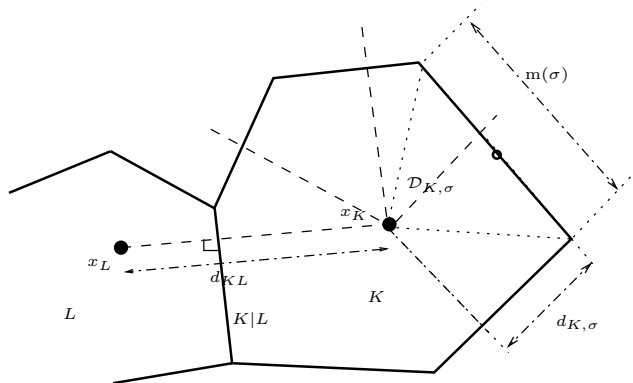


FIGURE 1. Notations for a control volume K in the case $d = 2$

Item (iv) of the above Definition will be referred to in the sequel as the “orthogonality property”.

We refer to [17] for a description of such admissible meshes, which include triangular meshes, rectangular meshes, or Voronoï meshes. Here, for the sake of simplicity, we assume that the points $x_K \in K$.

The finite volume approximations $\phi_{\mathcal{T}}$ and $u_{\mathcal{T}}$ of ϕ and u solution to (6) are sought in the space $X(\mathcal{T})$ of functions from Ω to \mathbb{R} which are constant over each control volume of the mesh, that is:

$$X(\mathcal{T}) = \{u \in L^2(\Omega); u|_K \in \mathbb{P}_0 \text{ for all } K \in \mathcal{T}\}, \quad (11)$$

where \mathbb{P}_0 denotes the set of constant functions.

Remark 3.2. Any element $u_T \in X(\mathcal{T})$ can be written as: $u_T = \sum_{K \in \mathcal{T}} u_K 1_K$, where $1_K(x) = 1$ if $x \in K$ and $1_K(x) = 0$ otherwise, and u_K denotes the value taken by u_T on the control volume K . We shall naturally identify the set $\mathbb{R}^{\text{Card}(\mathcal{T})}$ to $X(\mathcal{T})$ and then we can write $u_T = (u_K)_{K \in \mathcal{T}}$.

The finite volume scheme is classically obtained from the balance form of Equations (1) and (3) on a control volume K , that is:

$$-\int_{\partial K} \kappa(\cdot, u) \nabla \phi \cdot \mathbf{n}_K d\gamma(x) = \int_K f(\cdot, u) dx \quad (12)$$

$$-\int_{\partial K} \lambda(\cdot, u) \nabla u \cdot \mathbf{n}_K d\gamma(x) = \int_K \kappa(\cdot, u) |\nabla \phi|^2 dx, \quad (13)$$

where \mathbf{n}_K denotes the unit normal vector to ∂K outward to K and $d\gamma(x)$ is the integration symbol for the $(d-1)$ -dimensional Lebesgue measure. Let \mathcal{E}_K denote the set of edges or faces of ∂K , decomposing the boundary of K into edges or faces, $\partial K = \cup_{\sigma \in \mathcal{E}_K} \sigma$, we may rewrite (12)-(13) as:

$$-\sum_{\sigma \in \mathcal{E}_K} \int_{\sigma} \kappa(\cdot, u) \nabla \phi \cdot \mathbf{n}_{K,\sigma} d\gamma(x) = \int_K f(\cdot, u) dx \quad (14)$$

$$-\sum_{\sigma \in \mathcal{E}_K} \int_{\sigma} \lambda(\cdot, u) \nabla u \cdot \mathbf{n}_{K,\sigma} d\gamma(x) = \int_K \kappa(\cdot, u) |\nabla \phi|^2 dx, \quad (15)$$

where $\mathbf{n}_{K,\sigma}$ denotes the normal unit vector to σ outward to K . Let us write the sought approximations as $\phi_T = \sum_{K \in \mathcal{T}} \phi_K 1_K$ and $u_T = \sum_{K \in \mathcal{T}} u_K 1_K$ (see Remark 3.2); we then set

$$f_K(u_K) = \frac{1}{m(K)} \int_K f(x, u_K) dx. \quad (16)$$

Let \mathcal{E} denote the set of edges (or faces in 3D) of the mesh, and \mathcal{E}_{int} (resp. \mathcal{E}_{ext}) the set of edges laying in Ω (resp. on $\partial\Omega$). For $\sigma \in \mathcal{E}$, let $F_{K,\sigma}^{\kappa}(\phi_T)$ (resp. $F_{K,\sigma}^{\lambda}(u_T)$) be an approximation of the flux $\int_{\sigma} \kappa(x, u(x)) \nabla \phi(x) \cdot \mathbf{n}_{K,\sigma} d\gamma(x)$ (resp. $\int_{\sigma} \lambda(x, u(x)) \nabla u(x) \cdot \mathbf{n}_{K,\sigma} d\gamma(x)$), and let $\mathcal{J}_K(u_T, \phi_T)$ denote an approximation of the nonlinear right-hand-side $\frac{1}{m(K)} \int_K \kappa(x, u(x)) |\nabla \phi|^2(x) dx$. With these notations, a finite volume approximation may then be written under the form:

$$\begin{cases} \sum_{\sigma \in \mathcal{E}_K} F_{K,\sigma}^{\kappa}(\phi_T) = m(K) f_K(u_K), \quad \forall K \in \mathcal{T}, \\ \sum_{\sigma \in \mathcal{E}_K} F_{K,\sigma}^{\lambda}(u_T) = m(K) \mathcal{J}_K(u_T, \phi_T), \quad \forall K \in \mathcal{T}, \end{cases} \quad (17)$$

provided one defines the expressions $F_{K,\sigma}^{\kappa}(\phi_T)$, $F_{K,\sigma}^{\lambda}(u_T)$ and $\mathcal{J}_K(u_T, \phi_T)$ with respect to the discrete unknowns $(\phi_K)_{K \in \mathcal{T}}$ and $(u_K)_{K \in \mathcal{T}}$. The discrete fluxes are given by the classical two-points formula:

$$F_{K,\sigma}^{\kappa}(\phi_T) = \begin{cases} m(\sigma) \tau_{\sigma}^{\kappa}(u_T) (\phi_K - \phi_L) & \text{if } \sigma \in \mathcal{E}_{\text{int}}, \sigma = K|L, \\ m(\sigma) \tau_{\sigma}^{\kappa}(u_T) \phi_K & \text{if } \sigma \in \mathcal{E}_K \cap \mathcal{E}_{\text{ext}}, \end{cases} \quad (18)$$

$$F_{K,\sigma}^{\lambda}(u_T) = \begin{cases} m(\sigma) \tau_{\sigma}^{\lambda}(u_T) (u_K - u_L) & \text{if } \sigma \in \mathcal{E}_{\text{int}}, \sigma = K|L, \\ m(\sigma) \tau_{\sigma}^{\lambda}(u_T) (u_K) & \text{if } \sigma \in \mathcal{E}_K \cap \mathcal{E}_{\text{ext}}, \end{cases} \quad (19)$$

where τ_σ^κ (and, similarly τ_σ^λ) is defined through a harmonic average, that is:

$$\tau_\sigma^\kappa(u_T) = \begin{cases} \frac{\kappa_K(u_K)\kappa_L(u_L)}{d_{K,\sigma}\kappa_L(u_L) + d_{L,\sigma}\kappa_K(u_K)} & \text{if } \sigma \in \mathcal{E}_{\text{int}}, \sigma = K|L, \\ \frac{\kappa_K(u_K)}{d_{K,\sigma}} & \text{if } \sigma \in \mathcal{E}_{\text{ext}} \cap \mathcal{E}_K, \end{cases} \quad (20)$$

where the values $\kappa_K(u_K)$ and $\lambda_K(u_K)$ are defined by (16), replacing f by κ or λ .

The term $\mathcal{J}_K(u_T, \phi_T)$ is defined as:

$$\mathcal{J}_K(u_T, \phi_T) = \frac{1}{\mathfrak{m}(K)} \sum_{\sigma \in \mathcal{E}_K} \mathfrak{m}(\mathcal{D}_{K,\sigma}) \mathcal{J}_\sigma(u_T, \phi_T), \quad (21)$$

where, for $K \in \mathcal{T}$ and $\sigma \in \mathcal{E}_K$, we define the half dual cell $\mathcal{D}_{K,\sigma}$ as the convex hull of x_K and σ (see Figure 1), that is:

$$\mathcal{D}_{K,\sigma} = \{tx_K + (1-t)x, (x, t) \in \sigma \times (0, 1)\},$$

and

$$\mathcal{J}_\sigma(u_T, \phi_T) = \frac{\tau_\sigma^\kappa(u_T)}{d_\sigma} (D_\sigma \phi)^2 d, \quad (22)$$

with

$$D_\sigma \phi = \begin{cases} |\phi_K - \phi_L| & \text{if } \sigma \in \mathcal{E}_{\text{int}}, \sigma = K|L, \\ |\phi_K| & \text{if } \sigma \in \mathcal{E}_{\text{ext}} \end{cases} \quad (23)$$

We show in Theorem 3.1 below the existence of $(\phi_K)_{K \in \mathcal{T}}$ and $(u_K)_{K \in \mathcal{T}}$ solution to (17)–(22). This entitles us to define the functions ϕ_T and $u_T \in X(\mathcal{T})$ with respective values ϕ_K and u_K on cell K , along with the function $\mathcal{J}_T(u_T, \phi_T) \in X(\mathcal{T})$ with value $\mathcal{J}_K(u_T, \phi_T)$ on cell K .

Remark 3.3 (Relation between the finite volume scheme (17)–(22) and the finite volume scheme of [23]). *The above finite volume scheme may be seen as slight modification of a scheme which was first introduced in [23]; this scheme was based on the following integration by parts of the right-hand-side of equation (3):*

$$\begin{aligned} \int_K \kappa(\cdot, u) |\nabla \phi|^2 dx &= \int_{\partial K} \kappa(\cdot, u) \nabla \phi \cdot \mathbf{n}_K \phi d\gamma(x) - \int_K \nabla(\kappa(\cdot, u) \nabla \phi) \phi dx \\ &= \int_{\partial K} \kappa(\cdot, u) \nabla \phi \cdot \mathbf{n}_K \phi d\gamma(x) + \int_K f(\cdot, u) \phi dx, \end{aligned}$$

where \mathbf{n}_K denotes the unit vector normal to ∂K , outward to K . This formulation suggests the following approximation $\tilde{\mathcal{J}}_K(u_T, \phi_T)$ to $\frac{1}{\mathfrak{m}(K)} \int_K \kappa(\cdot, u) |\nabla \phi|^2 dx$:

$$\tilde{\mathcal{J}}_K(u_T, \phi_T) = f_K(u_K) \phi_K - \frac{1}{\mathfrak{m}(K)} \sum_{\sigma \in \mathcal{E}_K} F_{K,\sigma}^\kappa(\phi_T) \phi_\sigma, \quad (24)$$

where $F_{K,\sigma}^\kappa(\phi_T)$ is defined by (18), $f_K(u_K)$ is defined by (16), and ϕ_σ is an auxiliary value of ϕ_T on the interface, which may be eliminated:

$$\kappa_K(u_K) \frac{\phi_\sigma - \phi_K}{d_{K,\sigma}} + \kappa_L(u_L) \frac{\phi_\sigma - \phi_L}{d_{L,\sigma}} = 0, \quad \forall \sigma = K|L, \text{ and } \phi_\sigma = 0, \quad \forall \sigma \in \mathcal{E}_{\text{ext}}. \quad (25)$$

An easy computation shows that in fact,

$$\tilde{\mathcal{J}}_K(u_T, \phi_T) = \frac{1}{\mathfrak{m}(K)} \sum_{\sigma \in \mathcal{E}_K} \mathfrak{m}(\mathcal{D}_{K,\sigma}) \tilde{\mathcal{J}}_\sigma(\phi_T),$$

with

$$\tilde{\mathcal{J}}_\sigma(\phi_T) = \frac{\tau_\sigma^\kappa(u_T)}{d_\sigma} \mu_{K,\sigma} (D_\sigma \phi)^2 d \text{ and } \mu_{K,\sigma} = \frac{\kappa_L(u_L) d_\sigma}{\kappa_L(u_L) d_{K,\sigma} + \kappa_K(u_K) d_{L,\sigma}}.$$

Therefore, $\tilde{\mathcal{J}}_K(u_T, \phi_T) = \mathcal{J}_K(u_T, \phi_T)$ in the case of a homogeneous coefficient κ .

We show in Theorem 3.1 below the existence of $(\phi_K)_{K \in \mathcal{T}}$ and $(u_K)_{K \in \mathcal{T}}$ solution to (17)–(22). This entitles us to define the functions ϕ_T and $u_T \in X(\mathcal{T})$ with respective values ϕ_K and u_K on cell K , along with the function $\mathcal{J}_T(u_T, \phi_T) \in X(\mathcal{T})$ with value $\mathcal{J}_K(u_T, \phi_T)$ on cell K .

3.3. Existence of a discrete solution. To prove the existence of a finite element solution (ϕ_M, u_M) to the Problem (8) and a finite volume solution (ϕ_T, u_T) to the Problem (17)–(22), we use Brouwer’s theorem.

Theorem 3.1. *Let (κ, λ, f) be three functions satisfying the Assumption 1.*

1. *Let \mathcal{M} be a finite element simplicial mesh satisfying Assumption 2, and $V_0^{\mathcal{M}}$ be the linear finite element space defined by (7). Then there exists at least a solution $(u_M, \phi_M) \in (V_0^{\mathcal{M}})^2$ to the problem (8).*
2. *Let \mathcal{T} be an admissible mesh in the sense of Definition 2. Let $X(\mathcal{T})$ be the finite volume space defined by (11). Then there exists at least a solution $(u_T, \phi_T) \in (X(\mathcal{T}))^2$ to the Problem (17)–(22).*

Proof. The proof is based on the fixed point theorem. In fact, the existence of a solution to (6) was proven in [25] using Schauder’s fixed point theorem; here, since the spaces are finite-dimensional, we need only use Brouwer’s theorem. The proof is a rather easy adaptation of that of [25] and we only outline it.

1. For $u_M \in V_M^0$, let $\bar{u}_M = \mathcal{F}_M(u_M)$ be the unique solution (thanks to the Lax-Milgram lemma) to

$$\int_{\Omega} \lambda(x, u_M(x)) \nabla \bar{u}_M(x) \cdot \nabla v(x) \, dx = \int_{\Omega} \kappa(x, u_M(x)) |\nabla \phi_M|^2(x) v(x) \, dx, \quad \forall v \in V_0^{\mathcal{M}},$$

where $\phi_M \in V_0^{\mathcal{M}}$ is the unique solution to:

$$\int_{\Omega} \kappa(x, u_M(x)) \nabla \phi_M(x) \cdot \nabla \psi_M(x) \, dx = \int_{\Omega} f(x, u_M(x)) \psi_M(x) \, dx, \quad \forall \psi_M \in V_0^{\mathcal{M}}. \quad (26)$$

2. For $u_T \in X(\mathcal{T})$, let $\bar{u}_T = \mathcal{F}_T(u_T)$ be the unique solution (thanks to the classical techniques of [17]) to

$$\sum_{\sigma \in \mathcal{E}_K} \tau_\sigma^{\lambda(u_T)} (\bar{u}_L - \bar{u}_K) = \sum_{\sigma \in \mathcal{E}_K} m(\mathcal{D}_{K,\sigma}) \tau_\sigma^{\kappa(u_T)} \frac{(\phi_L - \phi_K)^2}{d_{K|L}} d, \quad \forall K \in \mathcal{T},$$

where $\tau_\sigma^{\lambda(u_T)}$ and $\tau_\sigma^{\kappa(u_T)}$ are defined by (20), (Note that we have denoted $\sigma = K|L$ if $\sigma \in \mathcal{E}_{\text{int}}$ and if $\sigma \in \mathcal{E}_{\text{ext}} \cap \mathcal{E}_K$, $(\bar{u}_L, \phi_L) = (0, 0)$), and let $\phi_T \in X(\mathcal{T})$ be the unique solution to:

$$\sum_{\sigma \in \mathcal{E}_K} \tau_\sigma^{\kappa(u_T)} (\phi_L - \phi_K) = m(K) f_K, \quad \forall K \in \mathcal{T}. \quad (27)$$

It is clear that if $u_{\mathcal{M}} = \mathcal{F}_{\mathcal{M}}(u_{\mathcal{M}})$ (resp. $u_{\mathcal{T}} = \mathcal{F}_{\mathcal{T}}(u_{\mathcal{T}})$) then $(u_{\mathcal{M}}, \phi_{\mathcal{M}})$ (resp. $(u_{\mathcal{T}}, \phi_{\mathcal{T}})$) is a solution to (8) (resp. (17)-(21)), where $\phi_{\mathcal{M}}$ (resp. $\phi_{\mathcal{T}}$) is defined by (26) (resp. (27)).

We then remark that, thanks to Assumption 1, the mappings $\mathcal{F}_{\mathcal{M}}$ and $\mathcal{F}_{\mathcal{T}}$ map the spaces $V_{\mathcal{M}}^0$ and $X(\mathcal{T})$ into a closed ball, and that they are continuous. Hence we may apply Brouwer's theorem which implies the existence of a solution to both schemes. \square

4. CONVERGENCE OF THE FINITE ELEMENT APPROXIMATION

Let us start with the following easy result, which we shall use in the convergence proof:

Lemma 4.1. *Under Assumption 1, let \mathcal{M} be a finite element mesh in the sense of Definition 1. Let $u_{\mathcal{M}} = \sum_{i \in \mathcal{I}} u_i \xi_i$ and $v_{\mathcal{M}} = \sum_{i \in \mathcal{I}} v_i \xi_i$ be some functions of $V_0^{\mathcal{M}}$; then:*

$$\int_{\Omega} \lambda(\cdot, u_{\mathcal{M}}) \nabla u_{\mathcal{M}} \cdot \nabla v_{\mathcal{M}} \, dx = \sum_{(i,j) \in \mathcal{V}^2} \theta_{i,j}^{\lambda}(u_{\mathcal{M}}) (u_i - u_j) (v_i - v_j), \quad (28)$$

where $\theta_{i,j}^{\lambda}(u_{\mathcal{M}})$ is defined in (9). Of course, the same equality is true replacing λ by κ .

Proof. By definition of $u_{\mathcal{M}}$, $v_{\mathcal{M}}$ and $\theta_{i,j}^{\lambda}(u_{\mathcal{M}})$, one has:

$$\int_{\Omega} \lambda(\cdot, u_{\mathcal{M}}) \nabla u_{\mathcal{M}} \cdot \nabla v_{\mathcal{M}} = - \sum_{i \in \mathcal{I}} \sum_{j \in \mathcal{I}} \theta_{i,j}^{\lambda}(u_{\mathcal{M}}) u_i v_j.$$

Since $\sum_{j \in \mathcal{V}} \theta_{i,j}^{\lambda}(u_{\mathcal{M}}) = 0$ and $u_i = v_j = 0$, for all $(i, j) \in \mathcal{B}^2$, we obtain:

$$\int_{\Omega} \lambda(\cdot, u_{\mathcal{M}}) \nabla u_{\mathcal{M}} \cdot \nabla v_{\mathcal{M}} = - \sum_{i \in \mathcal{V}} \sum_{j \in \mathcal{V}} \theta_{i,j}^{\lambda}(u_{\mathcal{M}}) u_i (v_j - v_i).$$

Reordering the summations on i, j as a summation on the pairs (i, j) , we then get that:

$$\int_{\Omega} \lambda(\cdot, u_{\mathcal{M}}) \nabla u_{\mathcal{M}} \cdot \nabla v_{\mathcal{M}} = \sum_{(i,j) \in \mathcal{V}^2} \theta_{i,j}^{\lambda}(u_{\mathcal{M}}) (u_i - u_j) (v_i - v_j), \quad (29)$$

which proves the Lemma. \square

Let us define the usual interpolation operator:

Definition 3 (Finite element interpolator). *Let \mathcal{M} be a simplicial finite element mesh of Ω in the sense of Definition 1. The interpolation operator into the finite element space $V_0^{\mathcal{M}}$ is defined by:*

$$\Pi_{\mathcal{M}} u = \sum_{i \in \mathcal{I}} u(x_i) \xi_i,$$

for any $u \in \mathcal{C}(\bar{\Omega})$ such that $u = 0$ on $\partial\Omega$. Note that for any $u \in \mathcal{C}(\bar{\Omega})$ such that $u = 0$ on $\partial\Omega$, one has:

$$\|\Pi_{\mathcal{M}} u\|_{L^{\infty}(\bar{\Omega})} \leq \|u\|_{L^{\infty}(\bar{\Omega})}. \quad (30)$$

Lemma 4.2. *Under Assumption 1, let \mathcal{M} be a finite element mesh satisfying Assumption 2. Let $(\phi_{\mathcal{M}}, u_{\mathcal{M}})$ be a solution of (17)-(21). Then the the following estimates hold:*

$$\|\nabla \phi_{\mathcal{M}}\|_{L^2(\Omega)} \leq C_1 \quad (31)$$

$$\|\kappa(\cdot, u_{\mathcal{M}}) \nabla \phi_{\mathcal{M}} \cdot \nabla \phi_{\mathcal{M}}\|_{L^1(\Omega)} \leq C_2, \quad (32)$$

where

$$C_1 = \frac{C_p}{\alpha} \|f\|_{0,\Omega},$$

and

$$C_2 = C_1^2 \|\kappa\|_{L^\infty(\Omega \times \mathbb{R}, \mathbb{R})},$$

and C_p is the Poincaré's constant.

Let $\psi \in L^\infty(\mathbb{R})$ be an absolutely increasing continuous (that is almost everywhere derivable and integral of its derivative). Define $\Psi(s) = \int_0^s \sqrt{\psi'(t)} dt$. Then the following estimate holds:

$$\|\nabla \Pi_{\mathcal{M}} \Psi(u_{\mathcal{M}})\|_{L^2(\Omega)} \leq \frac{C_2}{\alpha} \|\psi\|_\infty, \quad (33)$$

where the interpolation $\Pi_{\mathcal{M}}$ is defined in Definition 3.

Proof. Estimate (31) is clearly obtained by taking $\phi_{\mathcal{M}}$ as a test function in the first equation of (8). Using the fact that κ is bounded, one immediately gets (32).

Finally, noting that $\Pi_{\mathcal{M}} \psi(u_{\mathcal{M}}) \in V_0^{\mathcal{M}}$, where $\Pi_{\mathcal{M}}$ is defined in Definition 3, we may take it as a test function in the second equation of (8), which yields:

$$\int_{\Omega} \lambda(x, u_{\mathcal{M}}(x)) \nabla u_{\mathcal{M}}(x) \cdot \nabla \Pi_{\mathcal{M}}(u_{\mathcal{M}})(x) dx = \int_{\Omega} \kappa(x, u_{\mathcal{M}}(x)) |\nabla \phi_{\mathcal{M}}|^2(x) \Pi_{\mathcal{M}} \psi(u_{\mathcal{M}})(x) dx.$$

Noting that $\Pi_{\mathcal{M}} \psi(u_{\mathcal{M}}) = \sum_{i \in \mathcal{I}} \psi(u_i) \xi_i$, where $u_i = u(x_i)$ for any $i \in \mathcal{I}$, and applying Lemma 4.1 yields

$$\sum_{(i,j) \in \mathcal{V}^2} \theta_{i,j}^\lambda(u_{\mathcal{M}})(u_i - u_j)(\psi(u_i) - \psi(u_j)) = \int_{\Omega} \kappa(x, u_{\mathcal{M}}(x)) |\nabla \phi_{\mathcal{M}}|^2(x) \Pi_{\mathcal{M}} \psi(u_{\mathcal{M}})(x) dx.$$

Now since $\psi \in L^\infty(\mathbb{R})$, we get from (32) and (30) that

$$\sum_{(i,j) \in \mathcal{V}^2} \theta_{i,j}^\lambda(u_{\mathcal{M}})(u_i - u_j)(\psi(u_i) - \psi(u_j)) \leq C_2 \|\psi\|_{L^\infty(\mathbb{R})}$$

Now by the Cauchy–Schwarz inequality, we have:

$$(\Psi(a) - \Psi(b))^2 \leq (a - b)(\psi(a) - \psi(b)), \quad \forall (a, b) \in \mathbb{R}^2,$$

and therefore, since $\theta_{i,j}^\lambda(u_{\mathcal{M}}) \geq 0$ for any $(i, j) \in \mathcal{I} \times \mathcal{V}$ such that $i \neq j$ (thanks to Assumption 2), and $\Psi(u_i) = \Psi(u_j) = 0$, for all $(i, j) \in \mathcal{B}^2$, we get that:

$$\sum_{(i,j) \in \mathcal{V}^2} \theta_{i,j}^\lambda(u_{\mathcal{M}})(\Psi(u_i) - \Psi(u_j))^2 \leq C_2 \|\psi\|_{L^\infty(\mathbb{R})}$$

Applying Lemma 4.1 once more, we finally get that:

$$\int_{\Omega} \lambda(\cdot, u) |\nabla \Pi_{\mathcal{M}} \Psi(u)|^2 \leq C_2 \|\psi\|_{L^\infty(\mathbb{R})} \quad (34)$$

which concludes the proof of the Lemma, since λ is bounded by below. \square

From the above Lemma, one deduces that (33) is true for the function ψ defined by $\psi(s) = \int_0^s \frac{dt}{1+|t|^\theta}$ for some given $\theta > 1$; hence if one could get rid of the interpolator $\Pi_{\mathcal{M}}$ in (33), then one could apply the result of [7] to get that $u_{\mathcal{M}}$ is bounded in $W_0^{1,q}(\Omega)$. However, this does not seem straightforward in general; hence, in order to get some compactness, one may adapt the technique of [4], as performed in [9, Theorem 3.1] to show that (33) implies that $u_{\mathcal{M}}$ is bounded in $W_0^{1,q}(\Omega)$. From this compactness result, we get the convergence Theorem 4.1 (given below).

Remark 4.1. *In the two dimensional case, under the assumption that the mesh \mathcal{M} satisfies Delaunay and “non degeneracy” conditions (see [24]), it is possible to prove that $u_{\mathcal{M}}$ is bounded in $W_0^{1,q}(\Omega)$ by using finite volume techniques.*

4.1. Convergence result.

Theorem 4.1. *Let Ω be a convex polygonal open subset of \mathbb{R}^d , $d = 2$ or 3 , and κ, λ and f be three functions satisfying Assumption 1. Let $(\mathcal{M}_n)_{n \in \mathbb{N}}$ be a sequence of finite element simplicial meshes satisfying Assumption 2, such that $h_{\mathcal{M}_n} \rightarrow 0$, as $n \rightarrow \infty$. Then, there exists a subsequence, still denoted by $(\mathcal{M}_n)_{n \in \mathbb{N}}$ and a solution $(\phi_{\mathcal{M}_n}, u_{\mathcal{M}_n})$ to (8), such that $(\phi_{\mathcal{M}_n}, u_{\mathcal{M}_n})$ converges to a weak solution $(\phi, u) \in H_0^1(\Omega) \times \bigcap_{p < \frac{d}{d-1}} W_0^{1,p}(\Omega)$ of (6), in the following sense:*

1. $\phi_{\mathcal{M}_n}$ converges to ϕ in $H_0^1(\Omega)$ as $n \rightarrow +\infty$.
2. $u_{\mathcal{M}_n}$ converges to u weakly in $W_0^{1,p}(\Omega)$, for all $p \in [1, \frac{d}{d-1})$.

Furthermore,

$$\kappa(\cdot, \mathcal{M}_n) \nabla \phi_{\mathcal{M}_n} \cdot \nabla \phi_{\mathcal{M}_n} \rightarrow \kappa(\cdot, u) \nabla \phi \cdot \nabla \phi \text{ in } L^1(\Omega) \text{ as } n \rightarrow +\infty. \quad (35)$$

Proof. Let ψ and v be two functions in $C_c^\infty(\Omega)$ (that is the space of infinitely differentiable functions, with compact support on Ω). Let $\psi_n = \Pi_{\mathcal{M}_n} \psi$ and $v_n = \Pi_{\mathcal{M}_n} v$, (see Definition 3). Since $\psi_n \in V_0^{\mathcal{M}_n}$ and $v_n \in V_0^{\mathcal{M}_n}$, we may take them as test functions in (8) (for $\mathcal{M} = \mathcal{M}_n$). Hence ϕ_n and u_n satisfy:

$$\begin{cases} \int_{\Omega} \kappa(\cdot, u_n) \nabla \phi_n \cdot \nabla \psi_n \, dx = \int_{\Omega} f(\cdot, u_n) \psi_n \, dx, \\ \int_{\Omega} \lambda(\cdot, u_n) \nabla u_n \cdot \nabla v_n \, dx = \int_{\Omega} \kappa(\cdot, u_n) |\nabla \phi_n|^2 v_n \, dx. \end{cases} \quad (36)$$

From Estimate (31), we get by Rellich’s theorem that ϕ_n tends (up to a subsequence) to some function $\phi \in H_0^1(\Omega)$ in $L^2(\Omega)$ as $n \rightarrow +\infty$. From estimate (33) and Theorem 3.1 of [9], we get that u_n tends (up to a subsequence) to some function $u \in W_0^{1,q}(\Omega)$ weakly in $W_0^{1,q}(\Omega)$ as $n \rightarrow +\infty$. Furthermore, $\psi_n \rightarrow \psi$ and $v_n \rightarrow v$ in $W^{1,\infty}(\Omega)$, as $n \rightarrow +\infty$. Thanks to Assumption 1, κ and f are bounded, so that, by the Lebesgue dominated theorem, up to a subsequence, $\kappa(\cdot, u_n) \nabla \psi_n \rightarrow \kappa(\cdot, u) \nabla \psi$ in $L^2(\Omega)$, and $f(\cdot, u_n) \psi_n \rightarrow f(\cdot, u) \psi$ in $L^1(\Omega)$, as $n \rightarrow +\infty$ (these two previous convergences also hold in $L^p(\Omega)$, for any $p \in [1, \infty)$). We may therefore pass to the limit in the first equation of (36), to obtain that u and ϕ satisfy:

$$\int_{\Omega} \kappa(\cdot, u) \nabla \phi \cdot \nabla \psi \, dx = \int_{\Omega} f(\cdot, u) \psi \, dx. \quad (37)$$

Since ψ is arbitrary in (37), then, thanks to the density of $C_c^\infty(\Omega)$ in $H_0^1(\Omega)$, we get

$$\int_{\Omega} \kappa(\cdot, u) \nabla \phi \cdot \nabla \psi \, dx = \int_{\Omega} f(\cdot, u) \psi \, dx, \quad \forall \psi \in H_0^1(\Omega). \quad (38)$$

Now, by Assumption 1, κ is bounded by below, and we get that:

$$\alpha \int_{\Omega} |\nabla(\phi_n - \phi)|^2 \, dx \leq \int_{\Omega} \kappa(\cdot, u_n) \nabla(\phi_n - \phi) \cdot \nabla(\phi_n - \phi) \, dx = \mathbb{T}_1^n + \mathbb{T}_2^n + \mathbb{T}_3^n, \quad (39)$$

with:

$$\begin{aligned} \mathbb{T}_1^n &= \int_{\Omega} \kappa(\cdot, u_n) \nabla \phi_n \cdot \nabla \phi_n \, dx, \\ \mathbb{T}_2^n &= -2 \int_{\Omega} \kappa(\cdot, u_n) \nabla \phi_n \cdot \nabla \phi \, dx, \end{aligned}$$

and

$$\mathbb{T}_3^n = \int_{\Omega} \kappa(\cdot, u_n) \nabla \phi \cdot \nabla \phi \, dx.$$

Since (ϕ_n, u_n) is a solution to (8), one could take ϕ_n as a test function in the first equation of (8):

$$\mathbb{T}_1^n = \int_{\Omega} f(\cdot, u_n) \phi_n \, dx \rightarrow \int_{\Omega} f(\cdot, u) \phi \, dx, \quad \text{as } n \rightarrow +\infty. \quad (40)$$

Hence, since ϕ and u satisfy (38), the previous limit becomes as:

$$\mathbb{T}_1^n \rightarrow \int_{\Omega} \kappa(\cdot, u) \nabla \phi \cdot \nabla \phi \, dx, \quad \text{as } n \rightarrow +\infty. \quad (41)$$

Furthermore, by Lebesgue's theorem, $\kappa(\cdot, u_n) \nabla \phi \rightarrow \kappa(\cdot, u) \nabla \phi$ in $(L^2(\Omega))^2$; since $\nabla \phi_n \rightarrow \nabla \phi$ weakly in $L^2(\Omega)$, one gets that

$$\mathbb{T}_2^n \rightarrow -2 \int_{\Omega} \kappa(\cdot, u) \nabla \phi \cdot \nabla \phi \, dx, \quad \text{as } n \rightarrow +\infty. \quad (42)$$

It is then clear that we also have $\mathbb{T}_3^n \rightarrow \int_{\Omega} \kappa(\cdot, u) \nabla \phi \cdot \nabla \phi \, dx$ as $n \rightarrow +\infty$, this with (39), (41) and (42) imply that ϕ_n tends to ϕ in $H_0^1(\Omega)$ as $n \rightarrow +\infty$. We then immediately obtain (35).

Let us then pass to the limit in the second equation of (36) to show that (ϕ, u) is a solution to (6). From (35), we immediately get that:

$$\int_{\Omega} \kappa(\cdot, u_n) |\nabla \phi_n|^2 v_n \, dx \rightarrow \int_{\Omega} \kappa(\cdot, u) |\nabla \phi|^2 v \, dx. \quad (43)$$

Since u_n converges to u weakly in $W_0^{1,p}(\Omega)$, again using the Lebesgue theorem, we may pass to the limit (up to a subsequence) in the left-hand-side of the second equation of (36) and using (43) obtain that:

$$\int_{\Omega} \lambda(\cdot, u) \nabla u \cdot \nabla v \, dx = \int_{\Omega} \kappa(\cdot, u) |\nabla \phi|^2 v \, dx.$$

This concludes the proof of Theorem 4.1. \square

Remark 4.2. Note that our convergence result does not require the usual regularity condition on the mesh which assumes that the ratio $\sup_{T \in \mathcal{M}} \frac{h_T}{\rho_T}$, where h_T and ρ_T are respectively the diameter of T and the diameter of the ball inscribed in T , be bounded independently on the mesh. This is due to the fact that in the proof, we pass to the limit directly in the scheme, without using the classical finite element error estimates which are known in the variational setting. However, this condition is required in order to obtain error estimates by interpolation, see [12, 9].

5. CONVERGENCE OF THE FINITE VOLUME APPROXIMATION

5.1. The convergence result. In this section, we shall prove that a solution of (17)-(22) converges, as $h_{\mathcal{T}} = \sup\{\text{diam}(K), K \in \mathcal{T}\}$ tends to 0, towards a solution of (6), as stated in the following theorem:

Theorem 5.1. Under Assumption 1, let $(\mathcal{T}_n)_{n \in \mathbb{N}}$ be a sequence of admissible meshes in the sense of Definition 2. Let (ϕ^n, u^n) be a solution of the system (17)-(22) for $\mathcal{T} = \mathcal{T}_n$, and let $\mathcal{J}^n(u^n, \phi^n)$ be defined by (21). Assume that $h_n = \sup\{\text{diam}(K), K \in \mathcal{T}_n\} \rightarrow 0$, as $n \rightarrow \infty$, and that there exists $\zeta > 0$ (not depending on n), such that:

$$d_{\sigma} \leq \zeta d_{K,\sigma}, \quad \forall \sigma \in \mathcal{E}_n, \quad \forall K \in \mathcal{T}_n. \quad (44)$$

Then, there exists a subsequence of $(\mathcal{T}_n)_{n \in \mathbb{N}}$, still denoted by $(\mathcal{T}_n)_{n \in \mathbb{N}}$, such that (ϕ^n, u^n) converges to a solution $(\phi, u) \in H_0^1(\Omega) \times \cap_{q < \frac{d}{d-1}} W_0^{1,q}(\Omega)$ of (6), as $n \rightarrow \infty$, in the following sense:

$$\|\phi^n - \phi\|_{L^2(\Omega)} \rightarrow 0, \text{ as } n \rightarrow +\infty, \quad (45)$$

$$\|u^n - u\|_{L^p(\Omega)} \rightarrow 0, \text{ as } n \rightarrow +\infty, \text{ for all } p < \frac{d}{d-2}. \quad (46)$$

Moreover,

$$\int_{\Omega} \mathcal{J}^n(u^n, \phi^n)(x) dx \rightarrow \int_{\Omega} \kappa(x, u(x)) |\nabla \phi|^2(x) dx \text{ as } n \rightarrow +\infty. \quad (47)$$

Proof. For the sake of clarity, we only list here the main steps of the proof and refer to the lemmata proven below for the details.

Let $(\phi^n)_{n \in \mathbb{N}} \subset L^2(\Omega)$ and $(u^n)_{n \in \mathbb{N}} \subset L^2(\Omega)$ be such that, for any $n \in \mathbb{N}$, the pair $(\phi^n, u^n) \in (X(\mathcal{T}_n))^2$ is a solution of (17)–(22), with $T = \mathcal{T}_n$ (recall that this solution exists by Theorem 3.1).

- (1) *A priori estimates.* We first show in Lemma 5.1 below that the sequences $(\phi^n)_{n \in \mathbb{N}}$ and $(u^n)_{n \in \mathbb{N}}$ are bounded for respectively, the L^2 norm and the L^p norm, with $p < \frac{d}{d-2}$. Note that the condition (44) is required when using the discrete Sobolev inequality, see e.g. [13], to obtain the uniform bound of $(u^n)_{n \in \mathbb{N}}$ in an L^q norm from a discrete $W_0^{1,p}$ estimate.
- (2) *Estimates on the space translates.* Following [17, Lemma 9.3 page 770], [18, Lemma 4] or [26], one may then easily, using (54) and (57), get some uniform estimates on the translates of ϕ^n in the L^2 norm and of u^n in the L^p norm.
- (3) *Relative compactness.* We may therefore use a discrete Rellich theorem (see e.g. [18, Theorem 1]) to obtain that the sequences $(\phi^n)_{n \in \mathbb{N}}$ and $(u^n)_{n \in \mathbb{N}}$ are relatively compact in, respectively, $L^2(\Omega)$ and $L^p(\Omega)$, for $p < \frac{d}{d-2}$. The estimates on the translations also yield the regularity of the limit, that is, if ϕ is a limit of the sequence $(\phi^n)_{n \in \mathbb{N}}$ in $L^2(\Omega)$, then $\phi \in H_0^1(\Omega)$; similarly, if u is a limit of the sequence $(u^n)_{n \in \mathbb{N}}$ in $L^p(\Omega)$, then $u \in \cap_{q < \frac{d}{d-1}} W_0^{1,q}(\Omega)$.
- (4) *Passage to the limit in the scheme.* From step (3), for any sequence $(\mathcal{T}_n)_{n \in \mathbb{N}}$ of admissible meshes satisfying (44) and such that $\text{size}(\mathcal{T}_n) \rightarrow 0$, as $n \rightarrow \infty$, there exists a subsequence, still denoted by $(\mathcal{T}_n)_{n \in \mathbb{N}}$, such that:

(a) u^n converges to some $u \in \cap_{q < \frac{d}{d-1}} W_0^{1,q}(\Omega)$ in $L^p(\Omega)$, for all $p < \frac{d}{d-2}$, as $n \rightarrow \infty$.

(b) ϕ^n converges to some $\phi \in H_0^1(\Omega)$, in $L^2(\Omega)$, as $n \rightarrow \infty$.

As in the proof of [18, Theorem 2], we first multiply the first equation of (17) by $\psi(x_K)$, with $\psi \in C_c^\infty(\Omega)$; thanks to a discrete summation by parts, we obtain:

$$\sum_{K \in \mathcal{T}_n} \sum_{\sigma \in \mathcal{E}_K} \frac{m(\sigma)}{d_{K,\sigma}} \kappa_K(u_K^n) (\phi_K^n - \phi_\sigma^n) (\psi_K^n - \psi_\sigma^n) = \sum_{K \in \mathcal{T}_n} m(K) f_K(u_K) \psi(x_K), \quad (48)$$

with $\psi_K^n = \psi(x_K)$, where ψ_σ^n is defined by:

$$\begin{cases} \psi_\sigma^n = \frac{d_{K,\sigma} \psi(x_L) + d_{L,\sigma} \psi(x_K)}{d_\sigma} & \text{if } \sigma = K|L \in \mathcal{E}_{\text{int}}, \\ \psi_\sigma^n = 0 & \text{if } \sigma \in \mathcal{E}_{\text{ext}}. \end{cases} \quad (49)$$

and ϕ_σ^n is defined by:

$$\begin{cases} \kappa_K(u_K^n) \frac{\phi_\sigma^n - \phi_K^n}{d_{K,\sigma}} + \kappa_L(u_L^n) \frac{\phi_\sigma^n - \phi_L^n}{d_{L,\sigma}} = 0 & \text{if } \sigma = K|L \in \mathcal{E}_{\text{int}}, \\ \phi_\sigma^n = 0 & \text{if } \sigma \in \mathcal{E}_{\text{ext}}. \end{cases} \quad (50)$$

Now thanks to the assumptions on f , using the Lebesgue dominated theorem, we get that $\sum_{K \in \mathcal{T}_n} m(K) f_K(u_K) \psi(x_K) \rightarrow \int_{\Omega} f(u, \cdot) \psi dx$. Then, by Lemma 5.2 given below (with $v_n \equiv 1$), the left-hand-side of (48) tends to $\int_{\Omega} \kappa(u, \cdot) \nabla \phi \cdot \nabla \psi dx$. Hence the function

$\phi \in H_0^1(\Omega)$ is the (unique, for the considered function u) weak solution of the first equation of (6), that is:

$$\int_{\Omega} \kappa(x, u(x)) \nabla \phi(x) \cdot \nabla \psi(x) \, dx = \int_{\Omega} f(x, u(x)) \psi(x) \, dx, \quad \forall \psi \in H_0^1(\Omega).$$

In order to prove (45) and (46), there now only remains to show that u satisfies the second equation of (6). In order to do so, we proceed in a now classical way, that is, we multiply the second equation of the scheme (17) by $\psi(x_K)$ where ψ in $\mathcal{C}_c^\infty(\Omega)$ (the set of infinitely differentiable functions, with compact support on Ω), we sum over $K \in \mathcal{T}_n$, and obtain:

$$\sum_{K \in \mathcal{T}} \sum_{\sigma \in \mathcal{E}_K} m(\sigma) \tau_\sigma^\lambda(u^n) (u_K^n - u_L^n) \psi(x_K) = \sum_{K \in \mathcal{T}} m(K) \mathcal{J}_K^n(u^n, \phi^n) \psi(x_K) \quad (51)$$

Let us now pass to the limit as $n \rightarrow +\infty$. Applying Lemma 5.2 given below with $v \equiv 1$, we get that the left hand side of (51) tends to $\int_{\Omega} \lambda(x, u(x)) \nabla u(x) \cdot \nabla \psi(x) \, dx$, as $n \rightarrow +\infty$. Moreover, we show in Lemma 5.3 below that the right hand side of (51) tends to $\int_{\Omega} \kappa(x, u(x)) |\nabla \phi|^2(x) \psi(x) \, dx$, so that, by density of $\mathcal{C}_c^\infty(\Omega)$ in $W_0^{1,q}(\Omega)$, we get that u satisfies

$$\int_{\Omega} \lambda(x, u(x)) \nabla u(x) \cdot \nabla \psi(x) \, dx = \int_{\Omega} \kappa(x, u(x)) |\nabla \phi|^2(x) \psi(x) \, dx, \quad \forall \psi \in \cup_{q>d} W_0^{1,q}(\Omega). \quad (52)$$

The proof of (47) then follows by an adaptation of the proof of the convergence of the discrete H_0^1 norm in [17] (Theorem 9.1, proof page 776): see Lemma 5.4 below. This concludes the proof of the theorem. \square

In the following sections, we shall derive the estimates and the intermediate convergence results which were used in the above proof.

5.2. Estimate on the approximate solutions and compactness. Recall that the approximate finite volume solutions are piecewise constant; hence they are not, in general, in the spaces $W^{1,p}$, and we need therefore to define a discrete $W^{1,p}$ norm (see also [13, 17]) in order to obtain some compactness results.

Definition 4 (Discrete $W^{1,p}$ norm). *Let Ω be an open bounded subset of \mathbb{R}^d , $d = 2$ or 3 , and let \mathcal{T} be an admissible finite volume mesh in the sense of Definition 2. For $u_{\mathcal{T}} \in X(\mathcal{T})$ (defined in (11)), $u_{\mathcal{T}} = \sum_{K \in \mathcal{T}} u_K 1_K$, and $p \in [1, +\infty)$,*

$$\|u_{\mathcal{T}}\|_{1,p,\mathcal{T}} = \left(\sum_{\sigma \in \mathcal{E}} m(\sigma) d_\sigma \left(\frac{D_\sigma u}{d_\sigma} \right)^p \right)^{\frac{1}{p}},$$

with the notation

$$D_\sigma u = \begin{cases} |u_K - u_L| & \text{if } \sigma \in \mathcal{E}_{\text{int}}, \sigma = K|L, \\ |u_K| & \text{if } \sigma \in \mathcal{E}_{\text{ext}} \cap \mathcal{E}_K \end{cases}$$

To prove the convergence of $(\phi_{\mathcal{T}}, u_{\mathcal{T}})$, we prove at first some estimates on $\phi_{\mathcal{T}}$ and $u_{\mathcal{T}}$.

Lemma 5.1. *Under Assumption 1, let \mathcal{T} be an admissible mesh in the sense of Definition 2, and let $\zeta_{\mathcal{T}} > 0$ be such that:*

$$d_\sigma \leq \zeta_{\mathcal{T}} d_{K,\sigma}, \quad \forall \sigma \in \mathcal{E}, \text{ and for any } K \in \mathcal{T}. \quad (53)$$

Let $(\phi_{\mathcal{T}}, u_{\mathcal{T}})$ be a solution of (17)–(22). Then there exists $(C_3, C_4, C_5) \in (\mathbb{R}_+^)^3$, only depending on Ω , $\|f\|_{L^\infty(\Omega \times \mathbb{R}, \mathbb{R})}$, $\|\kappa\|_{L^\infty(\Omega \times \mathbb{R}, \mathbb{R})}$ and α such that*

$$\|\phi_{\mathcal{T}}\|_{1,2,\mathcal{T}} \leq C_3, \quad (54)$$

$$\|\phi_{\mathcal{T}}\|_{L^2(\Omega)} \leq C_4, \quad (55)$$

and

$$\|\mathcal{J}_{\mathcal{T}}(u_{\mathcal{T}}, \phi_{\mathcal{T}})\|_{L^1(\Omega)} \leq C_5. \quad (56)$$

Moreover, for all $p \in [1, \frac{d}{d-1})$, there exists a constant $C_6 \in \mathbb{R}_+^*$ only depending on Ω , $\|f\|_{L^\infty(\Omega \times \mathbb{R}, \mathbb{R})}$, $\|\kappa\|_{L^\infty(\Omega \times \mathbb{R}, \mathbb{R})}$, $\|\lambda\|_{L^\infty(\Omega \times \mathbb{R}, \mathbb{R})}$, $\zeta_{\mathcal{T}}$, p and α such that

$$\|u_{\mathcal{T}}\|_{1,p,\mathcal{T}} \leq C_6, \quad (57)$$

and a constant $C_7 \in \mathbb{R}_+^*$ only depending on Ω , $\|f\|_{L^\infty(\Omega \times \mathbb{R}, \mathbb{R})}$, $\|\kappa\|_{L^\infty(\Omega \times \mathbb{R}, \mathbb{R})}$, $\|\lambda\|_{L^\infty(\Omega \times \mathbb{R}, \mathbb{R})}$, $\zeta_{\mathcal{T}}$, p , α and d such that

$$\|u_{\mathcal{T}}\|_{L^{p^*}} \leq C_7, \quad (58)$$

where $p^* = \frac{pd}{d-p}$

Proof. The proof of (54) follows [17, Lemma 9.2 page 768] and the estimate (55) is then obtained by the discrete Poincaré inequality [17, Lemma 9.1 page 765]. Let us then prove the L^1 estimate (56). Indeed, by definition of $\mathcal{J}_{\mathcal{T}}(u_{\mathcal{T}}, \phi_{\mathcal{T}})$,

$$\|\mathcal{J}_{\mathcal{T}}(u_{\mathcal{T}}, \phi_{\mathcal{T}})\|_{L^1(\Omega)} = \sum_{K \in \mathcal{T}} \sum_{\sigma \in \mathcal{E}_K} m(\mathcal{D}_{K,\sigma}) \mathcal{J}_{\sigma}(\phi_{\mathcal{T}}) \quad (59)$$

$$= \sum_{\sigma \in \mathcal{E}} m(\mathcal{D}_{\sigma}) \mathcal{J}_{\sigma}(\phi_{\mathcal{T}}), \quad (60)$$

where \mathcal{D}_{σ} denotes the “diamond cell” around σ , that is $\mathcal{D}_{\sigma} = \mathcal{D}_{K,\sigma} \cup \mathcal{D}_{L,\sigma}$ if $\sigma = K|L \in \mathcal{E}_{\text{int}}$, and $\mathcal{D}_{\sigma} = \mathcal{D}_{K,\sigma}$ if $\sigma \in \mathcal{E}_{\text{ext}} \cap \mathcal{E}_K$. From the definition of $\mathcal{J}_{\sigma}(\phi_{\mathcal{T}})$, noting that $m(\mathcal{D}_{\sigma}) = \frac{1}{d}m(\sigma)d_{\sigma}$, and using Assumption 1, one then obtains that:

$$\begin{aligned} \|\mathcal{J}_{\mathcal{T}}(u_{\mathcal{T}}, \phi_{\mathcal{T}})\|_{L^1(\Omega)} &= \sum_{\sigma \in \mathcal{E}} m(\sigma) \tau_{\sigma}^{\kappa}(u_{\mathcal{T}}) |D_{\sigma} \phi|^2 \\ &\leq \sum_{\sigma \in \mathcal{E}} m(\sigma) d_{\sigma} \frac{\|\kappa\|_{L^\infty(\Omega \times \mathbb{R}, \mathbb{R})}^2}{\alpha} \left(\frac{D_{\sigma} \phi}{d_{\sigma}}\right)^2 \\ &\leq \frac{\|\kappa\|_{L^\infty(\Omega \times \mathbb{R}, \mathbb{R})}^2}{\alpha} \|\phi_{\mathcal{T}}\|_{1,2,\mathcal{T}}, \end{aligned}$$

which proves (56). Thanks to L^1 estimate (56), one obtains (57) by a straightforward adaptation of [26, Lemma 1] (see also [15, Theorem 2.2]). The estimate (58) follows from a discrete Sobolev inequality [13]. \square

5.3. Passage to the limit. Let us begin by a technical lemma which is used for the convergence of various terms in the passage of the limit.

Lemma 5.2. *Under Assumption 1, let $(\mathcal{T}_n)_{n \in \mathbb{N}}$ be a sequence of admissible meshes in the sense of Definition 2, such that $h_n = \sup\{\text{diam}(K), K \in \mathcal{T}_n\} \rightarrow 0$, as $n \rightarrow \infty$. Let $(u^n, \phi^n, v^n)_{n \in \mathbb{N}} \subset L^2(\Omega)^3$, with $(u^n, \phi^n, v^n) \in X(\mathcal{T}_n)^3 \forall n \in \mathbb{N}$ and let $\phi \in H_0^1(\Omega)$, $v \in L^2(\Omega)$ and $u \in L^p(\Omega)$, be such that:*

$$(\phi^n, v^n, u^n) \rightarrow (\phi, v, u) \text{ in } L^2(\Omega) \times L^2(\Omega) \times L^p(\Omega), \forall p < \frac{d}{d-2}, \text{ as } n \rightarrow +\infty.$$

Moreover, assume that there exists $C > 0$ such that $\|\phi^n\|_{1,\mathcal{T}_n} \leq C$, for all $n \in \mathbb{N}$. Let $\psi \in \mathcal{C}_c^\infty(\Omega)$ and $\psi^n \in X(\mathcal{T}_n)$ be defined by:

$$\psi^n(x) = \psi(x_K), \text{ for a.e. } x \in K, \forall K \in \mathcal{T}_n. \quad (61)$$

Define $I_n(\phi^n, v^n, u^n, \psi)$ by:

$$I_n(\phi^n, v^n, u^n, \psi) = \sum_{K \in \mathcal{T}_n} \sum_{\sigma \in \mathcal{E}_K} \frac{m(\sigma)}{d_{K,\sigma}} \kappa_K(u_K^n) (\phi_\sigma^n - \phi_K^n) (\psi_\sigma^n - \psi_K^n) v_K^n,$$

where ψ_σ^n is defined by (49) and ϕ_σ^n by (50). Then

$$I_n(\phi^n, v^n, u^n, \psi) \rightarrow \int_{\Omega} \kappa(\cdot, u) \nabla \phi \cdot \nabla \psi v \, dx \text{ as } n \rightarrow +\infty.$$

Proof. Let $\mathcal{G}^n(\phi^n) \in (L^2(\Omega))^d$ be the piecewise constant function equal to $\frac{\phi_\sigma^n - \phi_K^n}{d_{K,\sigma}} d \mathbf{n}_{K,\sigma}$ on the half diamond $\mathcal{D}_{K,\sigma}$. Let $\nabla_{\mathcal{T}_n} \psi^n \in (L^2(\Omega))^d$ be the piecewise constant function equal to $\frac{\psi_\sigma^n - \psi_K^n}{d_{K,\sigma}} \mathbf{n}_{K,\sigma} + (\nabla \psi \cdot \mathbf{t}_\sigma) \cdot \mathbf{t}_\sigma$ on the half diamond $\mathcal{D}_{K,\sigma}$, where \mathbf{t}_σ denotes a unit tangent vector to σ . Noting that $m(\mathcal{D}_{K,\sigma}) = \frac{1}{d} m(\sigma) d_{K,\sigma}$, we may write $I_n(\phi^n, v^n, u^n, \psi)$ in the following way:

$$I_n(\phi^n, v^n, u^n, \psi) = \int_{\Omega} \mathcal{G}^n(\phi^n)(x) \cdot \nabla_{\mathcal{T}_n} \psi^n(x) \kappa^n(x, u^n(x)) v^n(x) \, dx.$$

(Recall that $\kappa^n(\cdot, u^n(\cdot)) \in X(\mathcal{T})$ and $\kappa^n(x, u^n(x)) = \kappa_K(u_K^n)$, a.e. $x \in K$, for any $K \in \mathcal{T}_n$.)

By assumption, the sequence $(\varphi^n)_{n \in \mathbb{N}}$ converges to $\varphi \in H_0^1(\Omega)$ in $L^2(\Omega)$ and is bounded in the discrete H^1 norm; therefore, we get from Lemma 2 in [19]:

$$\mathcal{G}^n(\phi^n) \rightharpoonup_{n \rightarrow +\infty} \nabla \phi \text{ weakly in } (L^2(\Omega))^d. \quad (62)$$

Thanks to the definition (49) of ψ_σ^n , the differential quotient $\frac{\psi_\sigma^n - \psi_K^n}{d_{K,\sigma}}$ is a consistent approximation (in the finite difference sense) of $\nabla \psi \cdot \mathbf{n}_{K,\sigma}$; therefore, the function $\nabla_{\mathcal{T}_n} \psi^n$ converges to $\nabla \psi$ in $(L^\infty(\Omega))^d$. Hence, since $u^n \rightarrow u$ in the L^p norm for $p < \frac{d}{d-2}$ and since κ is bounded, we obtain from the Lebesgue dominated theorem that:

$$\kappa^n(\cdot, u^n) v^n \nabla_{\mathcal{T}_n} \psi^n \rightarrow \kappa(\cdot, u) v \nabla \psi \text{ in } (L^2(\Omega))^d \text{ as } n \rightarrow +\infty.$$

This, together with (62), concludes the proof of the Lemma. \square

Lemma 5.3 (Right-hand-side of the heat equation). *Under Assumption 1, let $(\mathcal{T}_n)_{n \in \mathbb{N}}$ be a sequence of admissible meshes in the sense of Definition 2. Let (ϕ^n, u^n) be a solution of the system (17)-(21) for $\mathcal{T} = \mathcal{T}_n$, and let $\mathcal{J}^n(u^n, \phi^n) \in X(\mathcal{T}_n)$ be defined by (21). Assume that $h_n = \max\{\text{diam}(K), K \in \mathcal{T}_n\} \rightarrow 0$, as $n \rightarrow \infty$, and that there exists $\zeta > 0$, not depending on n , such that (44) holds. Assume that*

1. u^n converges to some $u \in \cap_{q < \frac{d}{d-1}} W_0^{1,q}(\Omega)$ in $L^p(\Omega)$, for all $p < \frac{d}{d-2}$, as $n \rightarrow \infty$.
2. ϕ^n converges to some $\phi \in H_0^1(\Omega)$, in $L^2(\Omega)$, as $n \rightarrow \infty$.

For any $\psi \in \mathcal{C}_c^\infty(\Omega)$ (the set of infinitely differentiable functions with compact support on Ω), let $\psi^n \in X(\mathcal{T}_n)$ be defined by (61). Then:

$$\int_{\Omega} \mathcal{J}^n(u^n, \phi^n)(x) \psi^n(x) \, dx \rightarrow \int_{\Omega} \kappa(x, u(x)) |\nabla \phi|^2(x) \psi(x) \, dx \text{ as } n \rightarrow +\infty. \quad (63)$$

Proof. Noting that $m(\mathcal{D}_{K,\sigma}) = \frac{1}{d}m(\sigma)d_{K,\sigma}$, one has:

$$\begin{aligned} \int_{\Omega} \mathcal{J}^n(u^n, \phi^n)(x) \psi^n(x) dx &= \sum_{K \in \mathcal{T}^n} \sum_{\sigma \in \mathcal{E}_K} m(\sigma) d_{K,\sigma} \tau_{\sigma}^{\kappa}(u^n) \frac{(D_{\sigma} \phi^n)^2}{d_{\sigma}} \psi(x_K) \\ &= \mathbb{T}_4^n + \mathbb{T}_5^n, \end{aligned} \quad (64)$$

where

$$\mathbb{T}_4^n = \sum_{K \in \mathcal{T}^n} \sum_{\sigma \in \mathcal{E}_K} m(\sigma) d_{K,\sigma} \tau_{\sigma}^{\kappa}(u^n) \frac{(D_{\sigma} \phi^n)^2}{d_{\sigma}} \psi(x_L), \quad (65)$$

and

$$\mathbb{T}_5^n = \sum_{K \in \mathcal{T}^n} \sum_{\sigma \in \mathcal{E}_K} m(\sigma) d_{K,\sigma} \tau_{\sigma}^{\kappa}(u^n) \frac{(D_{\sigma} \phi^n)^2}{d_{\sigma}} (\psi(x_K) - \psi(x_L)), \quad (66)$$

where we have denoted $\psi(x_L) = 0$ if $\sigma \in \mathcal{E}_{\text{ext}} \cap \mathcal{E}_K$. Since $|\psi(x_K) - \psi(x_L)| \leq 2 h_n \|\nabla \psi\|_{(L^{\infty}(\overline{\Omega}))^d}$ and $\tau_{\sigma}^{\kappa}(u^n) \leq \frac{\|\kappa\|_{L^{\infty}(\Omega \times \mathbb{R}, \mathbb{R})}^2}{\alpha d_{\sigma}}$, we have

$$|\mathbb{T}_5^n| \leq 2 \frac{\|\kappa\|_{L^{\infty}(\Omega \times \mathbb{R}, \mathbb{R})}^2}{\alpha} h_n \|\nabla \psi\|_{(L^{\infty}(\overline{\Omega}))^d} \|\phi^n\|_{1,2,X(\mathcal{T})}^2.$$

Using (54) we then obtain that:

$$|\mathbb{T}_5^n| \rightarrow 0, \text{ as } n \rightarrow +\infty. \quad (67)$$

We turn now to the term \mathbb{T}_4^n , reordering the sum on the edges in the right hand side of (65), we get

$$\mathbb{T}_4^n = \sum_{\sigma \in \mathcal{E}} m(\sigma) \tau_{\sigma}^{\kappa}(u^n) (D_{\sigma} \phi^n)^2 \psi_{\sigma}^n,$$

where ψ_{σ}^n is defined by (49). We may then decompose $\mathbb{T}_4^n = \mathbb{T}_6^n + \mathbb{T}_7^n$, with

$$\mathbb{T}_6^n = - \sum_{\sigma \in \mathcal{E}} m(\sigma) \tau_{\sigma}^{\kappa}(u^n) (\phi_L^n - \phi_K^n) (\phi_K^n \psi_K^n - \phi_L^n \psi_L^n), \quad (68)$$

and

$$\mathbb{T}_7^n = - \sum_{\sigma \in \mathcal{E}} m(\sigma) \tau_{\sigma}^{\kappa}(u^n) ((\phi_L^n - \phi_K^n) \phi_K^n (\psi_{\sigma}^n - \psi_K^n) - (\phi_L^n - \phi_K^n) \phi_L^n (\psi_{\sigma}^n - \psi_L^n)). \quad (69)$$

(where we have denoted $\psi_K^n = \psi(x_K)$, for any $K \in \mathcal{T}^n$). We shall show below that

$$\mathbb{T}_6^n \rightarrow \int_{\Omega} \kappa(x, u(x)) |\nabla \phi|^2(x) \psi(x) dx + \int_{\Omega} \kappa(x, u(x)) \nabla \phi(x) \cdot \nabla \psi(x) \phi(x) dx, \text{ as } n \rightarrow +\infty, \quad (70)$$

and that

$$\mathbb{T}_7^n \rightarrow - \int_{\Omega} \kappa(x, u(x)) \nabla \phi(x) \cdot \nabla \psi(x) \phi(x) dx \text{ as } n \rightarrow +\infty, \quad (71)$$

from which it is easy to see that

$$\int_{\Omega} \mathcal{J}^n(u^n, \phi^n)(x) \psi^n(x) dx \rightarrow \int_{\Omega} \kappa(x, u(x)) |\nabla \phi|^2(x) \psi(x) dx,$$

which proves (63). To conclude the proof of the Lemma, there only remains to prove (70) and (71).

Let us first prove (70). Reordering the sum of the right hand side of (68) on the control volumes and using the fact that ϕ^n is the solution of the first equation of the finite volume scheme (17),

we get

$$\begin{aligned}
\mathbb{T}_6^n &= - \sum_{K \in \mathcal{T}_n} \sum_{\sigma \in \mathcal{E}_K} m(\sigma) \tau_{K|L}^\kappa(u^n) (\phi_L^n - \phi_K^n) \phi_K^n \psi(x_K) \\
&= \sum_{K \in \mathcal{T}_n} m(K) f_K(u_K^n) \phi_K^n \psi_K \\
&= \int_{\Omega} f(x, u^n(x)) \phi^n(x) \psi^n(x) dx.
\end{aligned}$$

Now u^n converges to $u \in \cap_{q < \frac{d}{d-1}} W_0^{1,q}(\Omega)$ in $L^p(\Omega)$, for all $p < \frac{d}{d-2}$, as $n \rightarrow \infty$, so that, by the Lebesgue theorem, $f(\cdot, u^n) \rightarrow f(\cdot, u)$ in $L^2(\Omega)$ as $n \rightarrow +\infty$. Moreover, ϕ^n tends to ϕ in $L^2(\Omega)$. Finally, it is clear that $\psi^n \rightarrow \psi$ in $L^\infty(\Omega)$. Hence we get that

$$\mathbb{T}_6^n \rightarrow \int_{\Omega} f(x, u(x)) \phi(x) \psi(x) dx, \quad \text{as } n \rightarrow \infty.$$

Since $\phi\psi \in H_0^1(\Omega)$, one may take it as a test function in the first equation of (6), which gives

$$\begin{aligned}
\int_{\Omega} f(x, u(x)) \phi(x) \psi(x) dx &= \int_{\Omega} \kappa(x, u(x)) \nabla \phi(x) \cdot \nabla (\psi \phi)(x) dx \\
&= \int_{\Omega} \kappa(x, u(x)) |\nabla \phi|^2(x) \psi(x) dx + \int_{\Omega} \kappa(x, u(x)) \nabla \phi(x) \cdot \nabla \psi(x) \phi(x) dx,
\end{aligned}$$

which proves (70). Finally, reordering the sum of \mathbb{T}_7^n on the edges of the control volumes, we get

$$\mathbb{T}_7^n = \sum_{K \in \mathcal{T}_n} \sum_{\sigma \in \mathcal{E}_K} \frac{m(\sigma)}{d_{K,\sigma}} \kappa_K(u_K^n) (\phi_\sigma^n - \phi_K^n) (\psi_K^n - \psi_\sigma^n) \phi_K^n,$$

where ϕ_σ^n is defined by (50). Using Lemma 5.2 with $v = \phi$, we obtain (71), which concludes the proof. \square

For the sake of completeness, we then prove the convergence of the ohmic losses.

Lemma 5.4 (Ohmic losses). *Under the assumptions of Lemma 5.3, let $\mathcal{J}^n(u^n, \phi^n)$ be defined by (21) for $\mathcal{T} = \mathcal{T}_n$, then $\mathcal{J}^n(u^n, \phi^n)$ satisfies (47), that is:*

$$\int_{\Omega} \mathcal{J}^n(u^n, \phi^n)(x) dx \rightarrow \int_{\Omega} \kappa(x, u(x)) |\nabla \phi|^2(x) dx \quad \text{as } n \rightarrow +\infty,$$

and therefore, $\mathcal{J}^n(u^n, \phi^n) \rightarrow \kappa(\cdot, u) |\nabla \phi|^2$ as $n \rightarrow +\infty$ for the weak \star topology of $(\mathcal{C}(\overline{\Omega}))'$, where $\mathcal{C}(\overline{\Omega})$ denotes the set of continuous functions on $\overline{\Omega}$.

Proof. By definition of \mathcal{J}^n , and again noting that $m(\mathcal{D}_{K,\sigma}) = \frac{1}{d} m(\sigma) d_{K,\sigma}$, one has:

$$\begin{aligned}
\int_{\Omega} \mathcal{J}^n(u^n, \phi^n)(x) dx &= \sum_{K \in \mathcal{T}_n} \sum_{\sigma \in \mathcal{E}_K} m(\sigma) d_{K,\sigma} \tau_\sigma^\kappa(u^n) \frac{(D_\sigma \phi^n)^2}{d_\sigma} \\
&= \sum_{K \in \mathcal{T}_n} m(K) f_K(u_K^n) \phi_K^n \\
&= \int_{\Omega} f(x, u^n(x)) \phi^n(x) dx.
\end{aligned}$$

Since u^n converges to $u \in \cap_{q < \frac{d}{d-1}} W_0^{1,q}(\Omega)$ in $L^p(\Omega)$, for all $p < \frac{d}{d-2}$, as $n \rightarrow \infty$, and $\phi_n \rightarrow \phi \in H_0^1(\Omega)$ in $L^2(\Omega)$, we get that

$$\int_{\Omega} \mathcal{J}^n(u^n, \phi^n)(x) dx \rightarrow_{n \rightarrow +\infty} \int_{\Omega} f \phi dx;$$

hence, thanks to the fact that ϕ satisfies (6), $\mathcal{J}^n(u^n, \phi^n)$ satisfies (47).

Now from Lemma 5.3, we get that

$$\int_{\Omega} \mathcal{J}^n(u^n, \phi^n) \psi dx \rightarrow_{n \rightarrow +\infty} \int_{\Omega} \kappa(\cdot, u) |\nabla \phi|^2 \psi dx \text{ for any } \psi \in \mathcal{C}_c^\infty(\Omega).$$

This, together with (47), yields, by classical results in measure theory, that:

$$\int_{\Omega} \mathcal{J}^n(u^n, \phi^n) \psi dx \rightarrow_{n \rightarrow +\infty} \int_{\Omega} \kappa(\cdot, u) |\nabla \phi|^2 \psi dx \text{ for any } \psi \in \mathcal{C}(\bar{\Omega}),$$

which concludes the proof of the lemma. \square

6. CONCLUSION AND PERSPECTIVES

We proved here the convergence of a cell centred finite volume method and the linear finite element method for the coupled heat and potential equation; the condition on the considered meshes is such that the discrete maximum principle holds. Indeed, the technique of proof mimics the tools used for the existence the continuous case, which requires the monotonicity of the operator. In the case of the cell centred finite volume, the scheme satisfies the maximum principle for any admissible mesh. These include triangles and rectangles in two space dimensions, and Voronoï meshes in any dimension.

In two space dimensions, the linear finite element method satisfies the discrete maximum principle for triangular meshes under the Delaunay condition. It is easy to show that under this condition, the matrix of the scheme is identical to that of the cell-centred finite volume on the dual Voronoï mesh. Therefore, the convergence of the finite element scheme may be obtained from that of the finite volume scheme, as explained in [24].

In three space dimensions, there is no known way to build a Voronoï mesh from a tetrahedral one, and therefore one must proceed directly with the finite element interpolation operator, as in section 4 above, and in [9] in the case of a linear diffusion operator. In the three-dimensional case, a known sufficient condition for the maximum principle to hold on a tetrahedral meshes is that all angles of all the faces be strictly acute. Unfortunately, there does not seem to be an easy way to construct such meshes in practise [3, 16], so that our convergence result for the finite element scheme in 3D remains quite academic.

Let us also note that the proof of convergence for the finite element uses the strong convergence in H^1 of the gradient of the approximate solutions (item 1. of Theorem 4.1), which is quite easy to prove. In the case of the cell centred method presented here, we could also have used a discrete gradient that converges strongly, as in [20], but the natural implementation which was performed in [23] leads to a weakly converging gradient as introduced in [19].

Another open problem concerns anisotropic problems. Indeed, if the diffusion coefficients are tensors, no practical sufficient condition is known for the maximum principle to hold, neither for the finite element method, nor for the finite volume one: in fact, finite volume schemes built with a strongly converging gradient exist, either for admissible meshes [20], or for general meshes [1]. However, the stencil of these schemes is wider than the one considered here, and they do not, in general, satisfy the discrete maximum principle. Hence work is required to prove their convergence for an irregular (L^1 or measure) right-hand-side.

REFERENCES

- [1] R. EYMARD, T. GALLOUËT AND R. HERBIN, A new finite volume scheme for anisotropic diffusion problems on general grids: convergence analysis, *submitted*.
- [2] PH. BÉNILAN, L. BOCCARDO, T. GALLOUËT, R. GARIEPY, M. PIERRE AND J.-L. VAZQUEZ, An L^1 theory of existence and uniqueness of solutions of nonlinear elliptic equations, *Ann. Scuola. Norm. Sup. Pisa*, **22**, 241–273, 1995.
- [3] M. BERN AND D. EPPSTEIN, Mesh generation and optimal triangulation, *Computing in Euclidean geometry. Ding-Zhu Du and Frank Hwang (eds.)*, World Scientific, Lecture Notes Series on Computing, **1**, 23–90, 1992.
- [4] L. BOCCARDO AND T. GALLOUËT, Nonlinear elliptic and parabolic equations involving measures Data, *J. of Functional Analysis*, **87** (1), 149–169, 1989.
- [5] L. BOCCARDO AND T. GALLOUËT, Nonlinear elliptic with right hand side measure, *Comm. P.D.E.*, **17**, 641–655, 1992.
- [6] L. BOCCARDO, T. GALLOUËT AND L. ORSINA, Existence and uniqueness of entropy solutions for nonlinear elliptic equations with measure data, *Ann. I.H.P., Analyse non linéaire*, **13** (5), 539–551, 1996.
- [7] L. BOCCARDO, T. GALLOUËT AND J.-L. VAZQUEZ, Nonlinear Elliptic Equations in \mathbb{R}^n without Growth Restrictions on the Data, *Journal of Differential Equations*, **105** (2), 334–363, 1993.
- [8] H. BREZIS, *Analyse Fonctionnelle: Théorie et Applications* (Dunod, Paris), 1999.
- [9] J. CASADO-DIAZ, T. CHACON REBOLLO, V. GIRAULT, M. GOMEZ AND F. MURAT, Finite elements approximation of second order linear elliptic equations in divergence form with right-hand side in L^1 , *Submitted*.
- [10] PH.G. CIARLET, Basic error estimates for elliptic problems, in *Handbook of Numerical Analysis*, Vol. II, Finite Element Methods (Part 1), P. G. Ciarlet and J. L. Lions (eds.), **II** (North-Holand, Amsterdam), 17–352, 1991.
- [11] PH.G. CIARLET AND P.A. RAVIART, Maximum principle and uniform convergence for the finite element method, *Comput. Methods Appl. Mech. Engrg.*, **2**, 17–31, 1973.
- [12] S. CLAIN, Finite element approximations for Laplace operator with right hand-side measure, *Math. Models Meth. App. Sci.*, **6** (5), 713–719, 1996.
- [13] Y. COUDIÈRE, T. GALLOUËT AND R. HERBIN, Discrete Sobolev inequalities and L^p error estimates for approximate finite volume solutions of convection diffusion equations, *Modél. Math. Anal. Numér.*, **35** (4), 767–778, 2001.
- [14] A. DALL’AGLIO, Approximated solutions of equations with L^1 solutions with L^1 data. Application to the H -convergence of quasi-linear parabolic equations, *Ann. Mat. Pura Appl.*, **170**, 207–240, 1996.
- [15] J. DRONIOU, T. GALLOUËT AND R. HERBIN, A finite volumes schemes for a noncoercive elliptic equation with measure data, *SIAM J. Numer. Anal.*, **41** (6), 1997–2031, 2003.
- [16] D. EPPSTEIN, J. M. SULLIVAN AND A. ÜNGÖR, Tiling space and slabs with acute tetrahedra, *Computational Geometry*, **27** (3), 237–255, 2004.
- [17] R. EYMARD, T. GALLOUËT AND R. HERBIN, Finite volume methods, in *Handbook of Numerical Analysis. P. G. Ciarlet and J. L. Lions (eds.)*, **VII** (North-Holand, Amsterdam), 723–1020, 2000.
- [18] R. EYMARD, T. GALLOUËT AND R. HERBIN, Convergence of finite volume schemes for semilinear convection diffusion equations, *Numer. Math.*, **82**, 91–116, 1999.
- [19] R. EYMARD AND T. GALLOUËT, H -convergence and numerical schemes for elliptic problems, *Siam. J. Numer. Anal.*, **41** (2), 539–562, 2003.
- [20] R. EYMARD, T. GALLOUËT AND R. HERBIN, A cell-centred finite volume approximation for anisotropic diffusion operators on unstructured meshes in any space dimension, *IMA Journal of Numerical Analysis*, **26** (2), 326–353, 2006.
- [21] P. FABRIE AND T. GALLOUËT, Modeling wells in porous media flow, *Math. Models Methods App. Sci.*, **10** (5), 673–709, 2000.
- [22] J.R. FERGUSON, J. M. FIARD AND R. HERBIN, A mathematical model of solid oxide fuel cells, *Journal of Power Sources*, **58**, 109–122, 1996.
- [23] J. M. FIARD AND R. HERBIN, Comparison between finite volume and finite element methods for an elliptic system arising in electrochemical engineering, *Comput. Methods App. Mech. Engrg.*, **115**, 315–338, 1994.
- [24] T. GALLOUËT AND R. HERBIN, Convergence of linear finite element for diffusion equations with measure data, *C. R. Acad. Sci. Paris*, **338** (1), 81–84, 2004.
- [25] T. GALLOUËT AND R. HERBIN, Existence of a solution to a coupled elliptic system, *Appl. Math. Lett*, **7** (2), 49–55, 1994.
- [26] T. GALLOUËT AND R. HERBIN, Finite volume approximation of elliptic problems with irregular data, in *Finite Volumes for Complex Applications. F. Benkhaldoun, M. Hänel and R. Vilsmeier (eds.)*, **II** (Hermès, Paris), 155–162, 1999.

- [27] T. GALLOUËT, Measure data and numerical schemes for elliptic problems, *Progress in Nonlinear Differential Equations and their Applications*, **63**, 1–13, 2005.
- [28] R. HERBIN, On the effective resistance of an electrolytic membrane”, *Math. Engng. Ind.*, 4, 4, 311-326, 1993.
- [29] F. MURAT, Equations elliptiques non linéaires avec second membre L^1 ou mesure, in *Actes du 26-ème Congrès d'Analyse Numérique* (Les Karellis, Université de Lyon), 1994.
- [30] R. SCOTT: Finite element convergence for singular data. *Numer. Math.*, **21**, 317–327, 1973.
- [31] G. STAMPACCHIA, Le problème de Dirichlet pour les équations elliptiques du second ordre à coefficients discontinus, *Ann. Inst. Fourier*, **15**, 189–258, 1965.

† WEIERSTRASS INSTITUTE OF APPLIED ANALYSIS AND STOCHASTICS, MOHRENSTR. 39, 10117 BERLIN, GERMANY
E-mail address: `bradji@wias-berlin.de`

*UNIVERSITÉ DE PROVENCE, CENTRE DE MATHÉMATIQUES ET INFORMATIQUE, 39 RUE F. JOLIOT CURIE, 13453
MARSEILLE, FRANCE
E-mail address: `herbin@cmi.univ-mrs.fr`