Asymptotic preserving schemes on conical unstructured 2D meshes

X Blanc, V Delmas, P Hoch

Asymptotic preserving schemes on conical unstructured 2D meshes

X. Blanc[1], V. Delmas[2], P. Hoch[2]

[1] Université de Paris, Laboratoire Jacques-Louis Lions (LJLL), F-75005 Paris, France
Sorbonne-Université, CNRS, LJLL, F-75005 Paris, France

[2]CEA, DAM, DIF,
91297 ARPAJON Cedex, FRANCE.

May 14, 2020

Abstract

In this article, we consider the P1 approximation of the radiative transfer equation. This system is linear hyperbolic and satisfies a diffusion limit. Some scheme have been proposed which reproduce this diffusion limit. Here, we extend such schemes, originally defined on polygonal meshes, to conical meshes.

Contents

1 Introduction

In this article, we consider the following system

$$\begin{cases} \partial_t E + \dfrac{1}{\varepsilon}\operatorname{div}(\boldsymbol{F}) = 0, \\ \partial_t \boldsymbol{F} + \dfrac{1}{\varepsilon}\nabla E + \dfrac{\sigma}{\varepsilon^2}\boldsymbol{F} = 0, \end{cases} \tag{1.1}$$

Here, $\sigma = \sigma(t, \boldsymbol{x}) > 0$ is the cross-section, $\varepsilon > 0$ is a fixed (small) parameter. The scalar function $E = E(t, \boldsymbol{x})$ is the radiative energy, while the vector-valued function $\boldsymbol{F} = \boldsymbol{F}(t, x)$ is the radiative flux. In all the paper, the spatial variable is denoted by $\boldsymbol{x} \in \mathbb{R}^2$, and the time variable by $t > 0$. System (1.1) is completed by boundary conditions. Here, we restrict our attention either to periodic boundary conditions, or homogeneous Dirichlet boundary conditions.

The model (1.1) is an approximation of the radiative transfer equation [7, 8, 19, 20], in which the closure is made by assuming that the radiative intensity is an affine function of the velocity.

It is well-known that, in the limit $\varepsilon \to 0$, this system reduces to the diffusion equation

$$\partial_t E - \operatorname{div}\left(\frac{1}{\sigma}\nabla E\right) = 0. \tag{1.2}$$

This is easily seen by a Hilbert expansion of E and $\boldsymbol{F}$ in powers of ε. It is important, from a physical point of view, to use a numerical scheme which is consistent with this limit. Such a scheme is usually called asymptotic preserving (AP). Many such schemes exist in the literature. The first work we know of in this direction is [14]. In this work, the aim was to design a scheme which is well-balanced (WB), that is, which exactly preserves stationary solutions of the system. However, doing so, the scheme happens to be AP as well. Further works in this direction may be found in [3, 4, 5, 10, 12, 13, 16]. All these works deal with (possibly deformed) polygonal meshes. The aim of the present article is to generalize such AP schemes to conical meshes.

Conical meshes have already been used for instance in [1] for discretizing the compressible Euler equations of hydrodynamics. However, in such a case, the question of diffusion limit is irrelevant, and therefore has not been studied in [1]. Here, we choose the approach used in [3, 12], which is actually an extension of [16] to dimension 2, using nodal flux evaluations. Before doing so, we first extend the diffusion scheme of [11, Section 2.2.2] (see also [3, Section 4.2]) to conical meshes, using cell-centered finite volumes with nodal flux evaluations. It happens that, doing so, we find a scheme which, when particularized to a polygonal mesh, does not give the standard one presented in [11]. In particular, this new scheme does not suffer from the cross-stencil instability mentioned in [11].

The article is organized as follows: in Section 2, we define what we mean by conical meshes, introducing some notation. In Section 3, we present the generalization of diffusion schemes with nodal fluxes to conical meshes. The scheme is analyzed, and some numerical results are provided. Section 4 is devoted to the definition of AP scheme on conical meshes, using the method developed in [11, 3], and adapting it to conical meshes. Two versions are proposed, depending on how the source term in the second line of (1.1) is discretized. In both case, numerical results are presented.

2 Conical meshes

2.1 Definition and basic properties

In all the paper, we consider a domain $\Omega \subset \mathbb{R}^2$ which is Lipschitz and bounded. We assume that a mesh is given on Ω. By mesh, we mean the following:

1. $(\Omega_j)_{1\leq j\leq N}$ is a finite family of nonempty connected open disjoint subsets of Ω (the cells) such that $\overline{\Omega} = \bigcup_{j=1}^{N} \overline{\Omega}_j$;
2. $(\boldsymbol{M}_r)_{1\leq r\leq P}$ is a family of P distinct points (the nodes) of Ω such that, for all r, $\boldsymbol{M}_r \in \partial\Omega_j$, for some j;

3. for all j, the boundary $\partial\Omega_j$ of Ω_j is the union of a finite number of edges, each of which is a smooth one-dimensional curve joining two nodes (say, $\boldsymbol{M}_r$ and $\boldsymbol{M}_{r+1}$) of Ω_j. Moreover, two different edges share at most one node. Finally, any edge is contained either in $\partial\Omega$ or in $\overline{\Omega}_j \cap \overline{\Omega}_k$ for two distinct cells Ω_j and Ω_k.

The edges are not necessarily segments, so we describe below how we parametrize them.

Let $\mathbf{M}_0$ and $\mathbf{M}_2$ be two points linked by a regular arc $\boldsymbol{\gamma}^{arc}_{\mathbf{M}_0,\mathbf{M}_2}$. It is further assumed that this arc can be represented by a closed part zero-level curve of a function Γ:

$$\begin{array}{rccc} \Gamma: & \mathbb{R}^2 & \longrightarrow & \mathbb{R} \\ & (x,y) & \longmapsto & \Gamma(x,y) \end{array}$$

such that $\boldsymbol{\gamma}^{arc}_{\mathbf{M}_0,\mathbf{M}_2}$ admits an implicit representation:

$$\boldsymbol{\gamma}^{arc}_{\mathbf{M}_0,\mathbf{M}_2} \subset \{(x,y) \in \mathbb{R}^2; \Gamma(x,y) = 0\}.$$

Definition 2.1. *If the implicit function Γ is a bi-variate polynomial of (maximal) degree p, we say that the arc $\boldsymbol{\gamma}^{arc}_{\mathbf{M}_0,\mathbf{M}_2}$ is of order $p+1$. In particular:*

$$\text{Linear form } P^1(x,y) \ : \ a\,x + b\,y + c = 0, \qquad \textit{straight lines.} \tag{2.1}$$

$$\text{Quadratic form } P^2(x,y) \ : \ a\,x^2 + b\,y^2 + 2\,c\,x\,y + d\,x + e\,y + f = 0, \qquad \textbf{conics}. \tag{2.2}$$

We recall that solutions of the quadratic equation (2.2) generate the family of conics of the plane.

In the context of finite volume methods for which fluxes are computed on edges, instead of considering only straight edges (2.1), we are interested in conical curves (2.2).

Proposition 2.2 (Quadratic Rational Bezier curve)**.** *Conical arcs may be parameterized by rational Bezier curve of degree 2, charaterized by 2 endpoints $\mathbf{M}_0$ and $\mathbf{M}_2$, a control point $\mathbf{M}_1$, and a scalar weight ω (cf Figure 1):*

$$\mathbf{M}^\omega(q) = \begin{pmatrix} x(q) \\ y(q) \end{pmatrix} = \frac{\mathbf{M}_0(1-q)^2 + 2\omega q(1-q)\mathbf{M}_1 + q^2\mathbf{M}_2}{(1-q)^2 + 2\omega q(1-q) + q^2}, \qquad q \in [0,1]. \tag{2.3}$$

Hence, we have: $\boldsymbol{\gamma}^{arc}_{\mathbf{M}_0,\mathbf{M}_2} = \{\mathbf{M}^\omega(q), q \in [0,1]\}$.

Each conical edge depends on its two endpoints $\mathbf{M}_0$, $\mathbf{M}_2$ and 3 (in dimension 2) additional parameters: the control point $\mathbf{M}_1$ and a scalar weight ω. The link with conics comes from the following characterization:

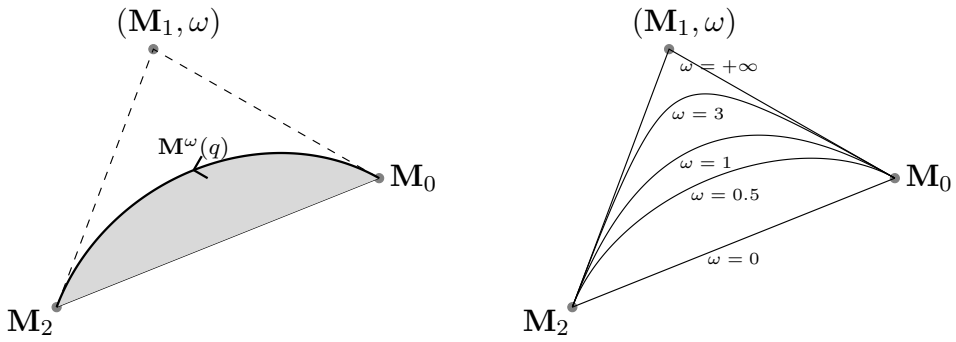

Figure 1: Conic parametrized by quadratic rational Bezier curves.

Proposition 2.3. *The edge is said to be of type:*

- *planar degenerate if* $\omega = 0$,
- *elliptic if* $\omega \in]0,1[$,

- *parabolic if* $\omega = 1$,
- *hyperbolic if* $\omega \in]1, +\infty[$.

In addition, we introduce the *shoulder point* $\mathbf{S}$ located at the midpoint of arc parametrization, $\mathbf{S} := \mathbf{M}^{\omega}(q = 0.5)$. It may be noted that the unit normal at this point is orthogonal to $[\mathbf{M}_0, \mathbf{M}_2]$

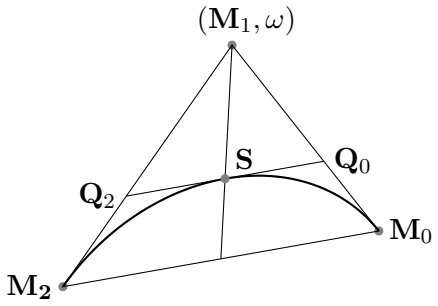

Figure 2: Shoulder point $\mathbf{S}$.

The shoulder point $\mathbf{S}$ satisfies the following property:

$$\mathbf{S} = \frac{1}{2}(\mathbf{Q_0} + \mathbf{Q_2}), \qquad \text{with } \mathbf{Q_0} = \frac{1}{1+\omega}(\omega\mathbf{M_1} + \mathbf{M_0}),\ \mathbf{Q_2} = \frac{1}{1+\omega}(\omega\mathbf{M_1} + \mathbf{M_2}).$$

2.2 Relationship on geometric objects

We recall that the area of a cell is given by:

$$|\Omega_j| = \int_{\Omega_j} 1dv = \frac{1}{2}\int_{\partial\Omega_j} \mathbf{OM}(s)\cdot\hat{\mathbf{N}}(s)ds \tag{2.4}$$

$$= \sum_e \frac{1}{2}\int_0^1 \mathbf{OM}^{\omega,e}(q)\cdot\mathbf{N}_e(q)dq \tag{2.5}$$

where $\hat{\mathbf{N}}(s)$ is the external **unit** normal of cell Ω_j whose boundary is parametrized by s and $\mathbf{N}_e(q)$ is the **non unit** one ($\mathbf{N}_e(q) = \hat{\mathbf{N}}_e(q)\frac{ds}{dq}$). The point $\mathbf{O}$ is the origin, but it can be any fixed point in the plane. In practice, it is usually set as the center of the cell Ω_j.

$$|\Omega_j| = \sum_e (A(\mathbf{O}, \mathbf{M}_0^e, \mathbf{M}_2^e) + A(\mathbf{M}^{\omega,e}(q), \mathbf{M}_0^e, \mathbf{M}_1^e, \mathbf{M}_2^e)), \tag{2.6}$$

where $A(\mathbf{O}, \mathbf{M}_0^e, \mathbf{M}_2^e)$ is the area of a simplex defined by origin $\mathbf{O}$, $\mathbf{M}_0^e$ and $\mathbf{M}_2^e$ vertices of edge e, and $A(\mathbf{M}^{\omega,e}(q), \mathbf{M}_0^e, \mathbf{M}_1^e, \mathbf{M}_2^e))$ corresponds to the area between straight edge $\mathbf{M}_0^e\mathbf{M}_2^e$ and the curved arc defined by (2.3). In addition,

$$A(\mathbf{M}^{\omega,e}(q), \mathbf{M}_0^e, \mathbf{M}_1^e, \mathbf{M}_2^e) = \frac{1}{2}\int_0^1 \left(x(q)\frac{dy}{dq}(q) - \frac{dx}{dq}(q)y(q)\right)dq. \tag{2.7}$$

Moreover:

$$A(\mathbf{M}^{\omega,e}(q), \mathbf{M}_0^e, \mathbf{M}_1^e, \mathbf{M}_2^e) = f(\omega)A(\mathbf{M}_0^e, \mathbf{M}_1^e, \mathbf{M}_2^e), \tag{2.8}$$

with $f(\omega)$ a piecewise analytical function of ω. The exact formula is [15], [2], [1]:

$$\begin{cases} (1)\ \text{if } \omega = 0,\ \text{then } f(\omega) = 0, \\ (2)\ \text{if } 0 < \omega < 1,\ \text{then } f(\omega) = \frac{2\omega}{1-\omega^2}\left(\frac{1}{\sqrt{1-\omega^2}}\text{atan}(\sqrt{\frac{1-\omega}{1+\omega}}) - \frac{\omega}{2}\right), \\ (3)\ \text{if } \omega = 1,\ \text{then } f(\omega) = \frac{2}{3}, \\ (4)\ \text{if } \omega > 1,\ \text{then } f(\omega) = \frac{\omega}{\omega^2-1}\left(\omega - \frac{1}{\sqrt{\omega^2-1}}\log(\omega + \sqrt{\omega^2-1})\right). \end{cases} \tag{2.9}$$

We recall that area may be expressed from vectors, cf [2] and Figure 3

$$|\Omega_j| = \frac{1}{2}\sum_{dof} \mathbf{C}_j^{dof,\omega}.\mathbf{OM}_{dof}. \tag{2.10}$$

Here, *dof* stands for degrees of freedom. Hence, the sum runs over all degrees of freedom in the cell (that is, values at the nodes and values at the control points of the edges).

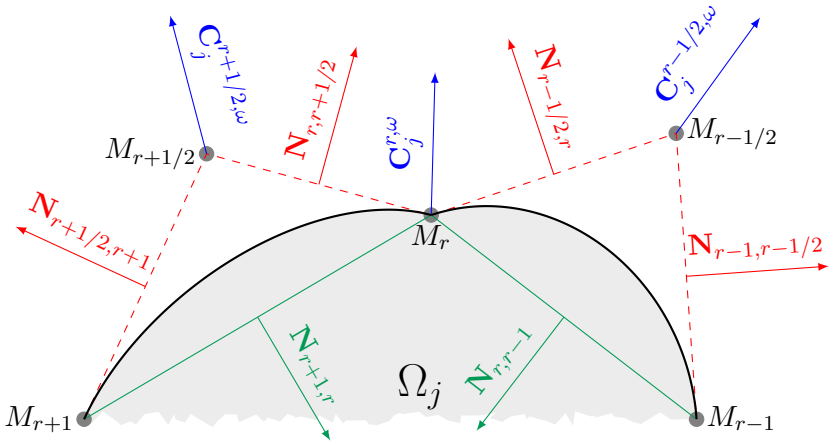

Figure 3: Normal nodal vector expressed with control point

Despite the concise writing of (2.10), a disadvantage is that the control point $\mathbf{M}_{dof}$ does not lie on the arc. This is why we will use a description of the arc with its shoulder point $\mathbf{S}$, cf Figure 4 (at the end), as in [1].

To summarize, (2.8) rewrites:

$$A(\mathbf{M}^{\omega,e}(q), \mathbf{M}_0^e, \mathbf{M}_1^e, \mathbf{M}_2^e) = h(\omega)A(\mathbf{M}_0^e, \mathbf{S}^e, \mathbf{M}_2^e), \tag{2.11}$$

This change in the spatial location of the degree of freedom involves now the function

$$h(\omega) := f(\omega)\frac{\omega+1}{\omega} \tag{2.12}$$

which is strictly decreasing on $\mathbb{R}^+$ with values in $]1, \frac{\pi}{2}]$.

This allows us to derive the following formulas, involving only the points that belong to the edges of cell Ω_j:

1. GLACE type:

$$|\Omega_j| = \frac{1}{2}\sum_{dof} \tilde{\mathbf{C}}_{\mathbf{j}}^{\mathbf{dof},\omega}.\mathbf{OM}_{dof}^{\omega}(q), \tag{2.13}$$

that is,

$$|\Omega_j| = \frac{1}{2}\left(\sum_r \tilde{\mathbf{C}}_j^{r,\omega}.\mathbf{OM}_r + \sum_{r+1/2} \tilde{\mathbf{C}}_j^{r+1/2,\omega}.\mathbf{OS}_{r+1/2}^{\omega}\right) := \frac{1}{2}\sum_{\text{newdof}} \tilde{\mathbf{C}}_j^{\text{newdof},\omega}.\mathbf{OM}_{\text{newdof}} \tag{2.14}$$

with

$$\tilde{\mathbf{C}}_j^{r,\omega} = \tfrac{1}{2}\left((1-h(\omega_{r-1/2}))\tilde{\mathbf{N}}_{r-1,r} + (1-h(\omega_{r+1/2}))\tilde{\mathbf{N}}_{r,r+1} + h(\omega_{r-1/2})\tilde{\mathbf{N}}_{r-1/2,r} + h(\omega_{r+1/2})\tilde{\mathbf{N}}_{r,r+1/2}\right), \tag{2.15}$$

$$\tilde{\mathbf{C}}_j^{r+1/2,\omega} = \tfrac{h(\omega_{r+1/2})}{2}(\tilde{\mathbf{N}}_{r,r+1/2} + \tilde{\mathbf{N}}_{r+1/2,r+1}). \tag{2.16}$$

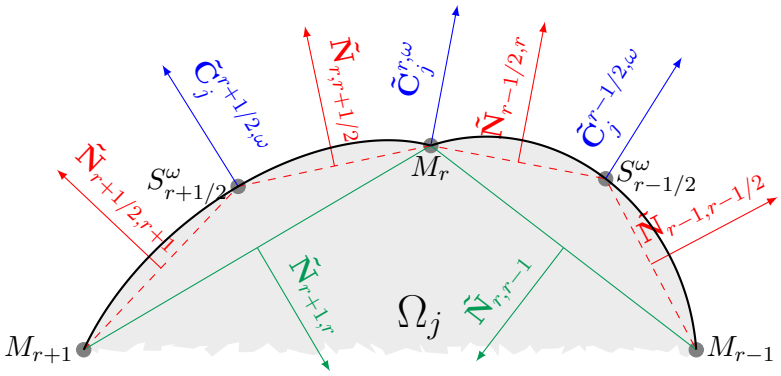

Figure 4: Normal vector from a degree of freedom defined on boundary cell Ω_j. Two types: endpoints $\mathbf{M}_r$ denoted by $\tilde{\mathbf{C}}_j^{r,\omega}$ or shoulder point $\mathbf{S}^{\omega}_{r+1/2}$ denoted by $\tilde{\mathbf{C}}_j^{r+1/2,\omega}$

Unknowns are located at the same points as $\tilde{\mathbf{C}}_j^{dof,\omega}$: points $\mathbf{M}_r, \mathbf{S}^{\omega}_{r+1/2}$ inside cell, moreover they satisfy [2][1]:

$$\text{For cell } \Omega_j : \qquad \sum_{dof} \tilde{\mathbf{C}}_j^{dof,\omega} = \mathbf{0} \tag{2.17}$$

$$\text{For all degree of freedom (dof) inside the domain (Figure 6) : } \sum_{j} \tilde{\mathbf{C}}_j^{dof,\omega} = \mathbf{0} \tag{2.18}$$

2. EUCCLHYD type:

$$\begin{cases} \tilde{\mathbf{N}}_j^{r,\omega-} := \frac{1}{2}(1-h(\omega_{r-1/2}))\tilde{\mathbf{N}}_{r-1,r} + \frac{1}{2}h(\omega_{r-1/2})\tilde{\mathbf{N}}_{r-1/2,r}, \\ \tilde{\mathbf{N}}_j^{r,\omega+} := \frac{1}{2}(1-h(\omega_{r+1/2}))\tilde{\mathbf{N}}_{r,r+1} + \frac{1}{2}h(\omega_{r+1/2})\tilde{\mathbf{N}}_{r,r+1/2}, \\ \tilde{\mathbf{N}}_j^{r+1/2,\omega-} := \frac{1}{2}h(\omega_{r+1/2})\tilde{\mathbf{N}}_{r,r+1/2}, \\ \tilde{\mathbf{N}}_j^{r+1/2,\omega+} := \frac{1}{2}h(\omega_{r+1/2})\tilde{\mathbf{N}}_{r+1/2,r+1}. \end{cases} \tag{2.19}$$

We recall that the unknowns are located at the same points as $\tilde{\mathbf{N}}_j^{dof,\omega\pm}$: points $\mathbf{M}_r, \mathbf{S}^{\omega}_{r+1/2}$ of the cell, morevover they satisfy [2][1]:

$$\text{For all cells } \Omega_j : \qquad \sum_{dof} \tilde{\mathbf{N}}_j^{dof,\omega,+} + \tilde{\mathbf{N}}_j^{dof,\omega,-} = \mathbf{0} \tag{2.20}$$

$$\text{For all dof inside the domain (Figure 6) : } \sum_{j} \tilde{\mathbf{N}}_j^{dof,\omega,+} + \tilde{\mathbf{N}}_j^{dof,\omega,-} = \mathbf{0} \tag{2.21}$$

3. Relatonship between "GLACE type vectors" (2.15) (2.16) cf Figure (4) and "EUCCLHYD type vectors" defined by (2.19) cf Figure (5):

$$\begin{cases} \tilde{\mathbf{C}}_j^{r,\omega} = (\tilde{\mathbf{N}}_j^{r,\omega-} + \tilde{\mathbf{N}}_j^{r,\omega+}), \\ \tilde{\mathbf{C}}_j^{r+1/2,\omega} = (\tilde{\mathbf{N}}_j^{r+1/2,\omega-} + \tilde{\mathbf{N}}_j^{r+1/2,\omega+}). \end{cases} \tag{2.22}$$

and also:

$$\tilde{\mathbf{C}}_j^{dof,\omega} = (\tilde{\mathbf{N}}_j^{dof,\omega-} + \tilde{\mathbf{N}}_j^{dof,\omega+}) \tag{2.23}$$

Remark 2.4. *Property for planar degenerate conic case.*
By a direct computation on (2.12), we have the following:

$$\lim_{\omega\to 0} h(\omega) > 0 \tag{2.24}$$

As we will see thereafter, for planar degenerate conics, the property (2.24) enable us to design new finite volume schemes (when compared to original polygonal ones) for straight meshes.

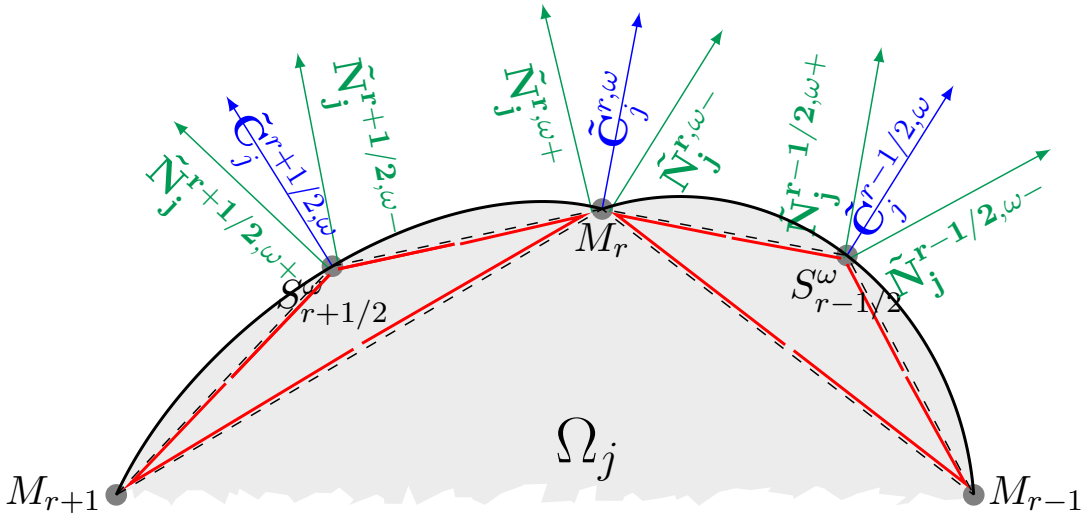

Figure 5: Vector decomposition at nodes (of GLACE type vector) $\tilde{\mathbf{C}}_j^{dof,\omega}$ into two half-vectors to left $\tilde{\mathbf{N}}_j^{dof,\omega+}$ and to the right $\tilde{\mathbf{N}}_j^{dof,\omega-}$ for the scheme of EUCCLHYD type (cf (2.22))

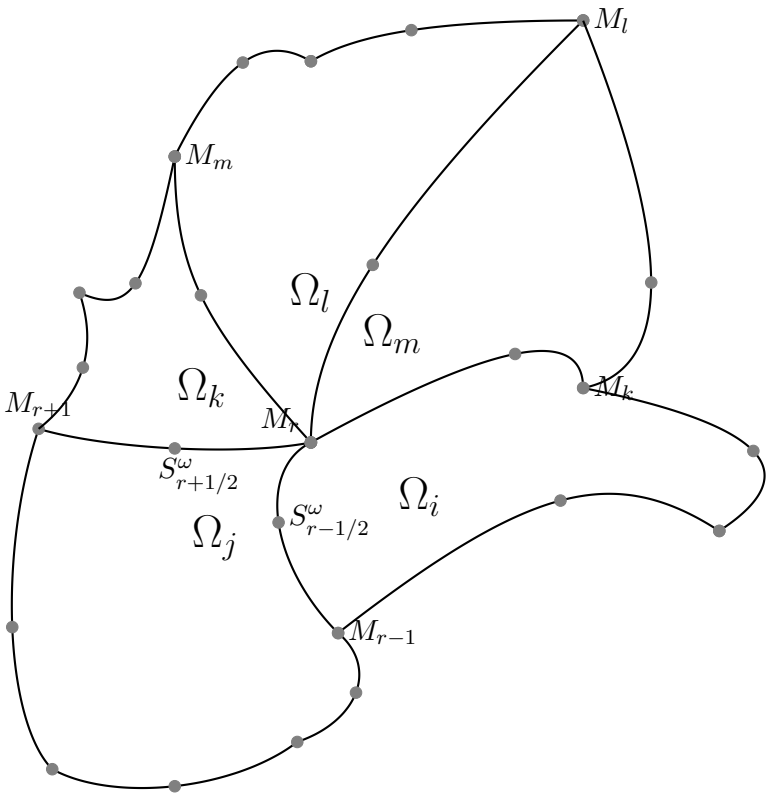

Figure 6: A degree of freedom of the edge of cell Ω_j: the points $\mathbf{M}_r$ and $\mathbf{S}^{\omega}_{r+1/2}$ where the Riemann problems are solved (exact or approximate) and the set of cells close to Ω_j containing this degree of freedom. For example, from the point of view of cell Ω_j, at the point M_r (resp. at the point $\mathbf{S}^{\omega}_{r+1/2}$) we need to define a numerical flux $G_j^r(f_{\Omega_m}, f_{\Omega_l}, f_{\Omega_k}, f_{\Omega_j}, f_{\Omega_i})$ (resp. a flux $G_j^{r+1/2}(f_{\Omega_j}, f_{\Omega_k})$).

3 Diffusion schemes on conical meshes

In this Section, we propose a discretization of the diffusion equation (1.2), using the formalism of [6, 9, 17] (see also [11, Chapter 6]).

3.1 GLACE scheme on polygonal meshes

Let us first recall the GLACE diffusion scheme on polygonal mesh. Using a finite volume approach, we integrate (1.2) on a cell Ω_j, and, denoting by $E_j(t)$ the average value of $E(t,\boldsymbol{x})$ on Ω_j, we have, after integrating by parts,

$$\partial_t E_j - \frac{1}{|\Omega_j|}\int_{\partial\Omega_j}\frac{1}{\sigma}\left(\nabla E(t,\boldsymbol{x}),n\right)=0.$$

In order to define the second term, we denote by $-\frac{1}{\sigma}\nabla E$ the flux, and approximate it using GLACE formalism, that is, nodal flux evaluations:

$$\int_{\partial\Omega_j}\frac{1}{\sigma}\left(\nabla E(t,\boldsymbol{x}),n\right)=\int_{\partial\Omega_j}\left(F(t,\boldsymbol{x}),n\right)\approx\sum_{r/\boldsymbol{M}_r\in\Omega_j}\left(\boldsymbol{F}_r(t),\boldsymbol{C}_j^r\right).$$

Here, the sum runs over all r such that the node $\boldsymbol{M}_r$ belongs to Ω_j, and

$$\boldsymbol{C}_j^r=\nabla_{\boldsymbol{M}_r}|\Omega_j|. \tag{3.1}$$

This definition coincides with (2.13)-(2.14)-(2.15)-(2.16) if one reduces the degrees of freedom to the nodes $\boldsymbol{M}_r$. In order to define the nodal flux $\boldsymbol{F}_r$, we impose the following equality:

$$\sigma(\boldsymbol{M}_r)\sum_{i/\boldsymbol{M}_r\in\Omega_i}\boldsymbol{C}_i^r\otimes(\boldsymbol{M}_r-\boldsymbol{x}_i)\boldsymbol{F}_r=-\sum_{i/\boldsymbol{M}_r\in\Omega_i}\boldsymbol{C}_i^r\left(\nabla E(\boldsymbol{M}_r),\boldsymbol{M}_r-\boldsymbol{x}_i\right),$$

where the sum runs over all i such that $\boldsymbol{M}_r\in\partial\Omega_i$. Now, we use a finite difference approximation, which states that $(\nabla E(\boldsymbol{M}_r),\boldsymbol{M}_r-\boldsymbol{x}_i)\approx E(\boldsymbol{M}_r)-E(\boldsymbol{x}_i)$. Hence, using (2.18), we see that the terms involving $E(\boldsymbol{M}_r)$ vanish, so that we finally have

$$\sigma(\boldsymbol{M}_r)\sum_{i/\boldsymbol{M}_r\in\Omega_i}\boldsymbol{C}_i^r\otimes(\boldsymbol{M}_r-\boldsymbol{x}_i)\boldsymbol{F}_r=\sum_{i/\boldsymbol{M}_r\in\Omega_i}\boldsymbol{C}_i^r E(\boldsymbol{x}_i), \tag{3.2}$$

We end up with the following semi-discrete scheme:

$$\partial_t E_j+\frac{1}{|\Omega_j|}\sum_{r/\boldsymbol{M}_r\in\Omega_i}(\boldsymbol{F}_r,\boldsymbol{C}_i^r)=0,$$

where the nodal fluxes $\boldsymbol{F}_r$ are defined by (3.2). Thus, as explained in [11], the discretization of (1.2) using GLACE formalism leads to the following scheme:

$$\begin{cases} E_i^{n+1}-E_i^n+\dfrac{\Delta t}{|\Omega_i|}\displaystyle\sum_{r/\boldsymbol{M}_r\in\Omega_i}\left(\boldsymbol{F}_r^{n+1},\boldsymbol{C}_i^r\right)=0, & \text{(3.3a)}\\[2ex] \sigma_r\displaystyle\sum_{i/\boldsymbol{M}_r\in\Omega_i}\hat{\beta}_{ir}\boldsymbol{F}_r^{n+1}=\sum_{i/\boldsymbol{M}_r\in\Omega_i}E_i^{n+1}\boldsymbol{C}_i^r, & \text{(3.3b)}\end{cases}$$

with $\hat{\beta}_{ir}=\boldsymbol{C}_i^r\otimes(\boldsymbol{M}_r-\boldsymbol{x}_i)$, where $\boldsymbol{x}_j$ is the center of cell Ω_j. Here, Δt is the time step, and n is the time index. We use an implicit scheme in order to avoid a parabolic stability condition which proves to be too restrictive for a fine mesh. The coefficient $\boldsymbol{C}_i^r$ is defined by (3.1).

As we pointed out above, the definition (3.1) coincides is the definition on conical meshes. However, if one uses the formalism (2.13)-(2.14)-(2.15)-(2.16) with $\omega=0$ (so that edges are straight lines), we still have degrees of freedom at the shoulder points. Hence the two schemes do not coincide, as we will see below.

Remark 3.1. *For the scheme* (3.3a)-(3.3b) *to be well-defined, we need the matrix* $\hat{\beta}_r=\sum_k\hat{\beta}_{kr}$ *to be invertible. This question was studied in [11], where it is proved that* $\hat{\beta}_r$ *is invertible under some regularity conditions on the mesh. For instance, a sufficient condition is that the angle between two neighboring edges should belong to the interval* $[\theta_{opt};\pi-2\theta_{opt}]$ *where* $\theta_{opt}\approx 0.189rad$.

Let us define $\boldsymbol{Q}_k^r = \hat{\beta}_r^{-1}\boldsymbol{C}_k^r$. We thus get the flux (3.3b),

$$\boldsymbol{F}_r^{n+1} = \frac{1}{\sigma_r} \sum_{k/\boldsymbol{M}_r \in \Omega_k} \boldsymbol{Q}_k^r E_k^{n+1}.$$

We can thus write the scheme (3.3a)-(3.3b) as follows

$$E_i^{n+1} - E_i^n + \frac{\Delta t}{|\Omega_i|} \sum_{r/\boldsymbol{M}_r \in \Omega_i} \frac{1}{\sigma_r} \sum_{k/\boldsymbol{M}_r \in \Omega_k} \left(\boldsymbol{C}_i^r, \boldsymbol{Q}_k^r\right) E_k^{n+1} = 0. \tag{3.4}$$

Setting $\boldsymbol{E}^n = \begin{pmatrix} \dots \\ E_{i-1}^n \\ E_i^n \\ E_{i+1}^n \\ \dots \end{pmatrix}$, we can write (3.4) $\hat{M}\boldsymbol{E}^{n+1} = \boldsymbol{E}^n$, where the coefficient M_{ij} of matrix $\hat{M}$ reads

$$M_{ij} = \delta_{ij} + \frac{\Delta t}{|\Omega_i|} \sum_{r/\boldsymbol{M}_r \in \Omega_i \cap \Omega_j} \frac{1}{\sigma_r} \left(\boldsymbol{C}_i^r, \boldsymbol{Q}_j^r\right). \tag{3.5}$$

3.2 GLACE scheme on conical meshes

We now go on and generalize the scheme to conical meshes. In order to do so, we need to introduce additional degrees of freedom, compared to (3.3a)-(3.3b). One can either use degrees of freedom at the control points of edges, or at the shoulder points. Here, we choose to use the shoulder points, because a control point is not on its edge. Hence, defining the unkwown at such a point by interpolation might be a difficult task. We will use the modified vectors $\tilde{\boldsymbol{C}}_j^r$ (2.15), as shown in Figure 3 and 4. We also introduce vectors $\tilde{\boldsymbol{C}}_j^{r+1/2}$ at the shoulder points. These vectors are defined by (2.16).

In this context, the natural extension of (3.3a)-(3.3b) is the following:

$$\begin{cases} E_i^{n+1} - E_i^n + \dfrac{\Delta t}{|\Omega_i|} \left(\displaystyle\sum_{r/\boldsymbol{M}_r \in \Omega_i} \left(\tilde{\boldsymbol{C}}_i^r, \boldsymbol{F}_r^{n+1}\right) + \sum_{r/\boldsymbol{M}_{r+1/2} \in \Omega_i} \left(\tilde{\boldsymbol{C}}_i^{r+1/2}, \boldsymbol{F}_{r+1/2}^{n+1}\right) \right) = 0 & \text{(3.6a)} \\ \sigma_r \displaystyle\sum_{i/\boldsymbol{M}_r \in \Omega_i} \hat{\beta}_{ir} \boldsymbol{F}_r^{n+1} = \sum_{i/\boldsymbol{M}_r \in \Omega_k} \tilde{\boldsymbol{C}}_i^r E_i^{n+1} & \text{(3.6b)} \\ \sigma_{r+1/2} \displaystyle\sum_{i/\boldsymbol{M}_{r+1/2} \in \Omega_i} \hat{\beta}_{ir+1/2} \boldsymbol{F}_{r+1/2}^{n+1} = \sum_{i/\boldsymbol{M}_{r+1/2} \in \Omega_i} \tilde{\boldsymbol{C}}_i^{r+1/2} E_i^{n+1} & \text{(3.6c)} \end{cases}$$

where $\hat{\beta}_{ir} = \tilde{\boldsymbol{C}}_i^r \otimes (\boldsymbol{M}_r - \boldsymbol{x}_i)$ and $\hat{\beta}_{ir+1/2} = \tilde{\boldsymbol{C}}_i^{r+1/2} \otimes (\boldsymbol{M}_{r+1/2} - \boldsymbol{x}_i)$. We define the matrices $\hat{\beta}_r$ and $\hat{\beta}_{r+1/2}$ by

$$\hat{\beta}_r = \sum_{k/\boldsymbol{M}_r \in \Omega_k} \hat{\beta}_{kr}, \quad \hat{\beta}_{r+1/2} = \sum_{k/\boldsymbol{M}_{r+1/2} \in \Omega_k} \hat{\beta}_{kr+1/2}. \tag{3.7}$$

As in the polygonal case, the matrix $\hat{\beta}_r$ is invertible, at least under some regularity conditions on the mesh. On the contrary, the matrix $\hat{\beta}_{r+1/2}$ is always singular, as the following result shows:

Lemma 3.2. *The matrix $\hat{\beta}_{r+1/2}$ defined by* (3.7) *has rank* 1.

Proof. Given a shoulder point $\boldsymbol{M}_{r+1/2}$, we have only two cells Ω_i and Ω_j which share the corresponding edge. Applying (2.18), we thus have

$$\tilde{\boldsymbol{C}}_i^{r+1/2} + \tilde{\boldsymbol{C}}_j^{r+1/2} = 0.$$

Hence, using (3.7),

$$\hat{\beta}_{r+1/2} = \tilde{\boldsymbol{C}}_i^{r+1/2} \otimes (\boldsymbol{x}_j - \boldsymbol{x}_i),$$

which has rank one. □

In the following, we will use the following definition, for any shoulder point $\boldsymbol{M}_{r+1/2}$:

$$\text{If } \boldsymbol{M}_{r+1/2} \in \Omega_i \cap \Omega_j, \quad \boldsymbol{N}_{r+1/2} = \boldsymbol{x}_j - \boldsymbol{x}_i \quad \text{and} \quad \boldsymbol{T}_{r+1/2} = \left(\boldsymbol{N}_{r+1/2}\right)^{\perp}. \tag{3.8}$$

Computing the scalar product of (3.6c) with $\boldsymbol{N}_{r+1/2}$ and using (2.18) again, we have

$$\left(\boldsymbol{F}_{r+1/2}^{n+1}, \boldsymbol{N}_{r+1/2}\right) = \frac{E_i^{n+1} - E_j^{n+1}}{\sigma_{r+1/2}}. \tag{3.9}$$

According to Lemma 3.2, we are free to choose any value for the component of the flux $\boldsymbol{F}_{r+1/2}^{n+1}$ along $\boldsymbol{T}_{r+1/2}$. Here, we choose to use the average between the fluxes at nodes r and $r+1$. Of course, some other possibilities may be considered (see for instance [1]), as far as it remains consistent. We thus define

$$\left(\boldsymbol{F}_{r+1/2}^{n+1}, \boldsymbol{T}_{r+1/2}\right) = \frac{1}{2}\left(\boldsymbol{F}_{r}^{n+1} + \boldsymbol{F}_{r+1}^{n+1}, \boldsymbol{T}_{r+1/2}\right) \tag{3.10}$$

Next, using the vector $\boldsymbol{Q}_j^r = \hat{\beta}_r^{-1}\boldsymbol{C}_j^r$ as in the polygonal case we have,

$$\left(\boldsymbol{F}_{r+1/2}^{n+1}, \boldsymbol{T}_{r+1/2}\right) = \sum_{k/\boldsymbol{M}_{r+1/2}\in\Omega_k} \left(\boldsymbol{Q}_k^r, \boldsymbol{T}_{r+1/2}\right) \frac{E_k^n}{2\sigma_{r+1/2}} + \sum_{p/\boldsymbol{M}_{r+1/2}\in\Omega_p} \left(\boldsymbol{Q}_p^{r+1}, \boldsymbol{T}_{r+1/2}\right) \frac{E_p^n}{2\sigma_{r+1/2}}, \tag{3.11}$$

where the sum over k runs for all k such that $\boldsymbol{M}_r \in \Omega_k$, and the sum over p for all p such that $\boldsymbol{M}_{r+1} \in \Omega_p$. Combining (3.9) and (3.11) we are able to compute the flux at the shoulder point $r+1/2$:

$$\begin{aligned}\boldsymbol{F}_{r+1/2}^{n+1} = {} & \frac{E_i^{n+1} - E_j^{n+1}}{\sigma_{r+1/2}} \frac{\boldsymbol{N}_{r+1/2}}{\left\|\boldsymbol{N}_{r+1/2}\right\|^2} \\ & + \left(\sum_{k/\boldsymbol{M}_{r+1/2}\in\Omega_k} E_k^{n+1}\left(\boldsymbol{Q}_k^r, \boldsymbol{T}_{r+1/2}\right) + \sum_{p/\boldsymbol{M}_{r+1/2}\in\Omega_p} E_p^{n+1}\left(\boldsymbol{Q}_p^{r+1}, \boldsymbol{T}_{r+1/2}\right) \right) \frac{\boldsymbol{T}_{r+1/2}}{2\sigma\left\|\boldsymbol{T}_{r+1/2}\right\|^2}\end{aligned} \tag{3.12}$$

This allows to put the system under ther form $\hat{M}\boldsymbol{E^{n+1}} = \boldsymbol{E^n}$, where the matrix $\hat{M}$ is defined by inserting (3.12) and (3.6b) into (3.6a), that is,

$$\begin{aligned}\hat{M}_{ij} = {} & \delta_{ij}\left(1 + \frac{\Delta t}{|\Omega_i|} \sum_{r/\boldsymbol{M}_{r+1/2}\in\Omega_i} \frac{\left(\tilde{\boldsymbol{C}}_i^r, \boldsymbol{N}_{r+1/2}\right)}{\sigma_{r+1/2}\left\|\boldsymbol{N}_{r+1/2}\right\|^2}\right) \\ & + \frac{\Delta t}{|\Omega_i|} \sum_{r/\boldsymbol{M}_r\in\Omega_i\cap\Omega_j} \frac{\left(\tilde{\boldsymbol{C}}_i^r, Q_j^r\right)}{\sigma_r} \\ & - \frac{\Delta t}{|\Omega_i|} \sum_{r/\boldsymbol{M}_{r+1/2}\in\Omega_i\cap\Omega_j} \frac{\left(\tilde{\boldsymbol{C}}_i^r, \boldsymbol{N}_{r+1/2}\right)}{\sigma_{r+1/2}\left\|\boldsymbol{N}_{r+1/2}\right\|^2} \\ & + \frac{\Delta t}{|\Omega_i|} \sum_{r/\boldsymbol{M}_{r+1/2}\in\Omega_i\cap\Omega_j} \frac{\left(\tilde{\boldsymbol{C}}_i^{r+1/2}, \boldsymbol{T}_{r+1/2}\right)\left(Q_j^r, \boldsymbol{T}_{r+1/2}\right)}{2\sigma_{r+1/2}\left\|\boldsymbol{T}_{r+1/2}\right\|^2} \\ & + \frac{\Delta t}{|\Omega_i|} \sum_{r/\boldsymbol{M}_{r+1/2}\in\Omega_i\cap\Omega_j} \frac{\left(\tilde{\boldsymbol{C}}_i^{r+1/2}, \boldsymbol{T}_{r+1/2}\right)\left(Q_j^{r+1}, \boldsymbol{T}_{r+1/2}\right)}{2\sigma_{r+1/2}\left\|\boldsymbol{T}_{r+1/2}\right\|^2}\end{aligned} \tag{3.13}$$

As it was pointed out in [11], on a cartesian mesh, the scheme (3.3a)-(3.3b) has a stencil which induces coupling of two cells Ω_j and Ω_i only if they share a node. This is not the case in the scheme (3.6a)-(3.6b), as it is clear in the definition of the matrix (3.13). Indeed, additional terms are present, which may (and actually do) couple cells which have a common edge. This may be seen on the case of a small mesh (9 cells) displayed in Figure 7. On the left part, there is a circle if the corresponding coefficient of the matrix is larger than 10^{-14}. For the conical case, we add star under the same condition.

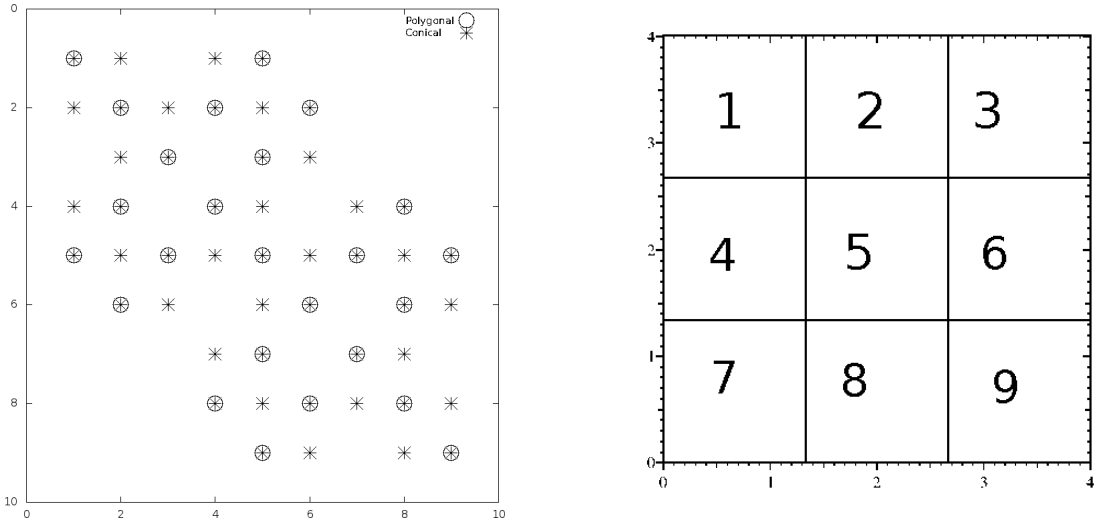

Figure 7: The stencil of the scheme in the case $\omega \geq 0$. Circles ($\circ$) correspond to the polygonal scheme, and crosses ($*$) to conical scheme.

In the pure polygonal case, we have a cross-stencil, as explained in [11]. In the conical case, additional terms are present. For instance, cell 5 is only coupled to cells $1, 3, 7, 9$ in the polygonal case, whereas it is coupled to all cells in the conical case.

3.3 Numerical tests

We present some numerical tests on various meshes. Our aim is first to show that using the conical mesh formalism has a stabilizing effect on the GLACE diffusion scheme. We also provide a convergence analysis on a case with an analytic solution.

3.3.1 Cartesian mesh

We use a cartesian mesh on the domain $\Omega = [-2; 2]^2$, with 31 cells in each direction. We then make the edges conical, adding control points at a distance d in the normal direction of a given edge, with $d = 0.2\Delta x$, Δx being the mesh size (see Figure 8). We impose that the control point is always closer to the origin than the center of the cell. Other configurations have been tested (control points on the other side of the edge, or random positions) with similar results.

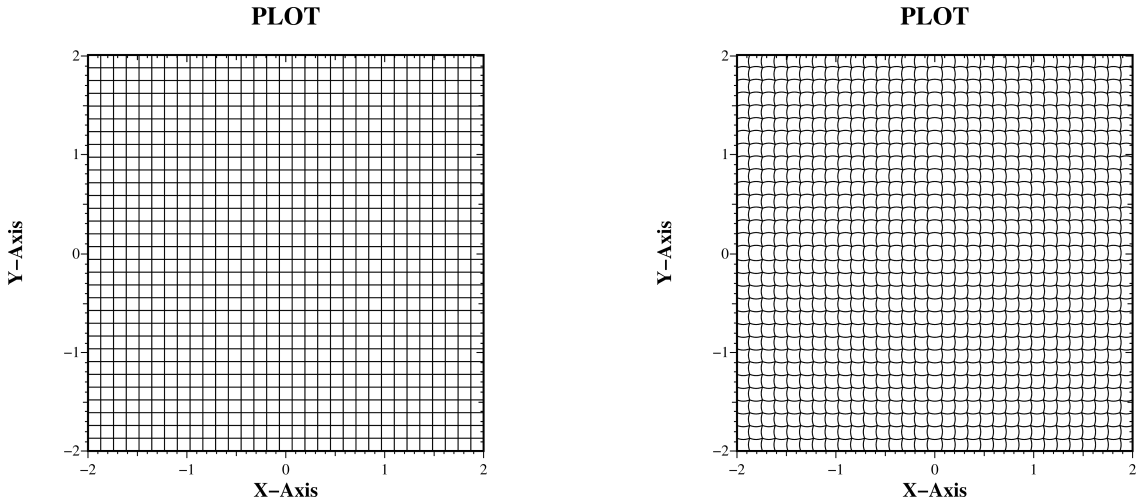

Figure 8: Left: classical cartesian mesh. Right: conical mesh with $\omega = 1$

We choose as an initial condition a Dirac mass, that is,

$$\begin{cases} E_j^0 = \frac{1}{|\Omega_j|} \text{ if } \boldsymbol{x}_j \in [1.995; 2.005]^2 \\ E_j^0 = 0 \text{ else} \end{cases}$$

In other words, energy is present only in the central cell, and we have a total energy equal to 1. We set $\sigma = 1$. The time step is $\Delta t = 0.003$s, and we give the results at $t = 0.03$s for different meshes.

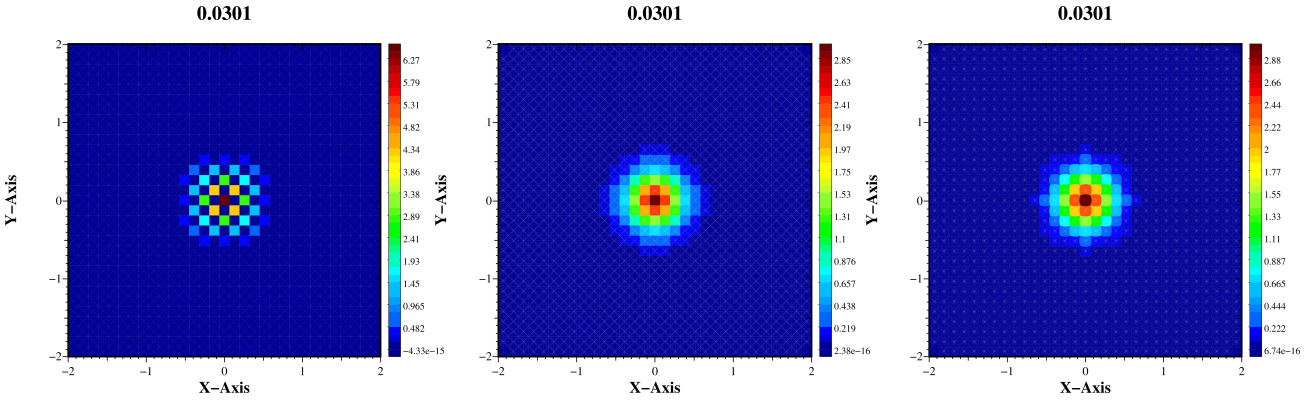

Figure 9: Left: classical cartesian mesh with GLACE polygonal scheme. Right: conical mesh with $\omega = 1$

Results are displayed on Figure 9. In the case of the cartesian mesh, we observe the cross stencil: one in two cells is not reached by the energy diffusion. This problem is no more present in the case of a conical mesh.

Next, we compute the numerical error of the scheme in this case. The equation reads

$$\begin{cases} E(t,x,y) - \dfrac{1}{\sigma}\Delta E(t,x,y) = 0, \\ E(0,x,y) = \delta_{(x=0,y=0)}. \end{cases}$$

The solution is given by

$$E_{\text{exact}}(t,x,y) = \frac{\sigma}{4\pi t} e^{-\frac{\sigma}{4t}(x^2+y^2)}, \quad \forall t > 0. \tag{3.14}$$

with

$$\int_0^{+\infty}\int_0^{+\infty} E_{\text{exact}}(t,x,y)dxdy = 1, \quad \forall t > 0. \tag{3.15}$$

The initialization is set at $t = 10^{-4}s$ in order to avoid the singularity of the Dirac mass. We choose $\sigma = 1$, $\Delta t = 10^{-2}$ and a final time equal to $t = 0.1$. Then we compute the L^2 error between the exact solution and the numerical solution. For a given mesh size Δx, we use several meshes (see Figure 10): the cartesian one (red curve), the conical one with control points outwards from the origin and $\omega = 0$ (green curve), $\omega = 2$ (blue curve), and the conical one with control point inwards from the origin and $\omega = 2$ (purple curve).

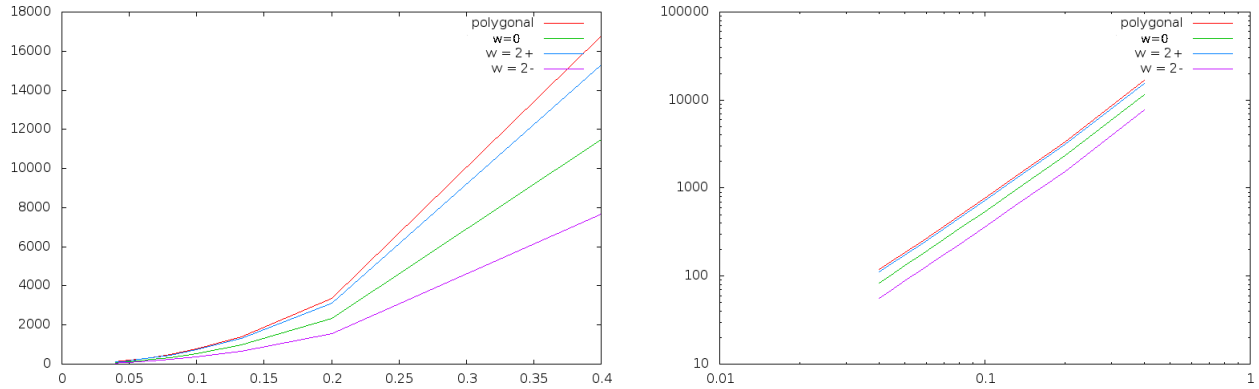

Figure 10: Left: error as a function of Δx. Right: log of the error as a function of Δx.

In all cases, we observe a convergence rate of approximately 2.14.

3.3.2 Radial mesh

Here we use a radial polygonal mesh on the one side, and a conical on the other side (Figure 11). For the conical mesh, we set $\omega = \frac{\sqrt{2}}{2}$ for orthoradial edges, and $\omega = 0$ for radial edges. The control points are defined in order to have edges which are portions of circles centered at the origin. We expect to have a much better solution with the conical mesh, if the exact solution is radially symmetric.

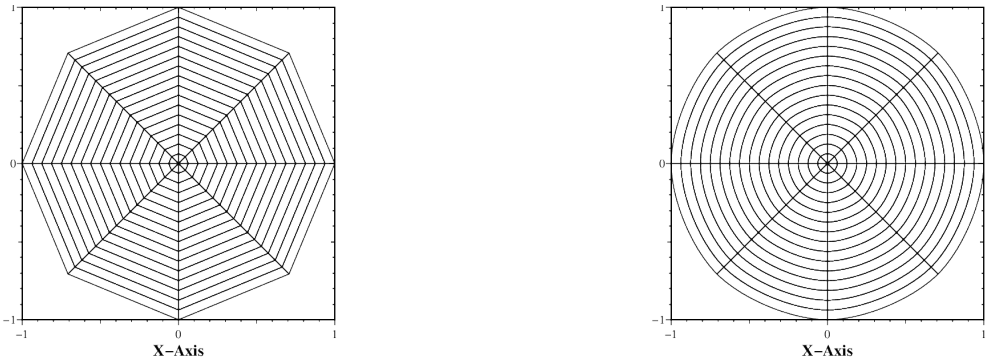

Figure 11: Left: polygonal mesh. Right: conical mesh with $\omega = \frac{\sqrt{2}}{2}$ for orthoradial edges, $\omega = 0$ for radial edges.

Here again, we use the exact solution (3.14), with an initial time equal to $t = 10^{-4}s$, and $\sigma = 1$.

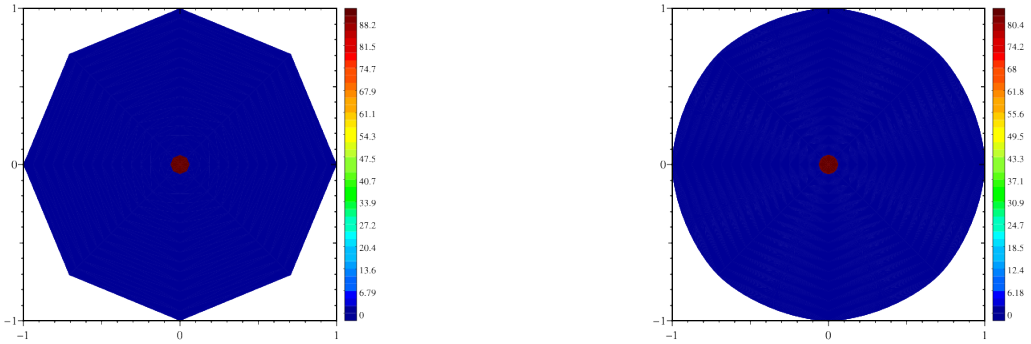

Figure 12: Left: polygonal mesh. Right: conical mesh with $\omega = \frac{\sqrt{2}}{2}$, $t = 0.0001s$

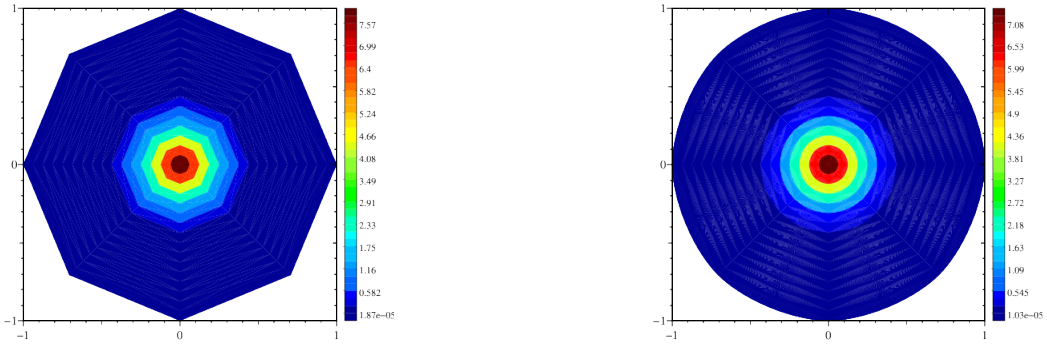

Figure 13: Left: polygonal mesh. Right: conical mesh with $\omega = \frac{\sqrt{2}}{2}$, $t = 0.0101s$

As expected, the solution is more accurately computed in the conical case. In particular, radial symmetry is preserved exactly (Figure 12 and Figure 13).

Next, we present a test case in which the polygonal scheme presents positivity issues. Here, we initialize the energy by 0 everywhere, except in one cell, so that the integral of E is equal to 1.

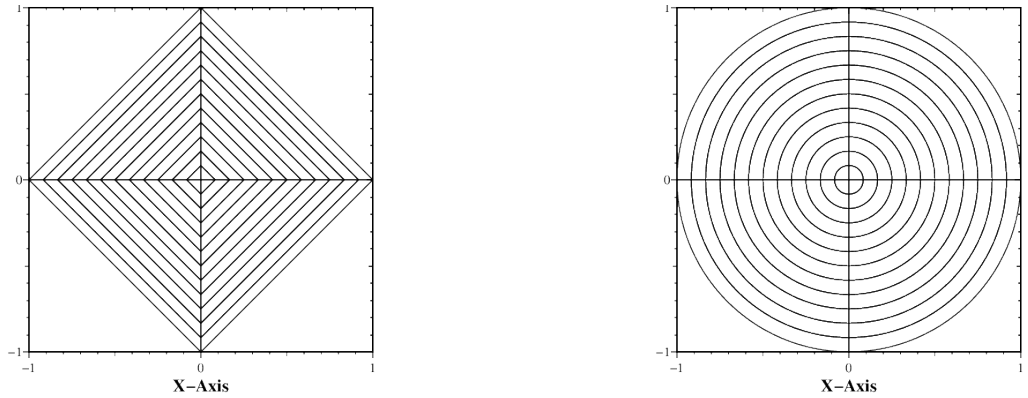

Figure 14: Left: polygonal mesh. Right: conical mesh with $\omega = \frac{\sqrt{2}}{2}$ for orthoradial edges, $\omega = 0$ for radial edges.

We present two conical meshes: the first one with $\omega = \frac{\sqrt{2}}{2}$, that is, with "real" conical edges (Figure 14, right). The second one satisfies $\omega = 0$, hence it is "almost" polygonal (Figure 14, left). This allows to observe the difference between the scheme designed for polygonal meshes (3.3a)-(3.3b), and the one designed for conical meshes (3.6a)-(3.6b)-(3.6c), but used on a polygonal mesh.

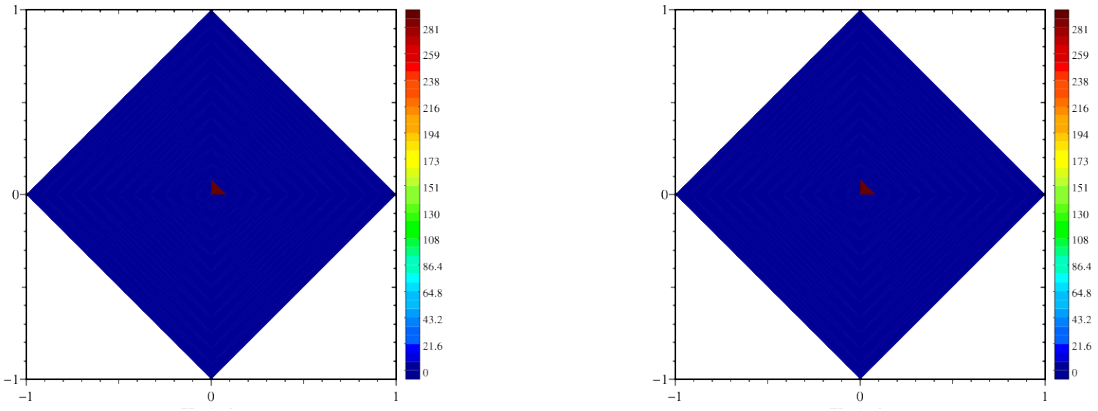

Figure 15: Left: polygonal mesh. Right: conical mesh with $\omega = 0$, $t = 0s$

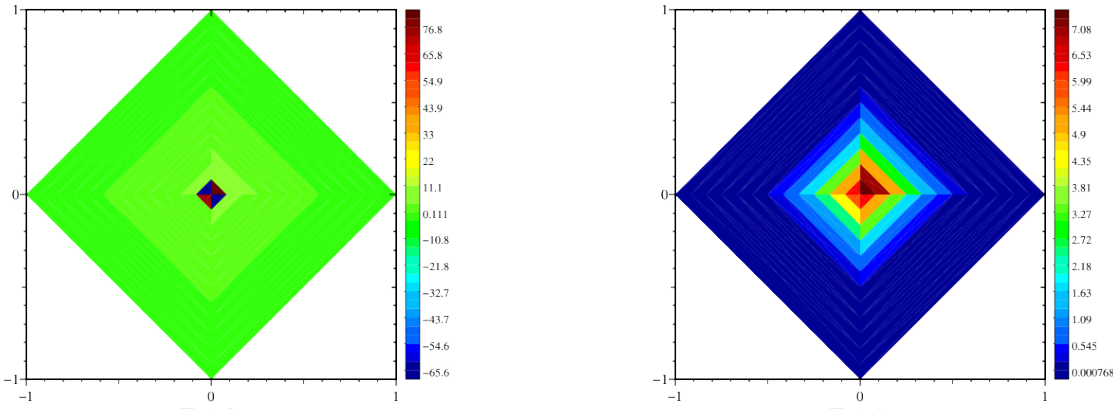

Figure 16: Left: polygonal mesh. Right: conical mesh with $\omega = 0$, $t = 0.01s$

3.3.3 Triangular mesh

In this Subsection, we use a regular triangular mesh of the domain$\Omega = [-1; 1]^2$, which consists of 6700 triangles. We add control points along the normal of each edge, at distance equal to 20% of the edge length (Figure 17).

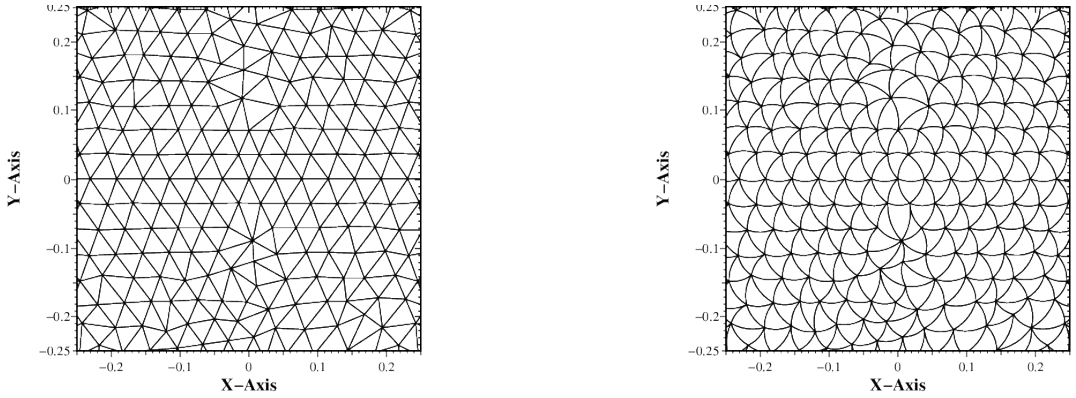

Figure 17: Left: triangular mesh. Right: conical mesh with $\omega = 1$.

The initial data is equal to 0 everywhere, except in the central hexagon, so that the integral of the energy is equal to 1. We set $\Delta t = 0.002$, $\sigma = 1$ and the final time is $t = 0.01$s.

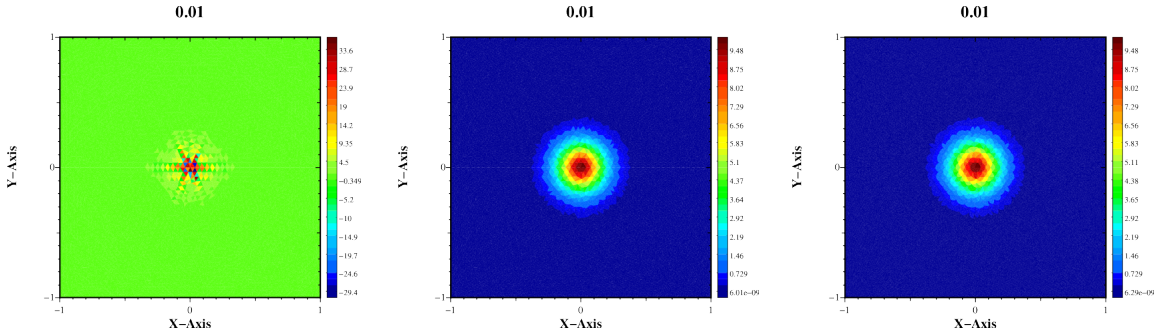

Figure 18: Left: polygonal mesh. Center: conical mesh with $\omega = 0$. Right: conical mesh with $\omega = 1$.

We note that the solution is much better on the conical mesh (Figure 18). The exact solution remains positive, while the numerical one computed on the polygonal meshes have large negative values.

4 AP schemes for the hyperbolic heat equation

In this section we define AP schemes for system (1.1) on conical meshes. Two variants are studied, in the spirit of [11], depending on how the source term $\frac{\sigma}{\varepsilon^2}\boldsymbol{F}$ is descretized. The first one (JLa) uses a cell-centered descretization, while the second one (JLb) uses a nodal discretization. It is shown in [11] that in dimension one, the first discretization corresponds to Jin-Levermore scheme [16], while the second one leads to Gosse-Toscani scheme [13].

4.1 JLa scheme

In this part we first recall the definition of GLACE AP schemes for system (1.1) on polygonal meshes. Then we give an adaptation on conical meshes.

4.1.1 JLa scheme for polygonal meshes

The JLa scheme on polygonal meshes reads [11]:

$$\begin{cases} E_i^{n+1} - E_i^n + \dfrac{\Delta t}{\varepsilon|\Omega_i|} \displaystyle\sum_{r/\boldsymbol{M}_r\in\Omega_i} \left(\boldsymbol{C}_i^r, \boldsymbol{F}_r^{n+1}\right) = 0, & \text{(4.1a)} \\ \boldsymbol{F}_i^{n+1} - \boldsymbol{F}_i^n + \dfrac{\Delta t}{\varepsilon|\Omega_i|} \displaystyle\sum_{r/\boldsymbol{M}_r\in\Omega_i} \boldsymbol{C}_i^r E_{ir}^{n+1} = -\dfrac{\sigma_i \Delta t}{\varepsilon^2} \boldsymbol{F}_i^{n+1}, & \text{(4.1b)} \end{cases}$$

where the fluxes are defined by

$$\begin{cases} E_{ir}^{n+1} = E_i^{n+1} + \left(\boldsymbol{F}_i^{n+1} - \boldsymbol{F}_r^{n+1}, \boldsymbol{n}_{ir}\right) - \dfrac{\sigma_r}{\varepsilon}\left(\boldsymbol{F}_r^{n+1}, \boldsymbol{M}_r - \boldsymbol{x}_i\right), & \text{(4.2a)} \\ \displaystyle\sum_{i/\boldsymbol{M}_r\in\Omega_i} \left(\hat{\alpha}_{ir} + \frac{\sigma_r}{\varepsilon}\hat{\beta}_{ir}\right) \boldsymbol{F}_r^{n+1} = \sum_{i/\boldsymbol{M}_r\in\Omega_i} \boldsymbol{C}_i^r E_i^{n+1} + \sum_{i/\boldsymbol{M}_r\in\Omega_i} \boldsymbol{C}_i^r \otimes \boldsymbol{n}_{ir} \boldsymbol{F}_i^{n+1}. & \text{(4.2b)} \end{cases}$$

where the local matrices read $\hat{\alpha}_{ir} = \boldsymbol{C}_i^r \otimes \boldsymbol{n}_{ir}$ and $\hat{\beta}_{ir} = \boldsymbol{C}_i^r \otimes (\boldsymbol{M}_r - \boldsymbol{x}_i)$ with $\boldsymbol{n}_{ir} = \frac{\boldsymbol{C}_i^r}{\|\boldsymbol{C}_i^r\|}$.

We also define

$$\hat{\alpha}_r = \sum_{i/\boldsymbol{M}_r\in\Omega_i} \hat{\alpha}_{ir} \quad \text{and} \quad \hat{\beta}_r = \sum_{i/\boldsymbol{M}_r\in\Omega_i} \hat{\beta}_{ir}. \tag{4.3}$$

Note that we have used an implicit version of the scheme, in order to avoid any stability issue. However, an explicit version is easily derived in the same way.

Remark 4.1. *The matrix $\hat{\alpha}_r$ is symmetric positive definite on any non-degenerate mesh. This fact is proved in [17]). The matrix $\hat{\beta}_r$ is invertible under the same conditions as in the diffusion case (see Remark 3.1).*

Let us now define the matrices $\hat{\gamma}_{ir} = \hat{\alpha}_{ir} + \frac{\sigma_r}{\varepsilon}\hat{\beta}_{ir}$ and $\hat{\gamma}_r = \hat{\alpha}_r + \frac{\sigma_r}{\varepsilon}\hat{\beta}_r$, so that the fluxes read

$$\begin{cases} E_{ir}^{n+1} = E_i^{n+1} + \left(\boldsymbol{F}_i^{n+1} - \boldsymbol{F}_r^{n+1}, \boldsymbol{n}_{ir}\right) - \dfrac{\sigma_r}{\varepsilon}\left(\boldsymbol{F}_r^{n+1}, \boldsymbol{M}_r - \boldsymbol{x}_i\right), & \text{(4.4a)} \\ \hat{\gamma}_r \boldsymbol{F}_r^{n+1} = \displaystyle\sum_{k/\boldsymbol{M}_r\in\Omega_k} \boldsymbol{C}_k^r E_k^{n+1} + \sum_{k/\boldsymbol{M}_r\in\Omega_k} \hat{\alpha}_{kr} \boldsymbol{F}_k^{n+1}. & \text{(4.4b)} \end{cases}$$

Thus, assuming that $\hat{\gamma}_r$ is invertible, we write (4.1a) as follows:

$$E_i^{n+1} - E_i^n + \frac{\Delta t}{\varepsilon|\Omega_i|} \sum_{r/\boldsymbol{M}_r\in\Omega_i} \sum_{k/\boldsymbol{M}_r\in\Omega_k} \left[\left(\boldsymbol{C}_i^r, \hat{\gamma}_r^{-1}\boldsymbol{C}_k^r\right) E_k^{n+1} + \left(\boldsymbol{C}_i^r, \hat{\gamma}_r^{-1}\hat{\alpha}_{kr}\boldsymbol{F}_k^{n+1}\right)\right] = 0. \tag{4.5}$$

In order to compute the nodal values E_{ir}^{n+1}, we insert (4.4b) into (4.4a):

$$E_{ir}^{n+1} = E_i^{n+1} + \left(\boldsymbol{F}_i^{n+1}, \boldsymbol{n}_{ir}\right) - \sum_{k/\boldsymbol{M}_r\in\Omega_k} \left(\hat{\gamma}_r^{-1}\boldsymbol{C}_k^r E_k^{n+1} + \hat{\gamma}_r^{-1}\hat{\alpha}_{kr}\boldsymbol{F}_k^{n+1}, \mathcal{R}\right) \tag{4.6}$$

with $\mathcal{R} = \boldsymbol{n}_{ir} + \frac{\sigma_r}{\varepsilon}(\boldsymbol{M}_r - \boldsymbol{x}_i)$.

Next, we insert (4.6) into (4.1b), which gives the following

$$\left(1+\frac{\sigma_i\Delta t}{\varepsilon^2}\right)\boldsymbol{F}_i^{n+1}+\frac{\Delta t}{\varepsilon|\Omega_i|}\sum_r\left[E_i^{n+1}+\left(\boldsymbol{F}_i^{n+1},\boldsymbol{n}_{ir}\right)-\sum_k\left(\hat{\gamma}_r^{-1}\boldsymbol{C}_k^rE_k^{n+1}+\hat{\gamma}_r^{-1}\hat{\alpha}_{kr}\boldsymbol{F}_k^{n+1},\mathcal{R}\right)\right]\boldsymbol{C}_i^r=\boldsymbol{F}_i^n. \quad (4.7)$$

Gathering (4.5) and (4.7), we get

$$\begin{cases} E_i^{n+1}-E_i^n+\dfrac{\Delta t}{\varepsilon|\Omega_i|}\displaystyle\sum_r\sum_k\left[\left(\boldsymbol{C}_i^r,\boldsymbol{Q}_k^r\right)E_k^{n+1}+\left(\boldsymbol{C}_i^r,\hat{\gamma}_r^{-1}\hat{\alpha}_{kr}\boldsymbol{F}_k^{n+1}\right)\right]=0, & (4.8a)\\ \left(1+\dfrac{\sigma_i\Delta t}{\varepsilon^2}\right)\boldsymbol{F}_i^{n+1}+\dfrac{\Delta t}{\varepsilon|\Omega_i|}\displaystyle\sum_r\left[E_i^{n+1}+\left(\boldsymbol{F}_i^{n+1},\boldsymbol{n}_{ir}\right)-\sum_k\left(\hat{\gamma}_r^{-1}\boldsymbol{C}_k^rE_k^{n+1}+\hat{\gamma}_r^{-1}\hat{\alpha}_{kr}\boldsymbol{F}_k^{n+1},\mathcal{R}\right)\right]\boldsymbol{C}_i^r=\boldsymbol{F}_i^n. & (4.8b)\end{cases}$$

with $\boldsymbol{Q}_k^r=\hat{\gamma}_r^{-1}\boldsymbol{C}_k^r$. In (4.11b), (4.11c) and (4.11d), the sums over r run for r such that $\boldsymbol{M}_r\in\Omega_i$, and the sums over k run over k such that $\boldsymbol{M}_r\in\Omega_k$.

Next, we write the scheme under the matrix form

$$\hat{M}^{\text{poly}}\boldsymbol{X^{n+1}}=\boldsymbol{X^n}\quad\text{with }\boldsymbol{X}^n=\begin{pmatrix}E^n\\F_x^n\\F_y^n\end{pmatrix}.$$

Each block of the vector $\boldsymbol{X}^n$ is of size N, the number of cells. Let us give the expression of the terms of the matrix $\hat{M}^{\text{poly}}$. We define

$$\mathcal{X}_k^r=\hat{\gamma}_r^{-1}\hat{\alpha}_{kr},$$

and we have

$$\begin{cases} E_i^{n+1}+\dfrac{\Delta t}{\varepsilon|\Omega_i|}\displaystyle\sum_r\sum_k\left[\left(\boldsymbol{C}_i^r,\boldsymbol{Q}_k^r\right)E_k^{n+1}+\left(\boldsymbol{C}_i^r,\mathcal{X}_k^r\boldsymbol{F}_k^{n+1}\right)\right]=E_i^n,\\ \left(1+\dfrac{\sigma_i\Delta t}{\varepsilon^2}\right)\boldsymbol{F}_i^{n+1}+\dfrac{\Delta t}{\varepsilon|\Omega_i|}\displaystyle\sum_r\left[E_i^{n+1}+\left(\boldsymbol{F}_i^{n+1},\boldsymbol{n}_{ir}\right)-\sum_k\left(\hat{\gamma}_r^{-1}\boldsymbol{C}_k^r,\mathcal{R}\right)E_k^{n+1}-\sum_k\left(\mathcal{X}_k^r\boldsymbol{F}_k^{n+1},\mathcal{R}\right)\right]\boldsymbol{C}_i^r=\boldsymbol{F}_i^n.\end{cases}$$

Thus, the coefficients of the matrix $\hat{M}^{\text{poly}}$ read, for all i,j such that $1\le i,j\le N$,

$$\begin{aligned}
\hat{M}_{ij}^{\text{poly}} &= \delta_{ij}+\frac{\Delta t}{\varepsilon|\Omega_i|}\sum_{r/\boldsymbol{M}_r\in\Omega_i\cap\Omega_j}\left(\boldsymbol{C}_i^r,\boldsymbol{Q}_j^r\right),\\
\hat{M}_{i,j+N}^{\text{poly}} &= \frac{\Delta t}{\varepsilon|\Omega_i|}\sum_{r/\boldsymbol{M}_r\in\Omega_i\cap\Omega_j}\left(\boldsymbol{C}_i^r,\mathcal{X}_j^r|_1\right),\\
\hat{M}_{i,j+2N}^{\text{poly}} &= \frac{\Delta t}{\varepsilon|\Omega_i|}\sum_{r/\boldsymbol{M}_r\in\Omega_i\cap\Omega_j}\left(\boldsymbol{C}_i^r,\mathcal{X}_j^r|_2\right),\\
\hat{M}_{i+N,j}^{\text{poly}} &= \delta_{ij}\frac{\Delta t}{\varepsilon|\Omega_i|}\sum_{r/\boldsymbol{M}_r\in\Omega_i}\boldsymbol{C}_i^r(1)-\frac{\Delta t}{\varepsilon|\Omega_i|}\sum_{r/\boldsymbol{M}_r\in\Omega_i\cap\Omega_j}\left(\hat{\gamma}_r^{-1}\boldsymbol{C}_j^r,\mathcal{R}\right)\boldsymbol{C}_i^r(1),\\
\hat{M}_{i+N,j+N}^{\text{poly}} &= \delta_{ij}\left(1+\frac{\sigma_i\Delta t}{\varepsilon^2}\right)+\delta_{ij}\frac{\Delta t}{\varepsilon|\Omega_i|}\sum_{r/\boldsymbol{M}_r\in\Omega_i}\boldsymbol{n}_{ir}(1)\boldsymbol{C}_i^r(1)-\frac{\Delta t}{\varepsilon|\Omega_i|}\sum_{r/\boldsymbol{M}_r\in\Omega_i\cap\Omega_j}\left(\mathcal{X}_j^r|_1,\mathcal{R}\right)\boldsymbol{C}_i^r(1),\\
\hat{M}_{i+N,j+2N}^{\text{poly}} &= \delta_{ij}\frac{\Delta t}{\varepsilon|\Omega_i|}\sum_{r/\boldsymbol{M}_r\in\Omega_i}\boldsymbol{n}_{ir}(2)\boldsymbol{C}_i^r(1)-\frac{\Delta t}{\varepsilon|\Omega_i|}\sum_{r/\boldsymbol{M}_r\in\Omega_i\cap\Omega_j}\left(\mathcal{X}_j^r|_2,\mathcal{R}\right)\boldsymbol{C}_i^r(1),\\
\hat{M}_{i+2N,j}^{\text{poly}} &= \delta_{ij}\frac{\Delta t}{\varepsilon|\Omega_i|}\sum_{r/\boldsymbol{M}_r\in\Omega_i}\boldsymbol{C}_i^r(2)-\frac{\Delta t}{\varepsilon|\Omega_i|}\sum_{r/\boldsymbol{M}_r\in\Omega_i\cap\Omega_j}\left(\hat{\gamma}_r^{-1}\boldsymbol{C}_j^r,\mathcal{R}\right)\boldsymbol{C}_i^r(2),\\
\hat{M}_{i+2N,j+N}^{\text{poly}} &= \delta_{ij}\frac{\Delta t}{\varepsilon|\Omega_i|}\sum_{r/\boldsymbol{M}_r\in\Omega_i}\boldsymbol{n}_{ir}(1)\boldsymbol{C}_i^r(2)-\frac{\Delta t}{\varepsilon|\Omega_i|}\sum_{r/\boldsymbol{M}_r\in\Omega_i\cap\Omega_j}\left(\mathcal{X}_k^r|_1,\mathcal{R}\right)\boldsymbol{C}_i^r(2),\\
\hat{M}_{i+2N,j+2N}^{\text{poly}} &= \delta_{ij}\left(1+\frac{\sigma_i\Delta t}{\varepsilon^2}\right)+\delta_{ij}\frac{\Delta t}{\varepsilon|\Omega_i|}\sum_{r/\boldsymbol{M}_r\in\Omega_i}\boldsymbol{n}_{ir}(2)\boldsymbol{C}_i^r(2)-\frac{\Delta t}{\varepsilon|\Omega_i|}\sum_{r/\boldsymbol{M}_r\in\Omega_i\cap\Omega_j}\left(\mathcal{X}_k^r|_2,\mathcal{R}\right)\boldsymbol{C}_i^r(2).
\end{aligned}$$

In the above formulas, we have used the following convention : for a vector $\boldsymbol{C}\in\mathbb{R}^2$, $\boldsymbol{C}(1)$ denotes its first component, while $\boldsymbol{C}(2)$ denotes its second component. For a matrix $\mathcal{X}\in\mathbb{R}^{2\times2}$, $\mathcal{X}_{|1}$ denotes its first column, while $\mathcal{X}_{|2}$ denotes its second column.

4.1.2 JLa scheme for conical meshes

We are now going to extend the scheme exposed in the previous subsection to conical meshes. As in the diffusion case, we add to the scheme contributions coming from the shoulder points.

This scheme reads

$$\begin{cases} E_i^{n+1} - E_i^n + \dfrac{\Delta t}{\varepsilon|\Omega_i|}\left(\displaystyle\sum_{r/\boldsymbol{M}_r\in\Omega_i}(\boldsymbol{C}_i^r,\boldsymbol{F}_r^{n+1}) + \sum_{r/\boldsymbol{M}_{r+1/2}\in\Omega_i}\left(\boldsymbol{C}_i^{r+1/2},\boldsymbol{F}_{r+1/2}^{n+1}\right)\right) = 0, & (4.10a)\\ \boldsymbol{F}_i^{n+1} - \boldsymbol{F}_i^n + \dfrac{\Delta t}{\varepsilon|\Omega_i|}\left(\displaystyle\sum_{r/\boldsymbol{M}_r\in\Omega_i}\boldsymbol{C}_i^r E_{ir}^{n+1} + \sum_{r/\boldsymbol{M}_{r+1/2}\in\Omega_i}\boldsymbol{C}_i^{r+1/2}E_{ir+1/2}^{n+1}\right) = -\dfrac{\sigma_i\Delta t}{\varepsilon^2}\boldsymbol{F}_i^{n+1}. & (4.10b)\end{cases}$$

The fluxes are defined by

$$\begin{cases} E_{ir}^{n+1} = E_i^{n+1} + \left(\boldsymbol{F}_i^{n+1} - \boldsymbol{F}_r^{n+1}, \boldsymbol{n}_{ir}\right) - \dfrac{\sigma_r}{\varepsilon}\left(\boldsymbol{F}_r^{n+1}, \boldsymbol{M}_r - \boldsymbol{x}_i\right), & (4.11a)\\ E_{ir+1/2}^{n+1} = E_i^{n+1} + \left(\boldsymbol{F}_i^{n+1} - \boldsymbol{F}_{r+1/2}^{n+1}, \boldsymbol{n}_{ir+1/2}\right) - \dfrac{\sigma_{r+1/2}}{\varepsilon}\left(\boldsymbol{F}_{r+1/2}^{n+1}, \boldsymbol{M}_{r+1/2} - \boldsymbol{x}_i\right), & (4.11b)\\ \hat{\gamma}_r\boldsymbol{F}_r^{n+1} = \displaystyle\sum_{i/\boldsymbol{M}_r\in\Omega_i}\left[\boldsymbol{C}_i^r E_i^{n+1} + \boldsymbol{C}_i^r\otimes\boldsymbol{n}_{ir}\boldsymbol{F}_i^{n+1}\right], & (4.11c)\\ \hat{\gamma}_{r+1/2}\boldsymbol{F}_{r+1/2}^{n+1} = \displaystyle\sum_{i/\boldsymbol{M}_{r+1/2}\in\Omega_i}\left[\boldsymbol{C}_i^{r+1/2}E_i^{n+1} + \boldsymbol{C}_i^{r+1/2}\otimes\boldsymbol{n}_{ir+1/2}\boldsymbol{F}_i^{n+1}\right]. & (4.11d)\end{cases}$$

Moreover, the matrices $\hat{\gamma}_r$ and $\hat{\gamma}_{r+1/2}$ read

$$\begin{aligned} \hat{\gamma}_r &= \sum_i \hat{\gamma}_{ir} = \sum_i\left(\hat{\alpha}_{ir} + \frac{\sigma_r}{\varepsilon}\hat{\beta}_{ir}\right), & \hat{\gamma}_{r+1/2} &= \sum_i \hat{\gamma}_{ir+1/2} = \sum_i\left(\hat{\alpha}_{ir+1/2} + \frac{\sigma_{r+1/2}}{\varepsilon}\hat{\beta}_{ir+1/2}\right),\\ \hat{\alpha}_r &= \sum_i \hat{\alpha}_{ir} = \sum_i \boldsymbol{C}_i^r\otimes\boldsymbol{n}_{ir}, & \hat{\alpha}_{r+1/2} &= \sum_i \hat{\alpha}_{ir+1/2} = \sum_i \boldsymbol{C}_i^{r+1/2}\otimes\boldsymbol{n}_{ir+1/2},\\ \hat{\beta}_r &= \sum_i \hat{\beta}_{ir} = \sum_i \hat{\beta}_{ir} = \boldsymbol{C}_i^r\otimes(\boldsymbol{M}_r - \boldsymbol{x}_i). & \hat{\beta}_{r+1/2} &= \sum_i \hat{\beta}_{ir+1/2} = \sum_i \boldsymbol{C}_i^{r+1/2}\otimes(\boldsymbol{M}_{r+1/2} - \boldsymbol{x}_i). \end{aligned}$$

Here, in order to simplify the notation, the sum over i runs for all i such that $\boldsymbol{M}_r\in\Omega_i$ (left part) or for all i such that $\boldsymbol{M}_{r+1/2}\in\Omega_i$ (right part).

As in the diffusion case, the matrix $\hat{\gamma}_r$ is invertible under some regularity condition on the mesh, but $\hat{\gamma}_{r+1/2}$ is, here again, of rank one (see Lemma 3.2). Hence, we apply the same treatment as in Section 3 to define the flux $\boldsymbol{F}_{r+1/2}$ at the shoulder point. We denote by j the index of the cell Ω_j such that $\boldsymbol{M}_{r+1/2}$ belongs to $\Omega_j\cap\Omega_i$. Then, we have

$$\boldsymbol{C}_i^{r+1/2} + \boldsymbol{C}_j^{r+1/2} = 0.$$

Hence,

$$\hat{\gamma}_{r+1/2} = 2\boldsymbol{C}_i^{r+1/2}\otimes\boldsymbol{n}_{ir+1/2} + \frac{\sigma_{r+1/2}}{\varepsilon}\boldsymbol{C}_i^{r+1/2}\otimes(\boldsymbol{x}_j - \boldsymbol{x}_i). \tag{4.12}$$

Inserting this equality into (4.11d), we infer

$$\left(\boldsymbol{F}_{r+1/2}^{n+1}, 2\boldsymbol{n}_{ir+1/2} + \frac{\sigma_{r+1/2}}{\varepsilon}(\boldsymbol{x}_j - \boldsymbol{x}_i)\right) = E_i^{n+1} - E_j^{n+1} + \left(\boldsymbol{F}_i^{n+1} + \boldsymbol{F}_j^{n+1}, \boldsymbol{n}_{ir+1/2}\right). \tag{4.13}$$

As in Section 3, we define the following vectors

$$\begin{cases} \boldsymbol{N}_{r+1/2} = 2\boldsymbol{n}_{ir+1/2} + \dfrac{\sigma_{r+1/2}}{\varepsilon}(\boldsymbol{x}_j - \boldsymbol{x}_i), \\ \boldsymbol{T}_{r+1/2} = \boldsymbol{N}_{r+1/2}^{\perp}. \end{cases}$$

This allows to write (4.13) as

$$\left(\boldsymbol{F}_{r+1/2}^{n+1}, \boldsymbol{N}_{r+1/2}\right) = E_i^{n+1} - E_j^{n+1} + \left(\boldsymbol{F}_i^{n+1} + \boldsymbol{F}_j^{n+1}, \boldsymbol{n}_{ir+1/2}\right). \tag{4.14}$$

Next, we need to compute the flux in the direction of $\boldsymbol{T}_{r+1/2}$. As in the diffusion case, we choose to define it as the average of the fluxes at nodes $\boldsymbol{M}_r$ and $\boldsymbol{M}_{r+1}$. Here again, other choices are possible (see for instance [1]). We thus have

$$\left(\boldsymbol{F}_{r+1/2}^{n+1}, \boldsymbol{T}_{r+1/2}\right) = \frac{1}{2}\left(\boldsymbol{F}_{r}^{n+1} + \boldsymbol{F}_{r+1}^{n+1}, \boldsymbol{T}_{r+1/2}\right), \tag{4.15}$$

that is, setting $\boldsymbol{Q}_i^r = \hat{\gamma}_r^{-1}\boldsymbol{C}_i^r$,

$$\begin{aligned}\left(\boldsymbol{F}_{r+1/2}^{n+1}, \boldsymbol{T}_{r+1/2}\right) = \frac{1}{2}\sum_{k/\boldsymbol{M}_r\in\Omega_k}\left[\left(\boldsymbol{Q}_k^r, \boldsymbol{T}_{r+1/2}\right)E_k^{n+1} + \left(\boldsymbol{F}_k^{n+1}, \boldsymbol{n}_{kr}\right)\left(\boldsymbol{Q}_k^r, \boldsymbol{T}_{r+1/2}\right)\right] \\ + \frac{1}{2}\sum_{p/\boldsymbol{M}_{r+1}\in\Omega_p}\left[\left(\boldsymbol{Q}_p^{r+1}, \boldsymbol{T}_{r+1/2}\right)E_p^{n+1} + \left(\boldsymbol{F}_p^{n+1}, \boldsymbol{n}_{pr+1}\right)\left(\boldsymbol{Q}_p^r, \boldsymbol{T}_{r+1/2}\right)\right].\end{aligned} \tag{4.16}$$

We are now going to compute the additional terms in the matrix $\hat{M}$, which are due to the fluxes at the shoulder points. We first write

$$\begin{aligned}\sum_{r/\boldsymbol{M}_{r+1/2}\in\Omega_i}\left(\boldsymbol{C}_i^{r+1/2}, \boldsymbol{F}_{r+1/2}^{n+1}\right) = \sum_{r/\boldsymbol{M}_{r+1/2}\in\Omega_i}\left(\boldsymbol{C}_i^{r+1/2}, \frac{\boldsymbol{N}_{r+1/2}}{\left\|\boldsymbol{N}_{r+1/2}\right\|^2}\right)\left(\boldsymbol{F}_{r+1/2}^{n+1}, \boldsymbol{N}_{r+1/2}\right) \\ + \sum_{r/\boldsymbol{M}_{r+1/2}\in\Omega_i}\left(\boldsymbol{F}_{r+1/2}^{n+1}, \boldsymbol{T}_{r+1/2}\right)\left(\boldsymbol{C}_i^{r+1/2}, \frac{\boldsymbol{T}_{r+1/2}}{\left\|\boldsymbol{T}_{r+1/2}\right\|^2}\right).\end{aligned}$$

Defining $\mathcal{N} = \left(\boldsymbol{C}_i^{r+1/2}, \frac{\boldsymbol{N}_{r+1/2}}{\left\|\boldsymbol{N}_{r+1/2}\right\|^2}\right)$ et $\mathcal{T} = \left(\boldsymbol{C}_i^{r+1/2}, \frac{\boldsymbol{T}_{r+1/2}}{\left\|\boldsymbol{T}_{r+1/2}\right\|^2}\right)$, we can write this term as follows:

$$\sum_{r/\boldsymbol{M}_{r+1/2}\in\Omega_i}\left(\boldsymbol{C}_i^{r+1/2}, \boldsymbol{F}_{r+1/2}^{n+1}\right) = \sum_{r/\boldsymbol{M}_{r+1/2}\in\Omega_i}\mathcal{N}\left(\boldsymbol{F}_{r+1/2}^{n+1}, \boldsymbol{N}_{r+1/2}\right) + \sum_{r/\boldsymbol{M}_{r+1/2}\in\Omega_i}\mathcal{T}\left(\boldsymbol{F}_{r+1/2}^{n+1}, \boldsymbol{T}_{r+1/2}\right). \tag{4.17}$$

Next, we insert (4.14) and (4.15) into (4.17). This gives

$$\begin{aligned}\sum_{r/\boldsymbol{M}_{r+1/2}\in\Omega_i}\left(\boldsymbol{C}_i^{r+1/2}, \boldsymbol{F}_{r+1/2}^{n+1}\right) = \sum_{r/\boldsymbol{M}_{r+1/2}\in\Omega_i}\mathcal{N}\left(E_i^{n+1} - E_j^{n+1} + \left(\boldsymbol{F}_i^{n+1} + \boldsymbol{F}_j^{n+1}, \boldsymbol{n}_{ir+1/2}\right)\right) \\ + \sum_{r/\boldsymbol{M}_{r+1/2}\in\Omega_i}\frac{\mathcal{T}}{2}\left(\boldsymbol{F}_r^{n+1} + \boldsymbol{F}_{r+1}^{n+1}, \boldsymbol{T}_{r+1/2}\right),\end{aligned} \tag{4.18}$$

where it is understood that j is the index of the cell Ω_j such that $\boldsymbol{M}_{r+1/2} \in \Omega_i \cap \Omega_j$. We split (4.18) into three termes, and we use (4.11c):

$$\begin{cases}\displaystyle\sum_{r+1/2}\mathcal{N}\left(E_i^{n+1} - E_j^{n+1} + \left(\boldsymbol{F}_i^{n+1} + \boldsymbol{F}_j^{n+1}, \boldsymbol{n}_{ir+1/2}\right)\right), & \text{(4.19a)} \\ \displaystyle\sum_{r+1/2}\frac{\mathcal{T}}{2}\left(\boldsymbol{F}_r^{n+1}, \boldsymbol{T}_{r+1/2}\right) = \sum_{r+1/2}\sum_k\left[\frac{\mathcal{T}}{2}\left(\boldsymbol{Q}_k^r, \boldsymbol{T}_{r+1/2}\right)E_k^{n+1} + \frac{\mathcal{T}}{2}\left(\hat{\gamma}_r^{-1}\hat{\alpha}_{kr}\boldsymbol{F}_k^{n+1}, \boldsymbol{T}_{r+1/2}\right)\right], & \text{(4.19b)} \\ \displaystyle\sum_{r+1/2}\frac{\mathcal{T}}{2}\left(\boldsymbol{F}_{r+1}^{n+1}, \boldsymbol{T}_{r+1/2}\right) = \sum_{r+1/2}\sum_k\left[\frac{\mathcal{T}}{2}\left(\boldsymbol{Q}_k^{r+1}, \boldsymbol{T}_{r+1/2}\right)E_k^{n+1} + \frac{\mathcal{T}}{2}\left(\hat{\gamma}_{r+1}^{-1}\hat{\alpha}_{kr+1}\boldsymbol{F}_k^{n+1}, \boldsymbol{T}_{r+1/2}\right)\right], & \text{(4.19c)}\end{cases}$$

where all sums run for r such that $\boldsymbol{M}_{r+1/2} \in \Omega_i$, and, here again, j is such that Ω_i and Ω_j share the shoulder point $\boldsymbol{M}_{r+1/2}$.

Keeping this in mind, we next insert the expression of the flux into the second eqation of the scheme (4.10b), in order to compute the corresponding additional terms. These terms read

$$\sum_{r+1/2}\boldsymbol{C}_i^{r+1/2}E_{ir+1/2}^{n+1}. \tag{4.20}$$

First, we use (4.11b) in order to compute $E_{ir+1/2}^{n+1}$

$$E_{ir+1/2}^{n+1} = E_i^{n+1} + \left(\boldsymbol{F}_i^{n+1} - \boldsymbol{F}_{r+1/2}^{n+1}, \boldsymbol{n}_{ir+1/2}\right) - \frac{\sigma_{r+1/2}}{\varepsilon}\left(\boldsymbol{F}_{r+1/2}^{n+1}, \boldsymbol{M}_{r+1/2} - \boldsymbol{x}_i\right),$$

that is,

$$E_{ir+1/2}^{n+1} = E_i^{n+1} + \left(\boldsymbol{F}_i^{n+1}, \boldsymbol{n}_{ir+1/2}\right) - \left(\boldsymbol{F}_{r+1/2}^{n+1}, \boldsymbol{n}_{ir+1/2} + \frac{\sigma_{r+1/2}}{\varepsilon}\left(\boldsymbol{M}_{r+1/2} - \boldsymbol{x}_i\right)\right).$$

We define $\mathcal{R}_{ir+1/2} = \boldsymbol{n}_{ir+1/2} + \frac{\sigma_{r+1/2}}{\varepsilon}\left(\boldsymbol{M}_{r+1/2} - \boldsymbol{x}_i\right)$, which allows us to write

$$E_{ir+1/2}^{n+1} = E_i^{n+1} + \left(\boldsymbol{F}_i^{n+1}, \boldsymbol{n}_{ir+1/2}\right) - \left(\boldsymbol{F}_{r+1/2}^{n+1}, \mathcal{R}_{ir+1/2}\right).$$

This also reads

$$E_{ir+1/2}^{n+1} = E_i^{n+1} + \left(\boldsymbol{F}_i^{n+1}, \boldsymbol{n}_{ir+1/2}\right) - \left(\mathcal{R}_{ir+1/2}, \frac{\boldsymbol{N}_{r+1/2}}{\left\|\boldsymbol{N}_{r+1/2}\right\|^2}\right)\left(\boldsymbol{F}_{r+1/2}^{n+1}, \boldsymbol{N}_{r+1/2}\right)$$
$$- \left(\boldsymbol{F}_{r+1/2}^{n+1}, \boldsymbol{T}_{r+1/2}\right)\left(\mathcal{R}_{ir+1/2}, \frac{\boldsymbol{T}_{r+1/2}}{\left\|\boldsymbol{T}_{r+1/2}\right\|^2}\right).$$

We set $\tilde{\mathcal{N}} = \left(\mathcal{R}_{ir+1/2}, \frac{\boldsymbol{N}_{r+1/2}}{\left\|\boldsymbol{N}_{r+1/2}\right\|^2}\right)$ and $\tilde{\mathcal{T}} = \left(\mathcal{R}_{ir+1/2}, \frac{\boldsymbol{T}_{r+1/2}}{\left\|\boldsymbol{T}_{r+1/2}\right\|^2}\right)$, hence

$$E_{ir+1/2}^{n+1} = E_i^{n+1} + \left(\boldsymbol{F}_i^{n+1}, \boldsymbol{n}_{ir+1/2}\right) - \left(\boldsymbol{F}_{r+1/2}^{n+1}, \boldsymbol{N}_{r+1/2}\right)\tilde{\mathcal{N}} - \left(\boldsymbol{F}_{r+1/2}^{n+1}, \boldsymbol{T}_{r+1/2}\right)\tilde{\mathcal{T}}.$$

We insert (4.14) and (4.15) into the above equality:

$$E_{ir+1/2}^{n+1} = E_i^{n+1} + \left(\boldsymbol{F}_i^{n+1}, \boldsymbol{n}_{ir+1/2}\right) - \left(E_i^{n+1} - E_j^{n+1} + \left(\boldsymbol{F}_i^{n+1} + \boldsymbol{F}_j^{n+1}, \boldsymbol{n}_{ir+1/2}\right)\right)\tilde{\mathcal{N}}$$
$$- \left(\boldsymbol{F}_r^{n+1}, \boldsymbol{T}_{r+1/2}\right)\frac{\tilde{\mathcal{T}}}{2} - \left(\boldsymbol{F}_{r+1}^{n+1}, \boldsymbol{T}_{r+1/2}\right)\frac{\tilde{\mathcal{T}}}{2},$$

where j is such that $\boldsymbol{M}_r \in \Omega_i \cap \Omega_j$. The additional term in the system reads $\sum_{r/\boldsymbol{M}_{r+1/2}\in\Omega_i} \boldsymbol{C}_i^{r+1/2} E_{ir+1/2}^{n+1}$, which we split into three terms:

$$\begin{cases} \displaystyle\sum_{r/\boldsymbol{M}_{r+1/2}\in\Omega_i} \left(E_i^{n+1} + \left(\boldsymbol{F}_i^{n+1}, \boldsymbol{n}_{ir+1/2}\right) - \left(E_i^{n+1} - E_j^{n+1} + \left(\boldsymbol{F}_i^{n+1} + \boldsymbol{F}_j^{n+1}, \boldsymbol{n}_{ir+1/2}\right)\right)\tilde{\mathcal{N}}\right)\boldsymbol{C}_i^{r+1/2}, & (4.21a) \\ \displaystyle -\sum_{r/\boldsymbol{M}_{r+1/2}\in\Omega_i} \left(\boldsymbol{F}_r^{n+1}, \boldsymbol{T}_{r+1/2}\right)\frac{\tilde{\mathcal{T}}}{2}\boldsymbol{C}_i^{r+1/2}, & (4.21b) \\ \displaystyle -\sum_{r/\boldsymbol{M}_{r+1/2}\in\Omega_i} \left(\boldsymbol{F}_{r+1}^{n+1}, \boldsymbol{T}_{r+1/2}\right)\frac{\tilde{\mathcal{T}}}{2}\boldsymbol{C}_i^{r+1/2}. & (4.21c) \end{cases}$$

In (4.21a), as before, the index j is such that $\boldsymbol{M}_{r+1/2} \in \Omega_i \cap \Omega_j$.

We are now going to use (4.19a)-(4.19b)-(4.19c) on the one hand, and (4.21a)-(4.21b)-(4.21c) on the other hand, to compute the additional terms in the matrix of the system. Let $\hat{M}$ be the matrix of the system. As before, the system reads

$$\hat{M}\boldsymbol{X}^{\boldsymbol{n+1}} = \boldsymbol{X}^{\boldsymbol{n}} \quad \text{with } \boldsymbol{X}^n = \begin{pmatrix} E^n \\ F_x^n \\ F_y^n \end{pmatrix}.$$

Each block of the vector $\boldsymbol{X}^n$ is of size N (the number of cells). Collecting all the terms, we have, denoting by $\hat{M}^{\text{poly}}$ the matrix of the polygonal case, for any i, j such that $1 \le i, j \le N$:

$$\hat{M}_{ij} \quad = \quad \hat{M}_{ij}^{\text{poly}} + \delta_{ij}\frac{\Delta t}{\varepsilon|\Omega_i|}\sum_{r/\boldsymbol{M}_{r+1/2}\in\Omega_i} \mathcal{N} - \frac{\Delta t}{\varepsilon|\Omega_i|}\sum_{r/\boldsymbol{M}_{r+1/2}\in\Omega_i\cap\Omega_j} \mathcal{N}$$

$$
\begin{aligned}
&+\frac{\Delta t}{\varepsilon|\Omega_i|}\sum_{r/\boldsymbol{M}_{r+1/2}\in\Omega_i\cap\Omega_j}\frac{\mathcal{T}}{2}\left(\boldsymbol{Q}_j^r+\boldsymbol{Q}_j^{r+1},\boldsymbol{T}_{r+1/2}\right)\\
\hat{M}_{i,j+N} &= \hat{M}^{\text{poly}}_{i,j+N}+\delta_{ij}\frac{\Delta t}{\varepsilon|\Omega_i|}\sum_{r/\boldsymbol{M}_{r+1/2}\in\Omega_i}\mathcal{N}\boldsymbol{n}_{ir+1/2}(1)+\frac{\Delta t}{\varepsilon|\Omega_i|}\sum_{r/\boldsymbol{M}_{r+1/2}\in\Omega_i\cap\Omega_j}\mathcal{N}\boldsymbol{n}_{ir+1/2}(1)\\
&+\frac{\Delta t}{\varepsilon|\Omega_i|}\sum_{r/\boldsymbol{M}_{r+1/2}\Omega_i\cap\Omega_j}\frac{\mathcal{T}}{2}\left(\mathcal{X}_j^r|_0+\mathcal{X}_j^{r+1}|_0,\boldsymbol{T}_{r+1/2}\right),\\
\hat{M}_{i,j+2N} &= \hat{M}^{\text{poly}}_{i,j+2N}+\delta_{ij}\frac{\Delta t}{\varepsilon|\Omega_i|}\sum_{r/\boldsymbol{M}_{r+1/2}\in\Omega_i}\mathcal{N}\boldsymbol{n}_{ir+1/2}(2)+\frac{\Delta t}{\varepsilon|\Omega_i|}\sum_{r,\boldsymbol{M}_{r+1/2}\in\Omega_i\cap\Omega_j}\mathcal{N}\boldsymbol{n}_{ir+1/2}(2)\\
&+\frac{\Delta t}{\varepsilon|\Omega_i|}\sum_{r/\boldsymbol{M}_{r+1/2}\in\Omega_i\cap\Omega_j}\frac{\mathcal{T}}{2}\left(\mathcal{X}_j^r|_2+\mathcal{X}_j^{r+1}|_2,\boldsymbol{T}_{r+1/2}\right)\\
\hat{M}_{i+N,j} &= \hat{M}^{\text{poly}}_{i+N,j}+\delta_{ij}\frac{\Delta t}{\varepsilon|\Omega_i|}\sum_{r/\boldsymbol{M}_{r+1/2}\in\Omega_i}\left(1-\tilde{\mathcal{N}}\right)\boldsymbol{C}_i^{r+1/2}(1)+\frac{\Delta t}{\varepsilon|\Omega_i|}\sum_{r/\boldsymbol{M}_{r+1/2}\in\Omega_i\cap\Omega_j}\tilde{\mathcal{N}}\boldsymbol{C}_i^{r+1/2}(1)\\
&-\frac{\Delta t}{\varepsilon|\Omega_i|}\sum_{r/\boldsymbol{M}_{r+1/2}\in\Omega_i\cap\Omega_j}\frac{\tilde{\mathcal{T}}}{2}\left(\boldsymbol{Q}_j^r+\boldsymbol{Q}_j^{r+1},\boldsymbol{T}_{r+1/2}\right)\boldsymbol{C}_i^{r+1/2}(1),\\
\hat{M}_{i+N,j+N} &= \hat{M}^{\text{poly}}_{i+N,j+N}+\delta_{ij}\frac{\Delta t}{\varepsilon|\Omega_i|}\sum_{r/\boldsymbol{M}_{r+1/2}\in\Omega_i}\left(1-\tilde{\mathcal{N}}\right)\boldsymbol{n}_{ir+1/2}(1)\boldsymbol{C}_i^{r+1/2}(1)\\
&-\frac{\Delta t}{\varepsilon|\Omega_i|}\sum_{r/\boldsymbol{M}_{r+1/2}\in\Omega_i\cap\Omega_j}\tilde{\mathcal{N}}\boldsymbol{n}_{ir+1/2}(1)\boldsymbol{C}_i^{r+1/2}(1)\\
&-\frac{\Delta t}{\varepsilon|\Omega_i|}\sum_{r/\boldsymbol{M}_{r+1/2}\in\Omega_i\cap\Omega_j}\frac{\tilde{\mathcal{T}}}{2}\left(\mathcal{X}_j^r|_0+\mathcal{X}_j^{r+1}|_0,\boldsymbol{T}_{r+1/2}\right)\boldsymbol{C}_i^{r+1/2}(1)\\
\hat{M}_{i+N,j+2N} &= \hat{M}^{\text{poly}}_{i+N,j+2N}+\delta_{ij}\frac{\Delta t}{\varepsilon|\Omega_i|}\sum_{r/\boldsymbol{M}_{r+1/2}\in\Omega_i}\left(1-\tilde{\mathcal{N}}\right)\boldsymbol{n}_{ir+1/2}(2)\boldsymbol{C}_i^{r+1/2}(1)\\
&-\frac{\Delta t}{\varepsilon|\Omega_i|}\sum_{r/\boldsymbol{M}_{r+1/2}\in\Omega_i\cap\Omega_j}\tilde{\mathcal{N}}\boldsymbol{n}_{ir+1/2}(2)\boldsymbol{C}_i^{r+1/2}(1)\\
&-\frac{\Delta t}{\varepsilon|\Omega_i|}\sum_{r/\boldsymbol{M}_{r+1/2}\in\Omega_i\cap\Omega_j}\frac{\tilde{\mathcal{T}}}{2}\left(\mathcal{X}_j^r|_2+\mathcal{X}_j^{r+1}|_2,\boldsymbol{T}_{r+1/2}\right)\boldsymbol{C}_i^{r+1/2}(1),\\
\hat{M}_{i+2N,j} &= \hat{M}^{\text{poly}}_{i+2N,j}+\delta_{ij}\frac{\Delta t}{\varepsilon|\Omega_i|}\sum_{r/\boldsymbol{M}_{r+1/2}\in\Omega_i}\left(1-\tilde{\mathcal{N}}\right)\boldsymbol{C}_i^{r+1/2}(2)+\frac{\Delta t}{\varepsilon|\Omega_i|}\sum_{r/\boldsymbol{M}_{r+1/2}\in\Omega_i\cap\Omega_j}\tilde{\mathcal{N}}\boldsymbol{C}_i^{r+1/2}(2)\\
&-\frac{\Delta t}{\varepsilon|\Omega_i|}\sum_{r/\boldsymbol{M}_{r+1/2}\in\Omega_i\cap\Omega_j}\frac{\tilde{\mathcal{T}}}{2}\left(\boldsymbol{Q}_j^r+\boldsymbol{Q}_j^{r+1},\boldsymbol{T}_{r+1/2}\right)\boldsymbol{C}_i^{r+1/2}(2),\\
\hat{M}_{i+2N,j+N} &= \hat{M}^{\text{poly}}_{i+2N,j+N}+\delta_{ij}\frac{\Delta t}{\varepsilon|\Omega_i|}\sum_{r/\boldsymbol{M}_{r+1/2}\in\Omega_i}\left(1-\tilde{\mathcal{N}}\right)\boldsymbol{n}_{ir+1/2}(1)\boldsymbol{C}_i^{r+1/2}(2)\\
&-\frac{\Delta t}{\varepsilon|\Omega_i|}\sum_{r/\boldsymbol{M}_{r+1/2}\in\Omega_i\cap\Omega_j}\tilde{\mathcal{N}}\boldsymbol{n}_{ir+1/2}(1)\boldsymbol{C}_i^{r+1/2}(2)\\
&-\frac{\Delta t}{\varepsilon|\Omega_i|}\sum_{r/\boldsymbol{M}_{r+1/2}\in\Omega_i\cap\Omega_j}\frac{\tilde{\mathcal{T}}}{2}\left(\mathcal{X}_j^r|_1+\mathcal{X}_j^{r+1}|_1,\boldsymbol{T}_{r+1/2}\right)\boldsymbol{C}_i^{r+1/2}(2)\\
\hat{M}_{i+2N,j+2N} &= \hat{M}^{\text{poly}}_{i+2N,j+2N}+\delta_{ij}\frac{\Delta t}{\varepsilon|\Omega_i|}\sum_{r/\boldsymbol{M}_{r+1/2}\in\Omega_i}\left(1-\tilde{\mathcal{N}}\right)\boldsymbol{n}_{ir+1/2}(2)\boldsymbol{C}_i^{r+1/2}(2)
\end{aligned}
$$

$$-\frac{\Delta t}{\varepsilon|\Omega_i|}\sum_{r/\boldsymbol{M}_{r+1/2}\in\Omega_i\cap\Omega_j}\tilde{\mathcal{N}}\boldsymbol{n}_{ir+1/2}(2)\boldsymbol{C}_i^{r+1/2}(2)$$
$$-\frac{\Delta t}{\varepsilon|\Omega_i|}\sum_{r+1/2}\frac{\tilde{\mathcal{T}}}{2}\left(\mathcal{X}_j^r|_2+\mathcal{X}_j^{r+1}|_2,\boldsymbol{T}_{r+1/2}\right)\boldsymbol{C}_i^{r+1/2}(2),$$

where we have used the notation $\mathcal{X}_k^r=\hat{\gamma}_r^{-1}\hat{\alpha}_{kr}$ and $\mathcal{X}_k^{r+1}=\hat{\gamma}_{r+1}^{-1}\hat{\alpha}_{kr+1}$. As in the polygonal case, for a vector $\boldsymbol{C}\in\mathbb{R}^2$, $\boldsymbol{C}(1)$ denotes its first component, while $\boldsymbol{C}(2)$ denotes its second component. For a matrix $\mathcal{X}\in\mathbb{R}^{2\times2}$, $\mathcal{X}_{|1}$ denotes its first column, while $\mathcal{X}_{|2}$ denotes its second column.

4.2 JLb scheme

In this section, we present a second version of the preceding scheme, in which the last term of the second equation of (1.1) is discretized at the nodes instead of being discretized at the cell centers.

4.2.1 JLb scheme for polygonal meshes

If the mesh is polygonal, the scheme reads [11]

$$\begin{cases}E_i^{n+1}-E_i^n+\dfrac{\Delta t}{\varepsilon|\Omega_i|}\displaystyle\sum_r\left(\boldsymbol{C}_i^r,\boldsymbol{F}_r^{n+1}\right)=0, & (4.22a)\\ \boldsymbol{F}_i^{n+1}-\boldsymbol{F}_i^n+\dfrac{\Delta t}{\varepsilon|\Omega_i|}\displaystyle\sum_r\boldsymbol{C}_i^rE_{ir}^{n+1}=-\dfrac{\Delta t}{|\Omega_i|\varepsilon^2}\sum_r\left(\boldsymbol{C}_i^r\otimes(\boldsymbol{M}_r-\boldsymbol{x}_i)\right)\sigma_r\boldsymbol{F}_r^{n+1}, & (4.22b)\end{cases}$$

where the fluxes are defined by

$$\begin{cases}E_{ir}^{n+1}=E_i^{n+1}+\left(\boldsymbol{F}_i^{n+1}-\boldsymbol{F}_r^{n+1},\boldsymbol{n}_{ir}\right)-\dfrac{\sigma_r}{\varepsilon}\left(\boldsymbol{F}_r^{n+1},\boldsymbol{M}_r-\boldsymbol{x}_i\right), & (4.23a)\\ \displaystyle\sum_{i/\boldsymbol{M}_r\in\Omega_i}\left(\hat{\alpha}_{ir}+\frac{\sigma_r}{\varepsilon}\hat{\beta}_{ir}\right)\boldsymbol{F}_r^{n+1}=\sum_{i/\boldsymbol{M}_r\in\Omega_i}\boldsymbol{C}_i^rE_i^{n+1}+\sum_{i/\boldsymbol{M}_r\in\Omega_i}\boldsymbol{C}_i^r\otimes\boldsymbol{n}_{ir}\boldsymbol{F}_i^{n+1}. & (4.23b)\end{cases}$$

and the local matrices are $\hat{\alpha}_{ir}=\boldsymbol{C}_i^r\otimes\boldsymbol{n}_{ir}$ and $\hat{\beta}_{ir}=\boldsymbol{C}_i^r\otimes(\boldsymbol{M}_r-\boldsymbol{x}_i)$ with $\boldsymbol{n}_{ir}=\frac{\boldsymbol{C}_i^r}{\|\boldsymbol{C}_i^r\|}$.

From now on, we set $\hat{\alpha}_r=\displaystyle\sum_{i/\boldsymbol{M}_r\in\Omega_i}\hat{\alpha}_{ir}$, $\hat{\beta}_r=\displaystyle\sum_{i/\boldsymbol{M}_r\in\Omega_i}\hat{\beta}_{ir}$ and $\hat{\gamma}_r=\displaystyle\sum_{i/\boldsymbol{M}_r\in\Omega_i}\hat{\gamma}_{ir}=\sum_{i/\boldsymbol{M}_r\in\Omega_i}\left(\hat{\alpha}_{ir}+\frac{\sigma}{\varepsilon}\hat{\beta}_{ir}\right)$

The scheme also reads

$$\begin{cases}E_i^{n+1}-E_i^n+\dfrac{\Delta t}{\varepsilon|\Omega_i|}\displaystyle\sum_{r/\boldsymbol{M}_r\in\Omega_i}\left(\boldsymbol{C}_i^r,\boldsymbol{F}_r^{n+1}\right)=0, & (4.24a)\\ \boldsymbol{F}_i^{n+1}-\boldsymbol{F}_i^n+\dfrac{\Delta t}{\varepsilon|\Omega_i|}\displaystyle\sum_{r/\boldsymbol{M}_r\in\Omega_i}\boldsymbol{C}_i^rE_{ir}^{*,n+1}=0, & (4.24b)\end{cases}$$

with the fluxes

$$\begin{cases}E_{ir}^{*,n+1}=E_i^{n+1}+\left(\boldsymbol{F}_i^{n+1}-\boldsymbol{F}_r^{n+1},\boldsymbol{n}_{ir}\right), & (4.25a)\\ \displaystyle\sum_{i/\boldsymbol{M}_r\in\Omega_i}\left(\hat{\alpha}_{ir}+\frac{\sigma}{\varepsilon}\hat{\beta}_{ir}\right)\boldsymbol{F}_r^{n+1}=\sum_{i/\boldsymbol{M}_r\in\Omega_i}\boldsymbol{C}_i^rE_i^{n+1}+\sum_{i/\boldsymbol{M}_r\in\Omega_i}\boldsymbol{C}_i^r\otimes\boldsymbol{n}_{ir}\boldsymbol{F}_i^{n+1}. & (4.25b)\end{cases}$$

With this form, it is clear that we can apply exactly the same method as with the JLa scheme. The only difference is the last term of (4.24b). These computations are left to the reader, and we only give below the matrix we obtain in such a way. We use the same convention as before, denoting by $\hat{M}^{\text{poly}}$ the matrix of the system. The indices i and j are such that $1\leq i,j\leq N$.

$$\hat{M}_{ij}^{\text{poly}}\quad=\quad\delta_{ij}+\frac{\Delta t}{\varepsilon|\Omega_i|}\sum_{r/\boldsymbol{M}_r\in\Omega_i\cap\Omega_j}\left(\boldsymbol{C}_i^r,\boldsymbol{Q}_j^r\right),$$

$$
\begin{aligned}
\hat{M}^{\mathrm{poly}}_{i,j+N} &= \frac{\Delta t}{\varepsilon|\Omega_i|}\sum_{r/\boldsymbol{M}_r\in\Omega_i\cap\Omega_j}\left(\boldsymbol{C}_i^r,\mathcal{X}_j^r|_1\right),\\
\hat{M}^{\mathrm{poly}}_{i,j+2N} &= \frac{\Delta t}{\varepsilon|\Omega_i|}\sum_{r/\boldsymbol{M}_r\in\Omega_i\cap\Omega_j}\left(\boldsymbol{C}_i^r,\mathcal{X}_j^r|_2\right),\\
\hat{M}^{\mathrm{poly}}_{i+N,j} &= \delta_{ij}\frac{\Delta t}{\varepsilon|\Omega_i|}\sum_{r/\boldsymbol{M}_r\in\Omega_i}\boldsymbol{C}_i^r(1)-\frac{\Delta t}{\varepsilon|\Omega_i|}\sum_{r/\boldsymbol{M}_r\in\Omega_i\cap\Omega_j}\left(\hat{\gamma}_r^{-1}\boldsymbol{C}_j^r,\mathcal{R}\right)\boldsymbol{C}_i^r(1),\\
\hat{M}^{\mathrm{poly}}_{i+N,j+N} &= \delta_{ij}+\delta_{ij}\frac{\Delta t}{\varepsilon|\Omega_i|}\sum_{r/\boldsymbol{M}_r\in\Omega_i}\boldsymbol{n}_{ir}(1)\boldsymbol{C}_i^r(1)-\frac{\Delta t}{\varepsilon|\Omega_i|}\sum_{r/\boldsymbol{M}_r\in\Omega_i\cap\Omega_j}\left(\mathcal{X}_j^r|_1,\mathcal{R}\right)\boldsymbol{C}_i^r(1),\\
\hat{M}^{\mathrm{poly}}_{i+N,j+2N} &= \delta_{ij}\frac{\Delta t}{\varepsilon|\Omega_i|}\sum_{r/\boldsymbol{M}_r\in\Omega_i}\boldsymbol{n}_{ir}(2)\boldsymbol{C}_i^r(1)-\frac{\Delta t}{\varepsilon|\Omega_i|}\sum_{r/\boldsymbol{M}_r\in\Omega_i\cap\Omega_j}\left(\mathcal{X}_j^r|_2,\mathcal{R}\right)\boldsymbol{C}_i^r(1),\\
\hat{M}^{\mathrm{poly}}_{i+2N,j} &= \delta_{ij}\frac{\Delta t}{\varepsilon|\Omega_i|}\sum_{r/\boldsymbol{M}_r\in\Omega_i}\boldsymbol{C}_i^r(2)-\frac{\Delta t}{\varepsilon|\Omega_i|}\sum_{r/\boldsymbol{M}_r\in\Omega_i\cap\Omega_j}\left(\hat{\gamma}_r^{-1}\boldsymbol{C}_j^r,\mathcal{R}\right)\boldsymbol{C}_i^r(2),\\
\hat{M}^{\mathrm{poly}}_{i+2N,j+N} &= \delta_{ij}\frac{\Delta t}{\varepsilon|\Omega_i|}\sum_{r/\boldsymbol{M}_r\in\Omega_i}\boldsymbol{n}_{ir}(1)\boldsymbol{C}_i^r(2)-\frac{\Delta t}{\varepsilon|\Omega_i|}\sum_{r/\boldsymbol{M}_r\in\Omega_i\cap\Omega_j}\left(\mathcal{X}_k^r|_1,\mathcal{R}\right)\boldsymbol{C}_i^r(2),\\
\hat{M}^{\mathrm{poly}}_{i+2N,j+2N} &= \delta_{ij}+\delta_{ij}\frac{\Delta t}{\varepsilon|\Omega_i|}\sum_{r/\boldsymbol{M}_r\in\Omega_i}\boldsymbol{n}_{ir}(2)\boldsymbol{C}_i^r(2)-\frac{\Delta t}{\varepsilon|\Omega_i|}\sum_{r/\boldsymbol{M}_r\in\Omega_i\cap\Omega_j}\left(\mathcal{X}_k^r|_2,\mathcal{R}\right)\boldsymbol{C}_i^r(2),
\end{aligned}
$$

where we have used the following notation: $\mathcal{R}=\boldsymbol{n}_{ir}$ (note that in the JLa case we have $\mathcal{R}=\boldsymbol{n}_{ir}+\frac{\sigma_r}{\varepsilon}\left(\boldsymbol{M}_r-\boldsymbol{x}_i\right)$), and $\mathcal{X}_k^r=\hat{\gamma}_r^{-1}\hat{\alpha}_{kr}$.

4.2.2 JLb scheme for conical meshes

We now extend the Jlb scheme to conical meshes. It reads (we recall that $\boldsymbol{C}_j^r=\nabla_{\boldsymbol{M}_r}|\Omega_j|$ and $\boldsymbol{C}_j^{r+1/2}=\nabla_{\boldsymbol{M}_{r+1/2}}|\Omega_j|$)

$$
\begin{cases}
E_i^{n+1}-E_i^n+\dfrac{\Delta t}{\varepsilon|\Omega_i|}\left(\displaystyle\sum_{r/\boldsymbol{M}_r\in\Omega_i}\left(\boldsymbol{C}_i^r,\boldsymbol{F}_r^{n+1}\right)+\sum_{r/\boldsymbol{M}_{r+1/2}\in\Omega_i}\left(\boldsymbol{C}_i^{r+1/2},\boldsymbol{F}_{r+1/2}^{n+1}\right)\right)=0, & (4.26a)\\
\boldsymbol{F}_i^{n+1}-\boldsymbol{F}_i^n+\dfrac{\Delta t}{\varepsilon|\Omega_i|}\left(\displaystyle\sum_{r/\boldsymbol{M}_r\in\Omega_i}\boldsymbol{C}_i^rE_{ir}^{n+1}+\sum_{r/\boldsymbol{M}_{r+1/2}\in\Omega_i}\boldsymbol{C}_i^{r+1/2}E_{ir+1/2}^{n+1}\right)= & (4.26b)\\
-\dfrac{\Delta t}{|\Omega_i|\varepsilon^2}\displaystyle\sum_{r/\boldsymbol{M}_r\in\Omega_i}\boldsymbol{C}_i^r\otimes\left(\boldsymbol{M}_r-\boldsymbol{x}_i\right)\sigma_r\boldsymbol{F}_r^{n+1} & (4.26c)\\
-\dfrac{\Delta t}{|\Omega_i|\varepsilon^2}\displaystyle\sum_{r/\boldsymbol{M}_{r+1/2}\in\Omega_i}\boldsymbol{C}_i^{r+1/2}\otimes\left(\boldsymbol{M}_{r+1/2}-\boldsymbol{x}_i\right)\sigma_{r+1/2}\boldsymbol{F}_{r+1/2}^{n+1}. & (4.26d)
\end{cases}
$$

with the fluxes

$$
\begin{cases}
E_{ir}^{n+1}=E_i^{n+1}+\left(\boldsymbol{F}_i^{n+1}-\boldsymbol{F}_r^{n+1},\boldsymbol{n}_{ir}\right), & (4.27a)\\
E_{ir+1/2}^{n+1}=E_i^{n+1}+\left(\boldsymbol{F}_i^{n+1}-\boldsymbol{F}_{r+1/2}^{n+1},\boldsymbol{n}_{ir+1/2}\right), & (4.27b)\\
\hat{\gamma}_r\boldsymbol{F}_r^{n+1}=\displaystyle\sum_{i/\boldsymbol{M}_r\in\Omega_i}\left[\boldsymbol{C}_i^rE_i^{n+1}+\boldsymbol{C}_i^r\otimes\boldsymbol{n}_{ir}\boldsymbol{F}_i^{n+1}\right], & (4.27c)\\
\hat{\gamma}_{r+1/2}\boldsymbol{F}_{r+1/2}^{n+1}=\displaystyle\sum_{i/\boldsymbol{M}_{r+1/2}\in\Omega_i}\left[\boldsymbol{C}_i^{r+1/2}E_i^{n+1}+\boldsymbol{C}_i^{r+1/2}\otimes\boldsymbol{n}_{ir+1/2}\boldsymbol{F}_i^{n+1}\right]. & (4.27d)
\end{cases}
$$

and the matrices

$$\hat{\gamma}_r = \sum_i \hat{\gamma}_{ir} = \sum_i \left(\hat{\alpha}_{ir} + \frac{\sigma_r}{\varepsilon}\hat{\beta}_{ir}\right), \qquad \hat{\gamma}_{r+1/2} = \sum_i \hat{\gamma}_{ir+1/2} = \sum_i \left(\hat{\alpha}_{r+1/2} + \frac{\sigma_{r+1/2}}{\varepsilon}\hat{\beta}_{r+1/2}\right),$$
$$\hat{\alpha}_r = \sum_i \hat{\alpha}_{ir} = \sum_i \boldsymbol{C}_i^r \otimes \boldsymbol{n}_{ir}, \qquad \hat{\alpha}_{r+1/2} = \sum_i \hat{\alpha}_{ir+1/2} = \sum_i \boldsymbol{C}_i^{r+1/2} \otimes \boldsymbol{n}_{ir+1/2},$$
$$\hat{\beta}_r = \sum_i \hat{\beta}_{ir} = \sum_i \hat{\beta}_{ir} = \boldsymbol{C}_i^r \otimes (\boldsymbol{M}_r - \boldsymbol{x}_i). \qquad \hat{\beta}_{r+1/2} = \sum_i \hat{\beta}_{ir+1/2} = \sum_i \boldsymbol{C}_i^{r+1/2} \otimes (\boldsymbol{M}_{r+1/2} - \boldsymbol{x}_i),$$

where all the sums run for i such that $\boldsymbol{M}_r \in \Omega_i$ (left column) and $\boldsymbol{M}_{r+1/2} \in \Omega_i$ (right column). As in the polygonal case, we need to change the definition of the flux (4.27b). We define $\mathcal{R}_{ir+1/2} = \boldsymbol{n}_{ir+1/2}$, and get

$$E_{ir+1/2}^{n+1} = E_i^{n+1} + \left(\boldsymbol{F}_i^{n+1}, \boldsymbol{n}_{ir+1/2}\right) - \left(\boldsymbol{F}_{r+1/2}^{n+1}, \mathcal{R}_{ir+1/2}\right).$$

Note that in the JLa case, we had $\mathcal{R}_{ir+1/2} = \boldsymbol{n}_{ir+1/2} + \frac{\sigma_{r+1/2}}{\varepsilon}\left(\boldsymbol{M}_{r+1/2} - \boldsymbol{x}_i\right)$. We proceed exactly as in the JLa case. For this purpose, we introduce

$$\begin{cases} \boldsymbol{N}_{r+1/2} = 2\boldsymbol{n}_{ir+1/2} + \dfrac{\sigma_{r+1/2}}{\varepsilon}\left(\boldsymbol{x}_j - \boldsymbol{x}_i\right), \\ \boldsymbol{T}_{r+1/2} = \boldsymbol{N}_{r+1/2}^{\perp}, \end{cases}$$

where j is such that $\boldsymbol{M}_{r+1/2} \in \Omega_i \cap \Omega_j$. We also define

$$\mathcal{N} = \left(\boldsymbol{C}_i^{r+1/2}, \frac{\boldsymbol{N}_{r+1/2}}{\left\|\boldsymbol{N}_{r+1/2}\right\|^2}\right), \quad \mathcal{T} = \left(\boldsymbol{C}_i^{r+1/2}, \frac{\boldsymbol{T}_{r+1/2}}{\left\|\boldsymbol{T}_{r+1/2}\right\|^2}\right), \quad \tilde{\mathcal{N}} = \left(\mathcal{R}_{ir+1/2}, \frac{\boldsymbol{N}_{r+1/2}}{\left\|\boldsymbol{N}_{r+1/2}\right\|^2}\right),$$
$$\tilde{\mathcal{T}} = \left(\mathcal{R}_{ir+1/2}, \frac{\boldsymbol{T}_{r+1/2}}{\left\|\boldsymbol{T}_{r+1/2}\right\|^2}\right), \quad \boldsymbol{Q}_i^r = \hat{\gamma}_r^{-1}\boldsymbol{C}_i^r, \quad \mathcal{X}_k^r = \hat{\gamma}_r^{-1}\hat{\alpha}_{kr}$$

We these notation, the matrix of the system reads

$$\begin{aligned}
\hat{M}_{ij} &= \hat{M}_{ij}^{\text{poly}} + \delta_{ij}\frac{\Delta t}{\varepsilon|\Omega_i|}\sum_{r/\boldsymbol{M}_{r+1/2}\in\Omega_i}\mathcal{N} - \frac{\Delta t}{\varepsilon|\Omega_i|}\sum_{r/\boldsymbol{M}_{r+1/2}\in\Omega_i\cap\Omega_j}\mathcal{N} \\
&\quad + \frac{\Delta t}{\varepsilon|\Omega_i|}\sum_{r/\boldsymbol{M}_{r+1/2}\in\Omega_i\cap\Omega_j}\frac{\mathcal{T}}{2}\left(\boldsymbol{Q}_j^r + \boldsymbol{Q}_j^{r+1}, \boldsymbol{T}_{r+1/2}\right) \\
\hat{M}_{i,j+N} &= \hat{M}_{i,j+N}^{\text{poly}} + \delta_{ij}\frac{\Delta t}{\varepsilon|\Omega_i|}\sum_{r/\boldsymbol{M}_{r+1/2}\in\Omega_i}\mathcal{N}\boldsymbol{n}_{ir+1/2}(1) + \frac{\Delta t}{\varepsilon|\Omega_i|}\sum_{r/\boldsymbol{M}_{r+1/2}\in\Omega_i\cap\Omega_j}\mathcal{N}\boldsymbol{n}_{ir+1/2}(1) \\
&\quad + \frac{\Delta t}{\varepsilon|\Omega_i|}\sum_{r/\boldsymbol{M}_{r+1/2}\Omega_i\cap\Omega_j}\frac{\mathcal{T}}{2}\left(\mathcal{X}_j^r|_0 + \mathcal{X}_j^{r+1}|_0, \boldsymbol{T}_{r+1/2}\right), \\
\hat{M}_{i,j+2N} &= \hat{M}_{i,j+2N}^{\text{poly}} + \delta_{ij}\frac{\Delta t}{\varepsilon|\Omega_i|}\sum_{r/\boldsymbol{M}_{r+1/2}\in\Omega_i}\mathcal{N}\boldsymbol{n}_{ir+1/2}(2) + \frac{\Delta t}{\varepsilon|\Omega_i|}\sum_{r,\boldsymbol{M}_{r+1/2}\in\Omega_i\cap\Omega_j}\mathcal{N}\boldsymbol{n}_{ir+1/2}(2) \\
&\quad + \frac{\Delta t}{\varepsilon|\Omega_i|}\sum_{r/\boldsymbol{M}_{r+1/2}\in\Omega_i\cap\Omega_j}\frac{\mathcal{T}}{2}\left(\mathcal{X}_j^r|_2 + \mathcal{X}_j^{r+1}|_2, \boldsymbol{T}_{r+1/2}\right) \\
\hat{M}_{i+N,j} &= \hat{M}_{i+N,j}^{\text{poly}} + \delta_{ij}\frac{\Delta t}{\varepsilon|\Omega_i|}\sum_{r/\boldsymbol{M}_{r+1/2}\in\Omega_i}\left(1-\tilde{\mathcal{N}}\right)\boldsymbol{C}_i^{r+1/2}(1) + \frac{\Delta t}{\varepsilon|\Omega_i|}\sum_{r/\boldsymbol{M}_{r+1/2}\in\Omega_i\cap\Omega_j}\tilde{\mathcal{N}}\boldsymbol{C}_i^{r+1/2}(1) \\
&\quad - \frac{\Delta t}{\varepsilon|\Omega_i|}\sum_{r/\boldsymbol{M}_{r+1/2}\in\Omega_i\cap\Omega_j}\frac{\tilde{\mathcal{T}}}{2}\left(\boldsymbol{Q}_j^r + \boldsymbol{Q}_j^{r+1}, \boldsymbol{T}_{r+1/2}\right)\boldsymbol{C}_i^{r+1/2}(1), \\
\hat{M}_{i+N,j+N} &= \hat{M}_{i+N,j+N}^{\text{poly}} + \delta_{ij}\frac{\Delta t}{\varepsilon|\Omega_i|}\sum_{r/\boldsymbol{M}_{r+1/2}\in\Omega_i}\left(1-\tilde{\mathcal{N}}\right)\boldsymbol{n}_{ir+1/2}(1)\boldsymbol{C}_i^{r+1/2}(1) \\
&\quad - \frac{\Delta t}{\varepsilon|\Omega_i|}\sum_{r/\boldsymbol{M}_{r+1/2}\in\Omega_i\cap\Omega_j}\tilde{\mathcal{N}}\boldsymbol{n}_{ir+1/2}(1)\boldsymbol{C}_i^{r+1/2}(1)
\end{aligned}$$

$$
\begin{aligned}
& -\frac{\Delta t}{\varepsilon|\Omega_i|}\sum_{r/\boldsymbol{M}_{r+1/2}\in\Omega_i\cap\Omega_j}\frac{\tilde{\mathcal{T}}}{2}\left(\mathcal{X}_j^r|_0+\mathcal{X}_j^{r+1}|_0,\boldsymbol{T}_{r+1/2}\right)\boldsymbol{C}_i^{r+1/2}(1)\\
\hat{M}_{i+N,j+2N} &= \hat{M}^{\text{poly}}_{i+N,j+2N}+\delta_{ij}\frac{\Delta t}{\varepsilon|\Omega_i|}\sum_{r/\boldsymbol{M}_{r+1/2}\in\Omega_i}\left(1-\tilde{\mathcal{N}}\right)\boldsymbol{n}_{ir+1/2}(2)\boldsymbol{C}_i^{r+1/2}(1)\\
& -\frac{\Delta t}{\varepsilon|\Omega_i|}\sum_{r/\boldsymbol{M}_{r+1/2}\in\Omega_i\cap\Omega_j}\tilde{\mathcal{N}}\boldsymbol{n}_{ir+1/2}(2)\boldsymbol{C}_i^{r+1/2}(1)\\
& -\frac{\Delta t}{\varepsilon|\Omega_i|}\sum_{r/\boldsymbol{M}_{r+1/2}\in\Omega_i\cap\Omega_j}\frac{\tilde{\mathcal{T}}}{2}\left(\mathcal{X}_j^r|_2+\mathcal{X}_j^{r+1}|_2,\boldsymbol{T}_{r+1/2}\right)\boldsymbol{C}_i^{r+1/2}(1),\\
\hat{M}_{i+2N,j} &= \hat{M}^{\text{poly}}_{i+2N,j}+\delta_{ij}\frac{\Delta t}{\varepsilon|\Omega_i|}\sum_{r/\boldsymbol{M}_{r+1/2}\in\Omega_i}\left(1-\tilde{\mathcal{N}}\right)\boldsymbol{C}_i^{r+1/2}(2)+\frac{\Delta t}{\varepsilon|\Omega_i|}\sum_{r/\boldsymbol{M}_{r+1/2}\in\Omega_i\cap\Omega_j}\tilde{\mathcal{N}}\boldsymbol{C}_i^{r+1/2}(2)\\
& -\frac{\Delta t}{\varepsilon|\Omega_i|}\sum_{r/\boldsymbol{M}_{r+1/2}\in\Omega_i\cap\Omega_j}\frac{\tilde{\mathcal{T}}}{2}\left(\boldsymbol{Q}_j^r+\boldsymbol{Q}_j^{r+1},\boldsymbol{T}_{r+1/2}\right)\boldsymbol{C}_i^{r+1/2}(2),\\
\hat{M}_{i+2N,j+N} &= \hat{M}^{\text{poly}}_{i+2N,j+N}+\delta_{ij}\frac{\Delta t}{\varepsilon|\Omega_i|}\sum_{r/\boldsymbol{M}_{r+1/2}\in\Omega_i}\left(1-\tilde{\mathcal{N}}\right)\boldsymbol{n}_{ir+1/2}(1)\boldsymbol{C}_i^{r+1/2}(2)\\
& -\frac{\Delta t}{\varepsilon|\Omega_i|}\sum_{r/\boldsymbol{M}_{r+1/2}\in\Omega_i\cap\Omega_j}\tilde{\mathcal{N}}\boldsymbol{n}_{ir+1/2}(1)\boldsymbol{C}_i^{r+1/2}(2)\\
& -\frac{\Delta t}{\varepsilon|\Omega_i|}\sum_{r/\boldsymbol{M}_{r+1/2}\in\Omega_i\cap\Omega_j}\frac{\tilde{\mathcal{T}}}{2}\left(\mathcal{X}_j^r|_1+\mathcal{X}_j^{r+1}|_1,\boldsymbol{T}_{r+1/2}\right)\boldsymbol{C}_i^{r+1/2}(2)\\
\hat{M}_{i+2N,j+2N} &= \hat{M}^{\text{poly}}_{i+2N,j+2N}+\delta_{ij}\frac{\Delta t}{\varepsilon|\Omega_i|}\sum_{r/\boldsymbol{M}_{r+1/2}\in\Omega_i}\left(1-\tilde{\mathcal{N}}\right)\boldsymbol{n}_{ir+1/2}(2)\boldsymbol{C}_i^{r+1/2}(2)\\
& -\frac{\Delta t}{\varepsilon|\Omega_i|}\sum_{r/\boldsymbol{M}_{r+1/2}\in\Omega_i\cap\Omega_j}\tilde{\mathcal{N}}\boldsymbol{n}_{ir+1/2}(2)\boldsymbol{C}_i^{r+1/2}(2)\\
& -\frac{\Delta t}{\varepsilon|\Omega_i|}\sum_{r+1/2}\frac{\tilde{\mathcal{T}}}{2}\left(\mathcal{X}_j^r|_2+\mathcal{X}_j^{r+1}|_2,\boldsymbol{T}_{r+1/2}\right)\boldsymbol{C}_i^{r+1/2}(2),
\end{aligned}
$$

As before, for a vector $\boldsymbol{C}\in\mathbb{R}^2$, $\boldsymbol{C}(1)$ denotes its first component, while $\boldsymbol{C}(2)$ denotes its second component. For a matrix $\mathcal{X}\in\mathbb{R}^{2\times 2}$, $\mathcal{X}_{|1}$ denotes its first column, while $\mathcal{X}_{|2}$ denotes its second column.

Note that, once the JLa scheme is implemented, JLb is only a slight modification: changing $\mathcal{R}$ on the one hand, and changing $\boldsymbol{F}_i^{n+1}$ in the second equation on the other hand.

4.3 Numerical tests

We present in this section some numerical tests for the schemes presented above. The two variants JLa and JLb give more or less the same results, so we do not make any difference between them.

We first present results for a small parameter ε. This allows to assess the fact that the scheme is asymptotic preserving (AP) in the diffusion limit. We also present results with larger values of ε, for which explicit solutions allow to assess the convergence of the scheme.

4.3.1 Diffusion limit $\varepsilon = 10^{-4}$

Here, we use a cartesian mesh on the domain $\Omega=[0;4]^2$, with 31 cells per direction. For each edge, we add a control point in the normal direction, at a distance equal to 20% of the edge length, towards the origin (see Figure 19).

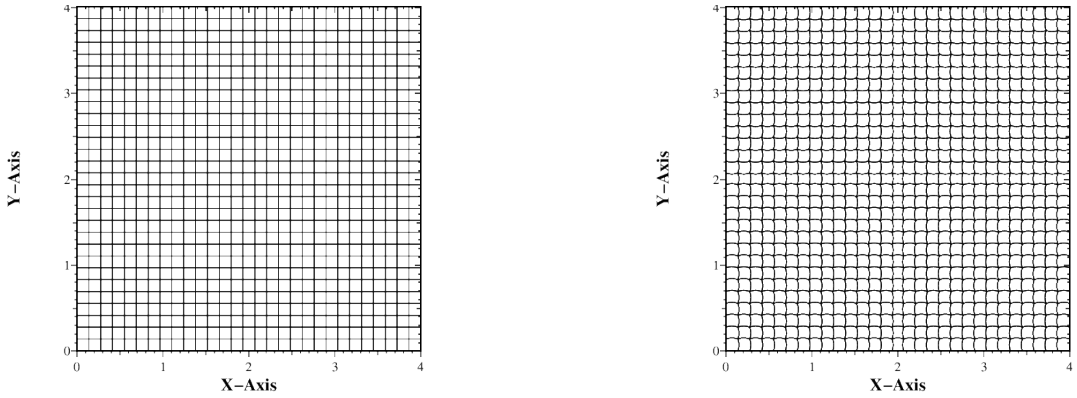

Figure 19: Left: cartesian mesh. Right: conical mesh with $\omega = 1$.

The initial condition is a Dirac mass, that is,

$$\begin{cases} E_j^0 = \frac{1}{|\Omega_j|} \text{ if } (2,2) \in \Omega_j \\ E_j^0 = 0 \text{ otherwise,} \end{cases}$$

and $\boldsymbol{F} = 0$. Thus, only the central cell contains some energy, and the total integral of E^0 is equal to 1. We set $\varepsilon = 10^{-4}$ and $\sigma = 1$, and the time step is $\Delta t = 0.003$. The results are displayed in Figure 20 at time $t = 0.03$.

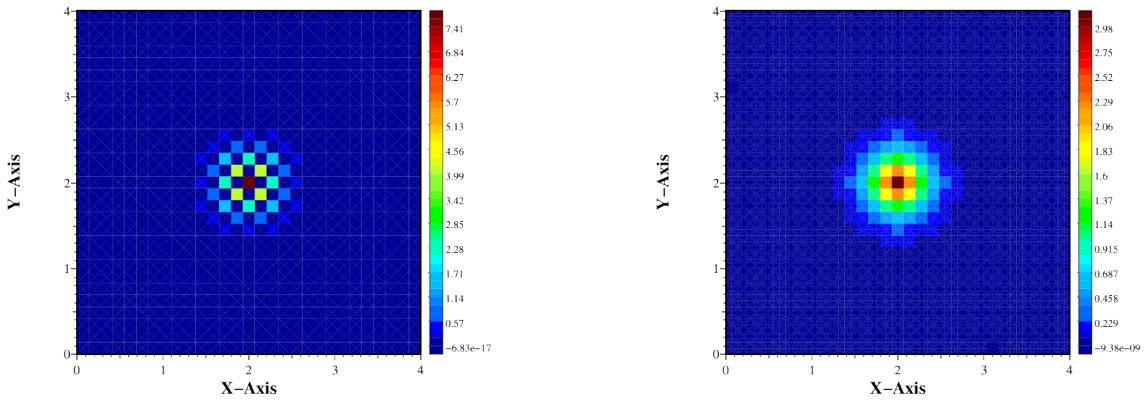

Figure 20: Results for a Dirac initial condition. Left: polygonal mesh. Right: conical mesh with $\omega = 0$

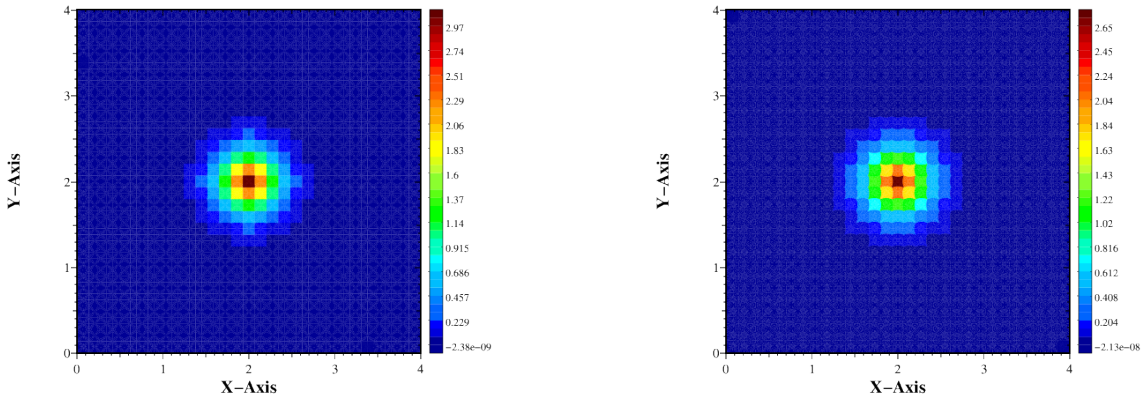

Figure 21: Left: conical mesh with $\omega = 0.01$. Right: conical mesh with $\omega = 1$.

We see that the scheme on polygonal mesh suffers from cross-stencil. We also plot in Figure 21 the same result for a different value of ω. Recall that in the case $\omega = 0$, we do not recover the scheme on a polygonal

mesh, although the mesh *is* polygonal. Indeed, the contribution of the shoulder points is still present and in some sense sovles the cross-stencil issue.

4.3.2 Transport case $\varepsilon = 1$

We now present the same kind of results for $\varepsilon = 1$. All other parameters of the simulation are unchanged. In particular, we use meshes displayed in Figure 19.

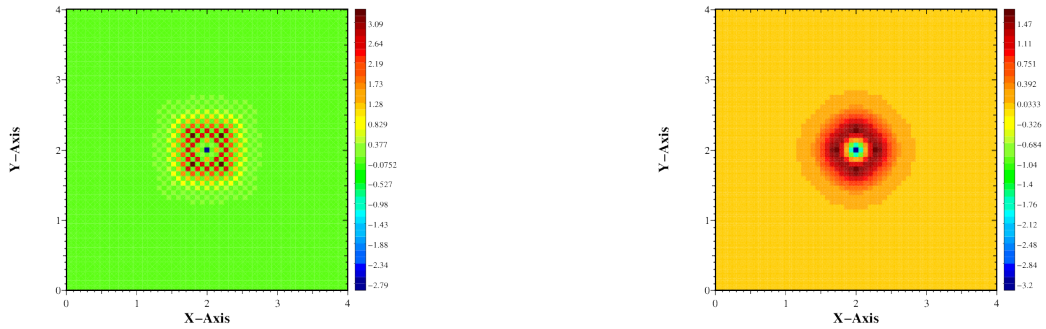

Figure 22: Left: polygonal mesh. Right: conical mesh with $\omega = 0$.

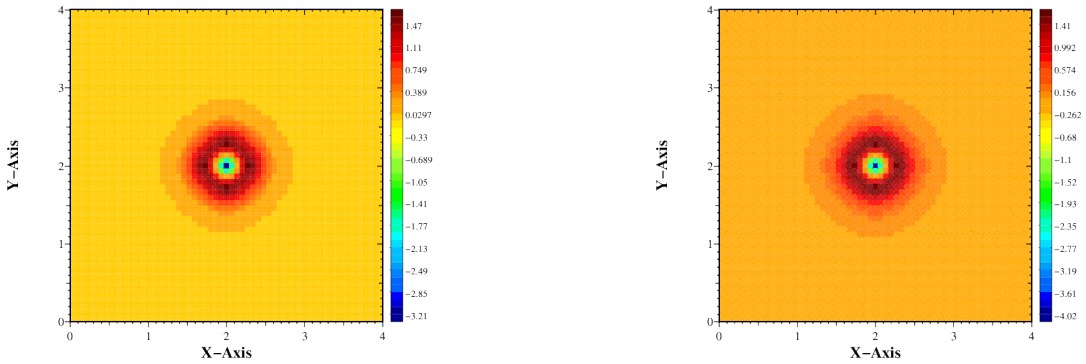

Figure 23: Left: conical mesh with $\omega = 0.01$. Right: conical mesh with $\omega = 1$.

Results are displayed in Figures 22 and 23. As in the diffusion limit case, we still observe the cross-stencil issue, with the same conclusion.

4.3.3 Radial mesh in the diffusion limit $\varepsilon = 10^{-4}$

We now go back to the diffusion limit case, but use different meshes, which are displayed in Figure 24. The weights are all set to $\omega = \frac{\sqrt{2}}{2}$, so that edges are portion of circles centered at the origin. We set $\varepsilon = 10^{-4}$ and $\sigma = 1$, and the time step is $\Delta t = 0.003$. The results are displayed in Figure 25 and 26 at time $t = 0.03$.

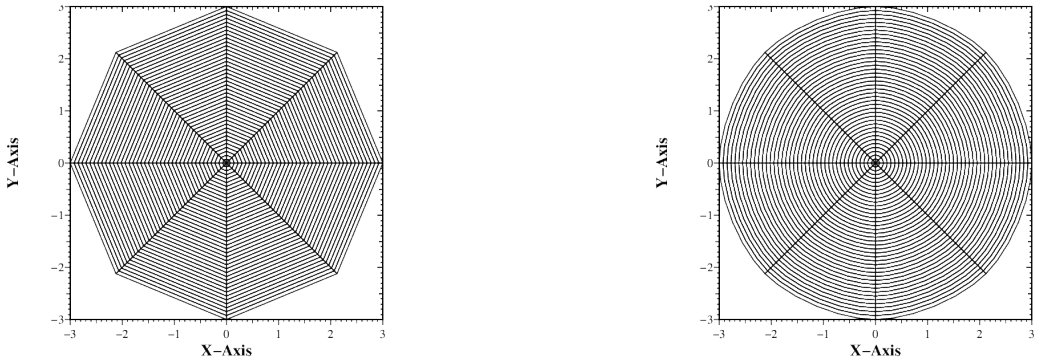

Figure 24: Left: polygonal mesh. Right: conical mesh with $\omega = \frac{\sqrt{2}}{2}$

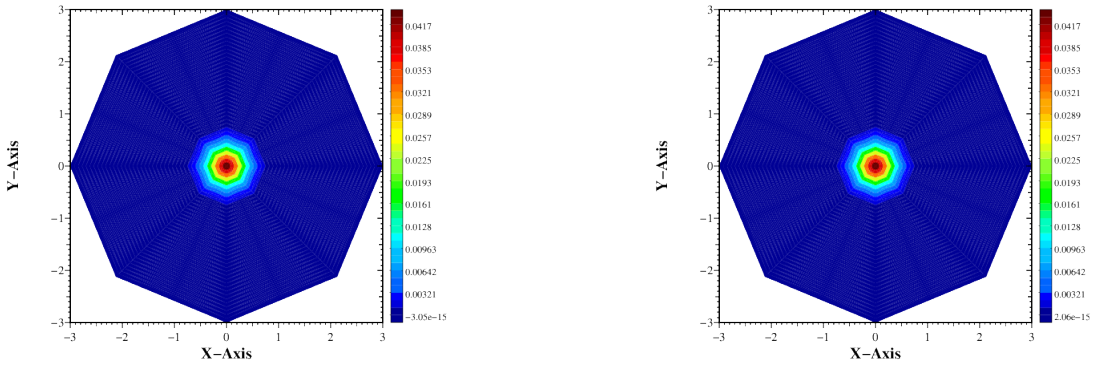

Figure 25: Left: polygonal mesh. Right: conical mesh with $\omega = 0$

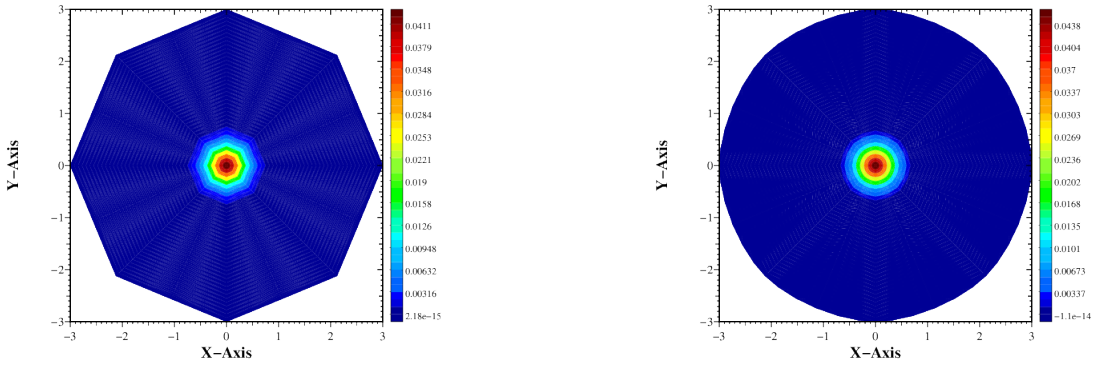

Figure 26: Left: conical mesh with $\omega = 0.01$. Right: conical mesh with $\omega = \frac{\sqrt{2}}{2}$.

4.3.4 Radial mesh in transport case $\varepsilon = 1$

This test case is exactly the same as the preceding one, except that we set $\varepsilon = 1$. Results may be observed in Figures 27.

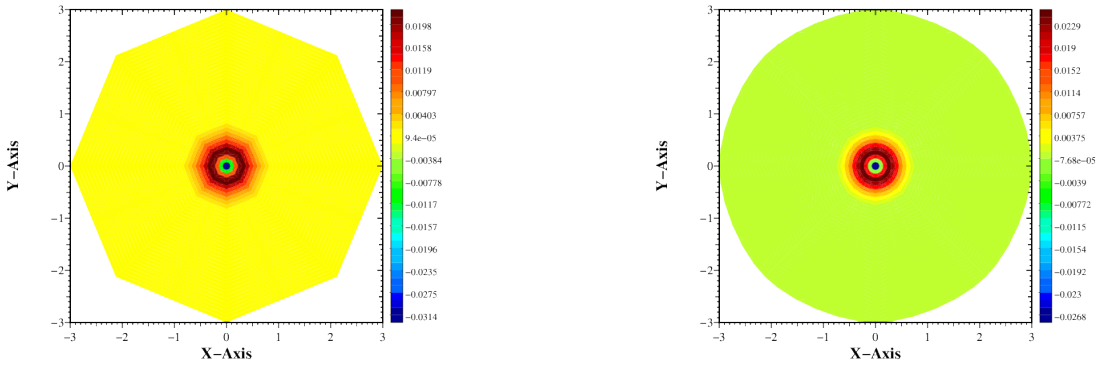

Figure 27: Left: polygonal mesh. Right: conical mesh with $\omega = \frac{\sqrt{2}}{2}$.

4.3.5 Triangular mesh in the diffusion limit $\varepsilon = 10^{-4}$

We now test the scheme on a triangular mesh. The domain is $\Omega = [-1; 1]^2$. The original triangular mesh is displayed in Figure 28 (left). For the conical mesh, we add a control point for each edge, at a distance equal to 20% of the edge length along the normal. This gives the mesh displayed in Figure 28 (right).

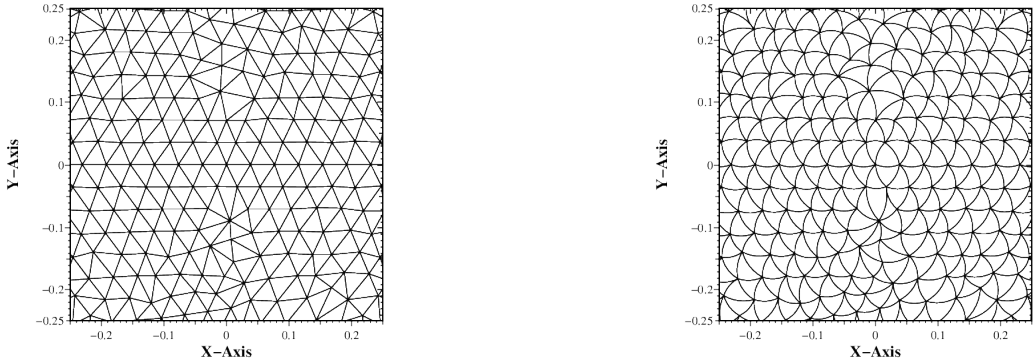

Figure 28: Left: polygonal mesh. Right: conical mesh with $\omega = 1$

Here again, $\sigma = 1$, $\varepsilon = 10^{-4}$, and the initial data is a Dirac mass at the origin. The time step is $\Delta t = 0.001$, and the results ar presented at $t = 0.014$ (Figures 29 and 30).

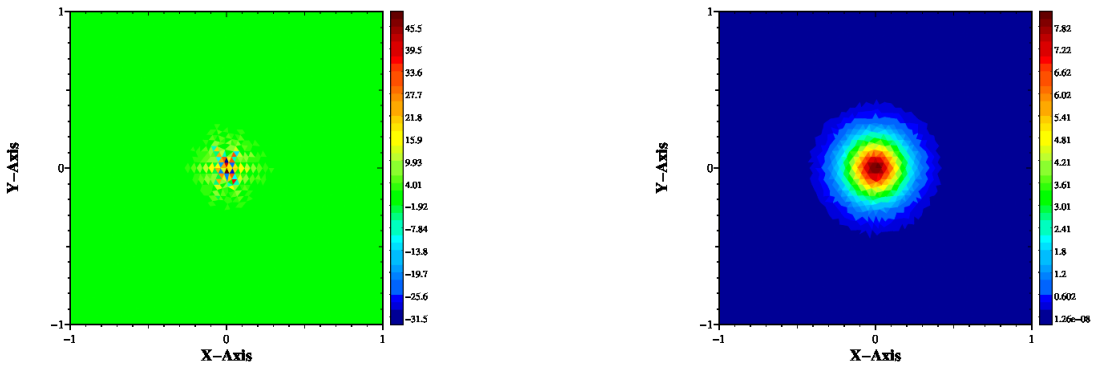

Figure 29: Left: polygonal mesh. Right: conical mesh with $\omega = 0$

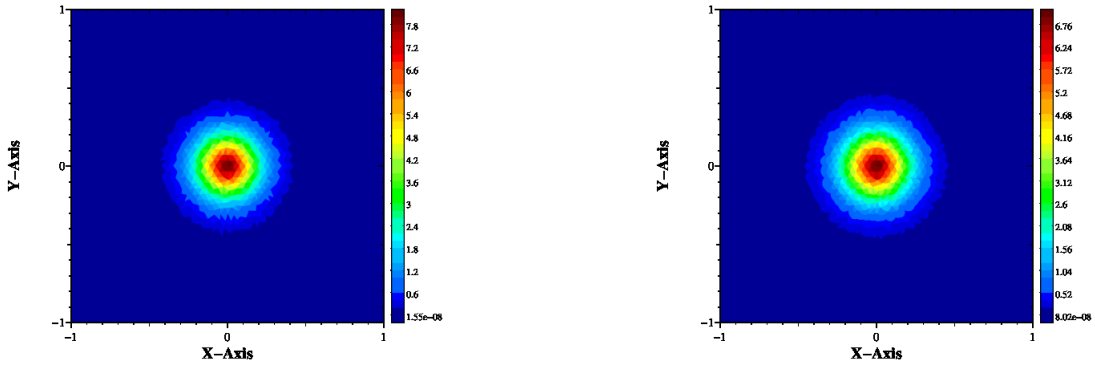

Figure 30: Left: conical mesh with $\omega = 0.01$. Right: conical mesh with $\omega = 1$.

As in the case of the diffusion scheme, the original scheme on polygonal meshes has positivity issues. This is not the case for the scheme on conical meshes, even in the limit $\omega = 0$. This limit corresponds to a polygonal mesh, but with a different scheme (see Remark 2.4).

4.3.6 Triangular mesh in the transport case $\varepsilon = 1$

This test is exactly the same as the preceding one, except that $\varepsilon = 1$. Results are displayed in Figures 31 and 32.

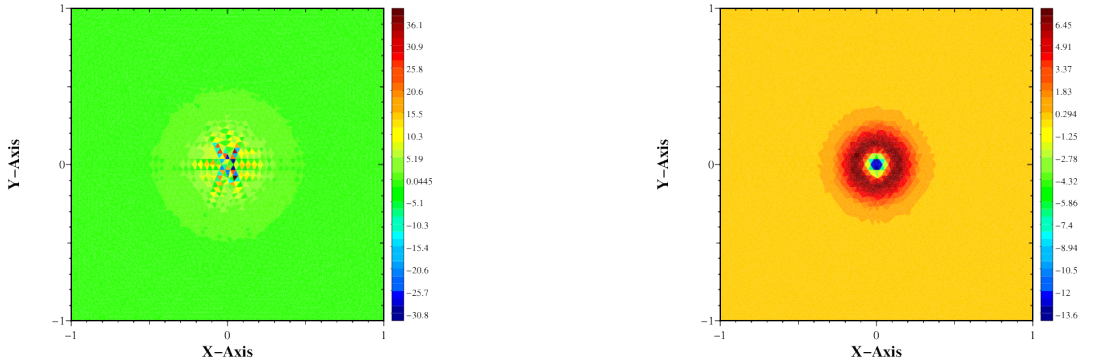

Figure 31: Left: polygonal mesh. Right: conical mesh with $\omega = 0$

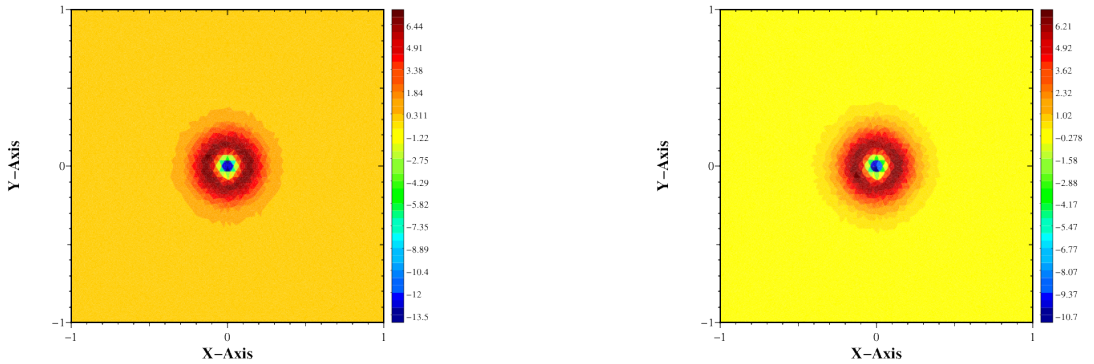

Figure 32: Left: conical mesh with $\omega = 0.01$. Right: conical mesh with $\omega = 1$

Here again, the cross-stencil issue is observed for the polygonal case. It is not present in the case of conical meshes.

4.3.7 Convergence study

Here we propose a convergence analysis of the scheme. For this purpose, we need an explicit solution. Using periodic boundary condition, and assuming that σ is constant, a natural way to compute solution is to use a plane wave decomposition, that is,

$$E(x,y,t) = \sum_{j\in\mathbb{Z}}\sum_{k\in\mathbb{Z}} \alpha_{j,j}(t) e^{2i\pi\frac{jx+ky}{L}}, \tag{4.28}$$

where L is the period (that is, the size of the domain). We also point out a simple computation lead from (1.1) to the damped wave equation for E:

$$\frac{\partial^2 E}{\partial t^2} + \frac{\sigma}{\varepsilon^2}\frac{\partial E}{\partial t} - \frac{1}{\varepsilon^2}\Delta E = 0, \tag{4.29}$$

where we explicitly used that σ is constant. Inserting (4.28) into (4.29), we find a linear differential equation solved by each $\alpha_{j,k}$. Solving it (we skip the computations, which are tedious but not difficult), we end up with the following solutions:

If $4\pi\sqrt{j^2+k^2} \leq \frac{\sigma L}{\varepsilon}$,

$$E(x,y,t) = \cos\left(\frac{2j\pi x}{L} + \frac{2k\pi y}{L}\right) e^{-\frac{\sigma t}{2\varepsilon^2}} \left[\alpha\left(\frac{\sigma}{\varepsilon^2}\frac{\sinh(\gamma t)}{2\gamma} + \cosh(\gamma t)\right) + \beta\frac{\sinh(\gamma t)}{\gamma}\right], \tag{4.30}$$

$$F(x,y,t) = \frac{\varepsilon L}{2\pi(j^2+k^2)}\sin\left(\frac{2j\pi x}{L} + \frac{2k\pi y}{L}\right) e^{-\frac{\sigma t}{2\varepsilon^2}} \left[\alpha\frac{4\pi^2(j^2+k^2)}{L^2\varepsilon^2}\frac{\sinh(\gamma t)}{\gamma} + \beta\left(\frac{\sigma}{2\varepsilon^2\gamma}\sinh(\gamma t) - \cosh(\gamma t)\right)\right]\begin{pmatrix} j \\ k\end{pmatrix},$$

where $\gamma = \frac{1}{2}\sqrt{\frac{\sigma^2}{\varepsilon^4} - 16\frac{\pi^2}{L^2\varepsilon^2}(j^2+k^2)}$, and the initial data are given by

$$E(x,y,0) = \alpha\cos\left(\frac{2j\pi x}{L} + \frac{2k\pi y}{L}\right), \quad F(x,y,0) = -\beta\frac{\varepsilon L}{2\pi(j^2+k^2)}\sin\left(\frac{2j\pi x}{L} + \frac{2k\pi y}{L}\right)\begin{pmatrix} j \\ k\end{pmatrix}.$$

If $4\pi\sqrt{j^2+k^2} > \frac{\sigma L}{\varepsilon}$,

$$E(x,y,t) = \cos\left(\frac{2j\pi x}{L} + \frac{2k\pi y}{L}\right) e^{-\frac{\sigma t}{2\varepsilon^2}} \left[\alpha\left(\frac{\sigma}{\varepsilon^2}\frac{\sin(\gamma t)}{2\gamma} + \cos(\gamma t)\right) + \beta\frac{\sin(\gamma t)}{\gamma}\right], \tag{4.31}$$

$$F(x,y,t) = \frac{\varepsilon L}{2\pi(j^2+k^2)}\sin\left(\frac{2j\pi x}{L} + \frac{2k\pi y}{L}\right) e^{-\frac{\sigma t}{2\varepsilon^2}} \left[\alpha\frac{4\pi^2(j^2+k^2)}{L^2\varepsilon^2}\frac{\sin(\gamma t)}{\gamma} + \beta\left(\frac{\sigma}{2\varepsilon^2\gamma}\sin(\gamma t) - \cos(\gamma t)\right)\right]\begin{pmatrix} j \\ k\end{pmatrix},$$

where $\gamma = \frac{1}{2}\sqrt{16\frac{\pi^2}{L^2\varepsilon^2}(j^2+k^2) - \frac{\sigma^2}{\varepsilon^4}}$, and the initial data are given by

$$E(x,y,0) = \alpha\cos\left(\frac{2j\pi x}{L} + \frac{2k\pi y}{L}\right), \quad F(x,y,0) = -\beta\frac{\varepsilon L}{2\pi(j^2+k^2)}\sin\left(\frac{2j\pi x}{L} + \frac{2k\pi y}{L}\right)\begin{pmatrix} j \\ k\end{pmatrix}.$$

In what follows, we use the following parameters: $\sigma = 1$, $\varepsilon = 1$, $L = 20$, $j = 1$, $k = 2$, $\alpha = 1$, $\beta = 0$. Hence, we are in the case (4.31). We add 1 to the energy in order to avoid negative non-physical energies. Hence, the energy reads

$$E(x,y,t) = 1 + \cos\left(\frac{2j\pi x}{L} + \frac{2k\pi y}{L}\right) e^{-\frac{\sigma t}{2\varepsilon^2}} \left(\frac{\sigma}{\varepsilon^2}\frac{\sin(\gamma t)}{2\gamma} + \cos(\gamma t)\right), \tag{4.32}$$

which is plotted in Figure 33

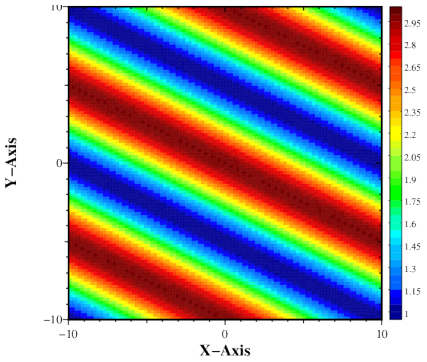

Figure 33: The exact solution given by (4.32).

We define the meshes as follows: first, we use a uniform grid if size $N_x \times N_y$, with $N_x = N_y$ ranging from 50 to 1000. This mesh is then made conical in three different ways:

1. For each edge, we add a control point at a distance equal to 20% of the edge length, in the direction of the normal going away from the origin. This mesh is shown in Figure 34 (upper right).
2. For each edge, we add a control point at a distance equal to 20% of the edge length, in the direction of the normal going toward the origin. This mesh is shown in Figure 34 (lower left).
3. The same as above, but with the direction chosen randomly outward or inward. This mesh is shwon in Figure 34 (lower right).

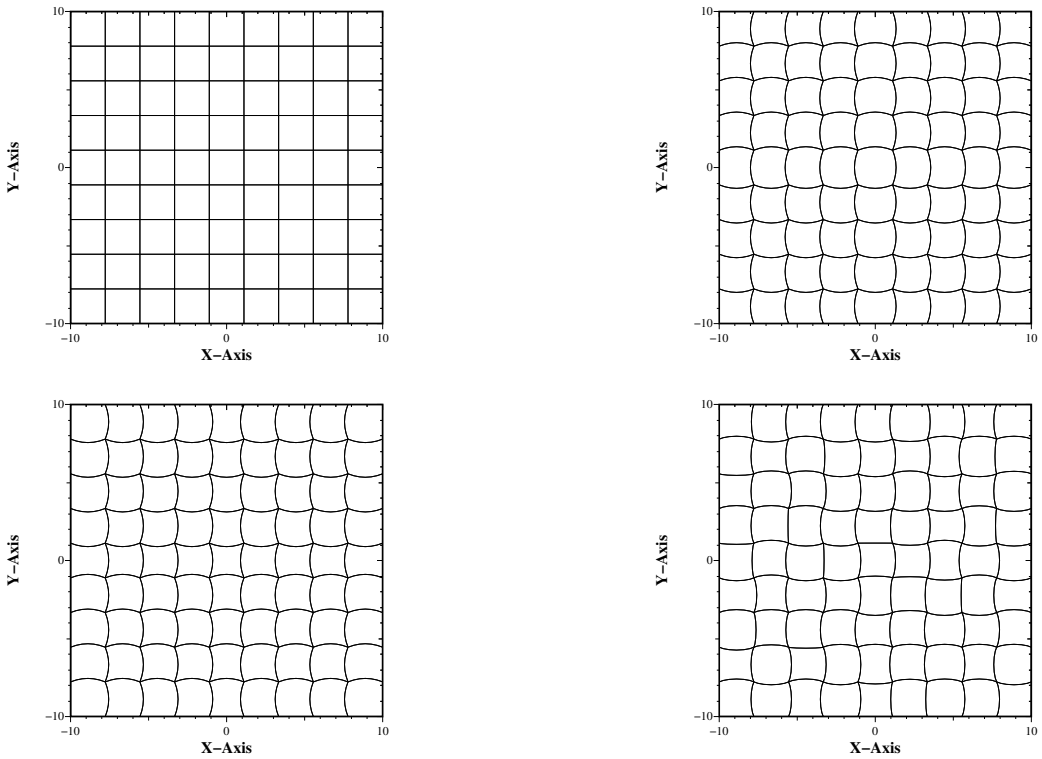

Figure 34: The different meshes used for the convergence analysis: upper left cartesian mesh, upper right conical mesh with control points away from the origin, lower left conical mesh with control points closer to the origin, lower right random control points.

We fix $\Delta t = 2.5 \times 10^{-5}$, so that the time discretization error is much smaller than the spacial discretization error. We compute the relative L^2 error on the energy at the final time $T = 1$, as a function of the number

of cells per direction. As shown in Figure 35, we observe convergence of order one, as expected. The error is almost the same for all conical meshes. It is slightly larger for the polygonal mesh.

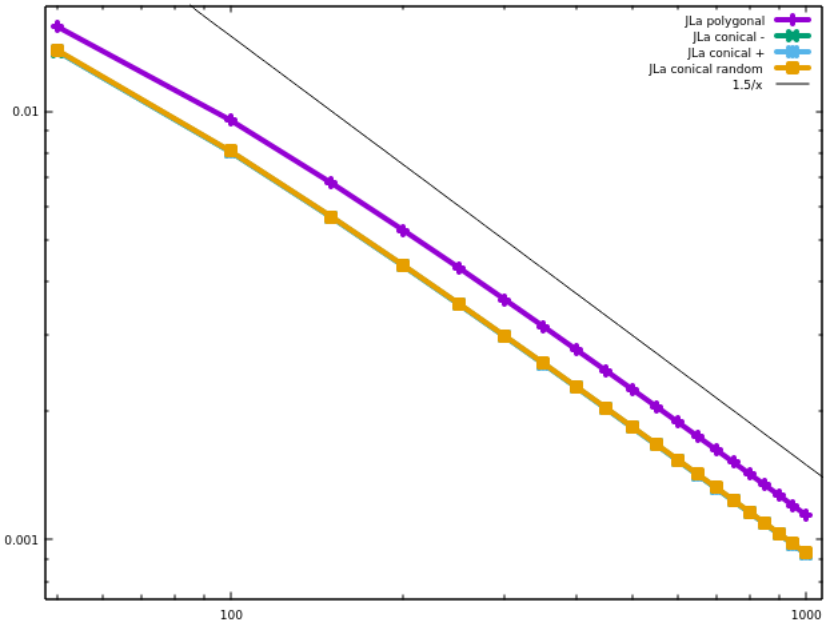

Figure 35: Error on the energy as a function of then number or cells per direction. The exact solution is given by (4.32), with $\varepsilon = \sigma = 1$ and $j = 1, k = 2$.

We also performed similar tests in the diffusion limit: first with $\varepsilon = 10^{-4}$, then $\varepsilon = 10^{-8}$. In each case, we used the above-mentionned exact solution with $\sigma = 1$ and $j = 1$, $k = 2$. The time step is set to $\Delta t = 10^{-4}$, and, due to the stiffness of the system, we use an implicit scheme. As can be seen on Figure 36, the error is second-order with respect to Δx. This is consistent with the fact that the limit diffusion scheme is second-order convergent.

4.3.8 A periodic test case with holes

In this last test, we have used the geometry displayed in Figure 37.

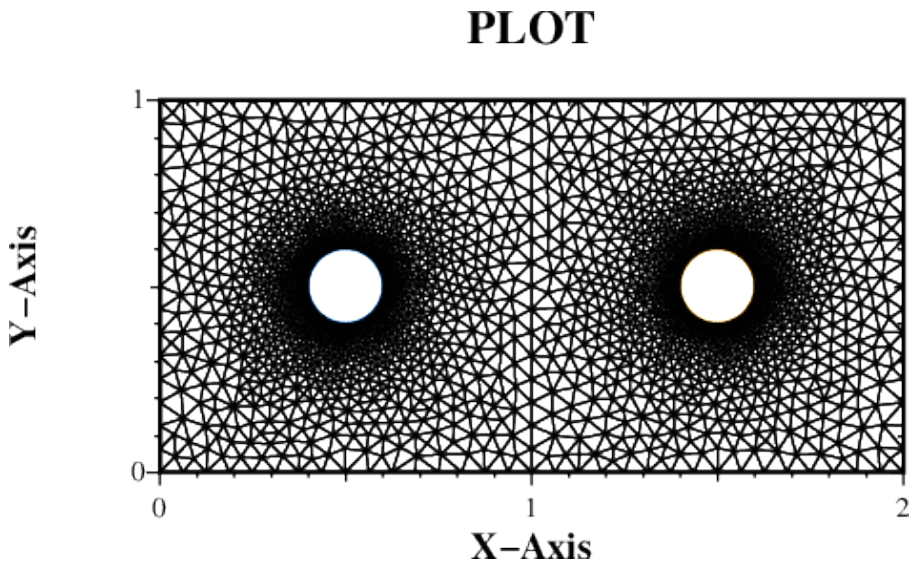

Figure 37: Geometry and mesh of the test case presented in Subsection 4.3.8

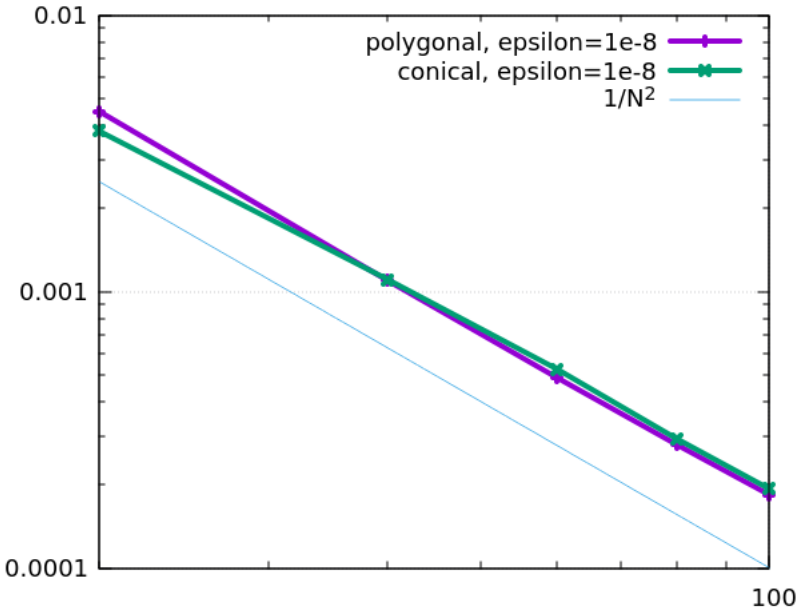

Figure 36: Error on the energy as a function of the number or cells per direction, in the diffusion limit case $\varepsilon = 10^{-8}$. The exact solution is given by (4.32), with $\sigma = 1$ and $j = 1, k = 2$.

We solve problem (1.2) with $\sigma = 1$, with periodic boundary conditions on the outer boundaries. On the boundary of the holes, homogeneous Neumann boundary conditions are used. The initial data is a Dirac mass at point $(0.7, 0.7)$, that is, close to the left-hand side hole, above it and on its right. When using a polygonal mesh, the code crashes due to (unphysical) negative values and numerical instabilities. However, using an exact description of the hole boundaries with conical edges allows for a more stable scheme, with the solution displayed on Figure 38.

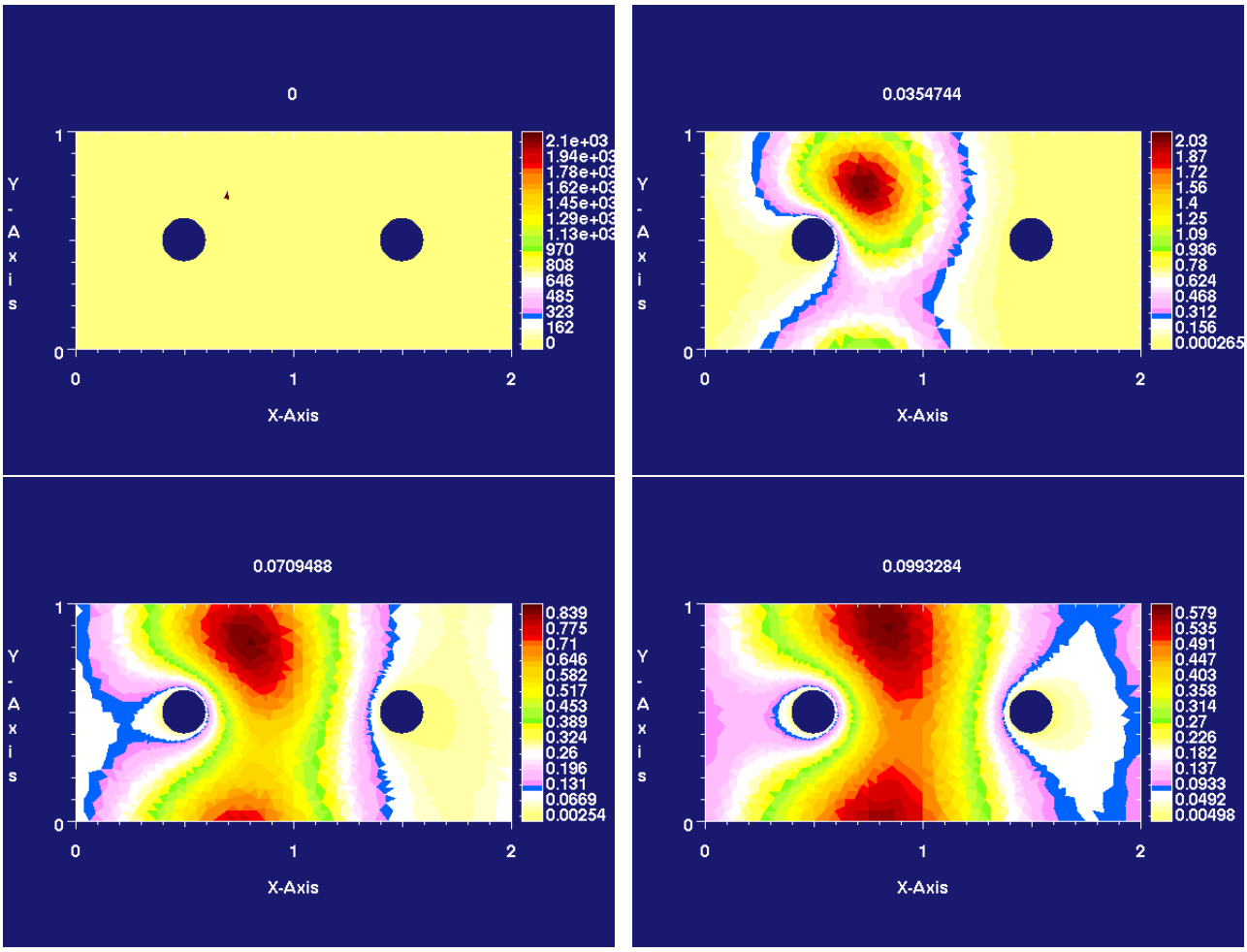

Figure 38: Solution of problem (1.2) with a Dirac mass as initial data

A similar study has been carried out in the case of (1.1), with $\sigma = \varepsilon = 1$. Here, we have no maximum principle and it is expected that the energy becomes negative [18]. However, we still have a much nicer behavior when the holes are described by exact (conical) edges. See Figure 39 for the polygonal case, and 40 for the conical case.

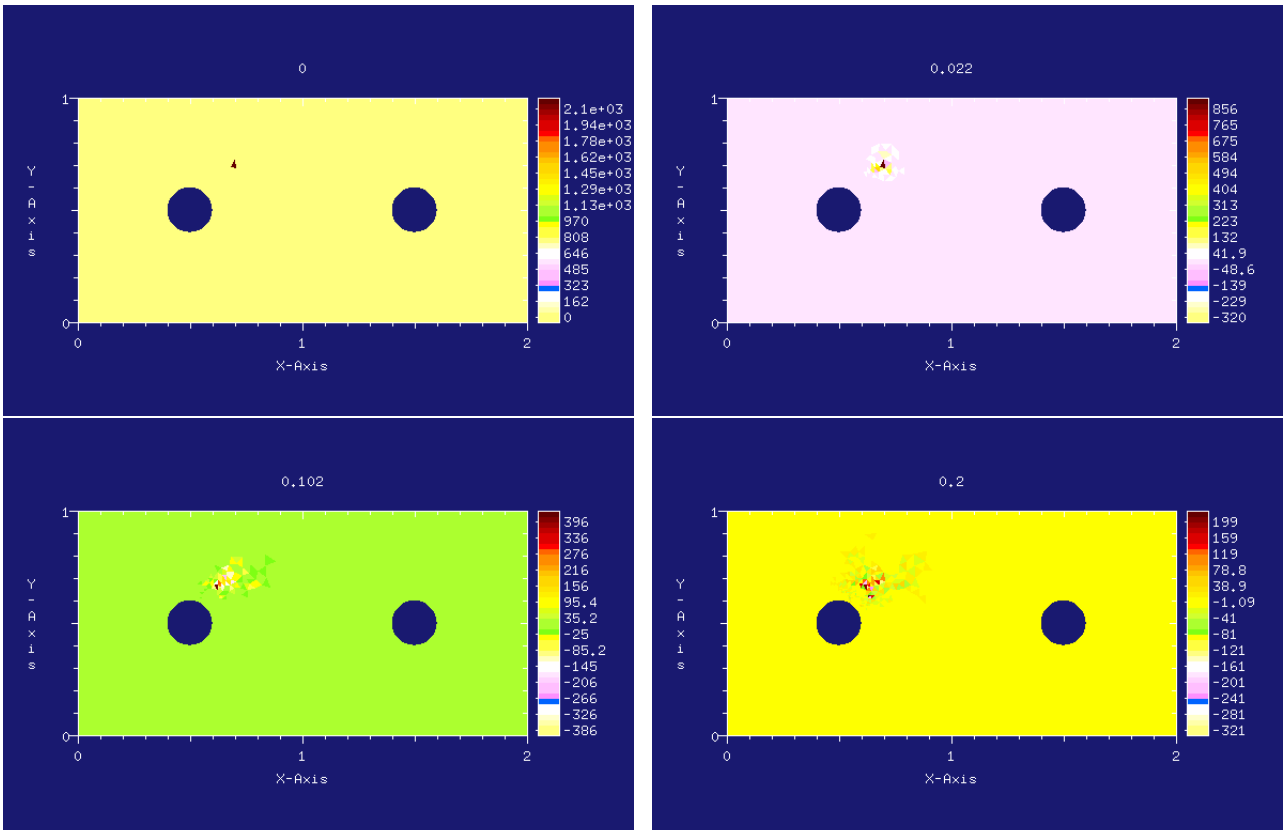

Figure 39: Solution of problem (1.1) with $\varepsilon = \sigma = 1$, with a Dirac mass as initial data, on a polygonal mesh.

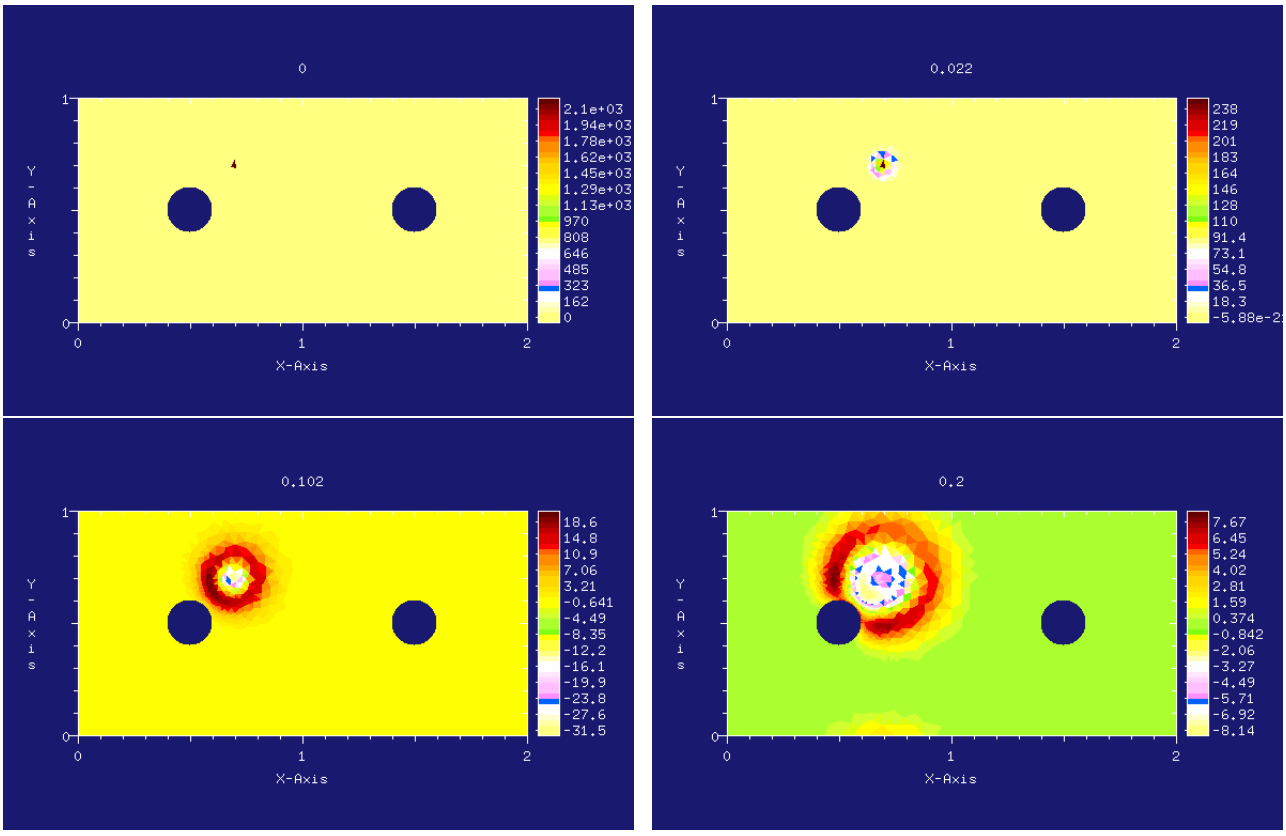

Figure 40: Solution of problem (1.1) with $\varepsilon = \sigma = 1$, with a Dirac mass as initial data, on a conical mesh with an exact description of the boundary edges on the holes.

References

[1] Aude Bernard-Champmartin, Erwan Deriaz, Philippe Hoch, Gerald Samba, and Michael Schaefer. Extension of centered hydrodynamical schemes to unstructured deforming conical meshes : the case of circles. In *ESAIM : proc.*, volume 135-162, 2012.

[2] Benjamin Boutin, Erwan Deriaz, Philippe Hoch, and Pierre Navaro. Extension of ale methodology to unstructured conical meshes. In *ESAIM: Proc.*, volume 32, pages 31–55, 2011.

[3] Christophe Buet, Bruno Després, and Emmanuel Franck. Design of asymptotic preserving finite volume schemes for the hyperbolic heat equation on unstructured meshes. *Numer. Math.*, 122(2):227–278, 2012.

[4] Christophe Buet, Bruno Després, and Emmanuel Franck. Asymptotic preserving schemes on distorted meshes for Friedrichs systems with stiff relaxation: application to angular models in linear transport. *J. Sci. Comput.*, 62(2):371–398, 2015.

[5] Christophe Buet, Bruno Després, Emmanuel Franck, and Thomas Leroy. Proof of uniform convergence for a cell-centered AP discretization of the hyperbolic heat equation on general meshes. *Math. Comp.*, 86(305):1147–1202, 2017.

[6] G. Carré, S. Del Pino, B. Després, and E. Labourasse. A cell-centered Lagrangian hydrodynamics scheme on general unstructured meshes in arbitrary dimension. *J. Comput. Phys.*, 228(14):5160–5183, 2009.

[7] John Castor. *Radiation Hydrodynamics.* Cambridge University Press, 2004.

[8] S. Chandrasekhar. *Radiative transfer.* Dover Publications, Inc., New York, 1950.

[9] Bruno Després. Weak consistency of the cell-centered Lagrangian GLACE scheme on general meshes in any dimension. *Comput. Methods Appl. Mech. Eng.*, 199(41-44):2669–2679, 2010.

[10] Bruno Després and Christophe Buet. The structure of well-balanced schemes for Friedrichs systems with linear relaxation. *Appl. Math. Comput.*, 272(part 2):440–459, 2016.

[11] Emmanuel Franck. *Design and numerical analysis of asymptotic preserving schemes on unstructured meshes. Application to the linear transport and Friedrichs systems.* Theses, Université Pierre et Marie Curie - Paris VI, October 2012.

[12] Franck, Emmanuel, Hoch, Philippe, Navaro, Pierre, and Samba, Gérald. An asymptotic preserving scheme for p1 model using classical diffusion schemes on unstructured polygonal meshes. *ESAIM: Proc.*, 32:56–75, 2011.

[13] Laurent Gosse and Giuseppe Toscani. An asymptotic-preserving well-balanced scheme for the hyperbolic heat equations. *C. R. Math. Acad. Sci. Paris*, 334(4):337–342, 2002.

[14] James M. Greenberg and Alain-Yves Le Roux. A well-balanced scheme for the numerical processing of source terms in hyperbolic equations. *SIAM J. Numer. Anal.*, 33(1):1–16, 1996.

[15] Wang Guojin. Computing integral values involving nurbs curves. *Jour. of Software*, 7:542–546, 1998.

[16] Shi Jin and C.David Levermore. Numerical schemes for hyperbolic conservation laws with stiff relaxation terms. *Journal of Computational Physics*, 126(2):449 – 467, 1996.

[17] Constant Mazeran. *Sur la structure mathématique et l'approximation numérique de l'hydrodynamique lagrangienne bidimensionnelle.* PhD thesis, Université de Bordeaux I, 2007.

[18] Ryan G. McClarren, James Paul Holloway, and Thomas A. Brunner. Analytic p1 solutions for time-dependent, thermal radiative transfer in several geometries. *Journal of Quantitative Spectroscopy and Radiative Transfer*, 109(3):389 – 403, 2008.

[19] Dimitri Mihalas and Barbara Weibel Mihalas. *Foundations of radiation hydrodynamics.* Oxford University Press, New York, 1984.

[20] Gerald C. Pomraning. *Linear kinetic theory and particle transport in stochastic mixtures.* Singapore etc.: World Scientific, 1991.