Finite volume schemes for the approximation via characteristics of linear convection equations with irregular data

François Bouchut, Robert Eymard, Alain Prignet

Finite volume schemes for the approximation via characteristics of linear convection equations with irregular data

F. Bouchut, R. Eymard and A. Prignet*

November 18, 2010

Abstract

We consider the approximation by multidimensional finite volume schemes of the transport of an initial measure by a Lipschitz flow. We first consider a scheme defined via characteristics, and we prove the convergence to the continuous solution, as the time-step and the ratio of the space step to the time-step tend to zero. We then consider a second finite volume scheme, obtained from the first one by addition of some uniform numerical viscosity. We prove that this scheme converges to the continuous solution, as the space step tends to zero whereas the ratio of the space step to the time-step remains bounded by below and by above, and under assumption of uniform regularity of the mesh. This is obtained via an improved discrete Sobolev inequality and a sharp weak BV estimate, under some additional assumptions on the transport flow. Examples show the optimality of these assumptions.

Keywords: Approximation of the transport of measures, Lipschitz flows, compression estimates, finite volume schemes, discrete Sobolev inequalities, weak BV estimates.

1 Introduction

We consider the problem of transport of non-negative Radon measures by a flow in $\mathbb{R}^N$. The family of transported non-negative Radon measures $(\mu_t)_{t\in\mathbb{R}_+}$ is defined as

$$\int f(x)\mathrm{d}\mu_s(x) = \int f(X(s,0,x))\mathrm{d}\mu_0(x), \quad f \in C_c(\mathbb{R}^N),\ s \in \mathbb{R}_+, \tag{1}$$

where $C_c(\mathbb{R}^N)$ denotes the space of continuous functions on $\mathbb{R}^N$ with compact support. We make the following natural assumptions.

Hypothesis 1. *We assume that*

1. $N \in \mathbb{N}$ *with* $N \geq 1$,
2. *defining* $D = \{(s,t) \in \mathbb{R}^2, 0 \leq t \leq s\}$, *the flow* $X : D\times\mathbb{R}^N \to \mathbb{R}^N$ *is a globally Lipschitz continuous function with Lipschitz constant* ξ, *such that the flow property holds, i.e.*

$$X(s_3,s_2,X(s_2,s_1,x)) = X(s_3,s_1,x), \quad \text{for all } 0 \leq s_1 \leq s_2 \leq s_3,\ x \in \mathbb{R}^N, \tag{2}$$

$$X(s,s,x) = x, \quad \text{for all } s \in \mathbb{R}_+,\ x \in \mathbb{R}^N, \tag{3}$$

3. $\mu_0 \in \mathcal{R}^+(\mathbb{R}^N)$, *where* $\mathcal{R}^+(\mathbb{R}^N)$ *is the set of non-negative Radon measures on* $\mathbb{R}^N$,
4. μ_0 *is finite, and* $\int_{\mathbb{R}^N} \mathrm{d}\mu_0(x) = M_0 \geq 0$.

*Université Paris-Est & CNRS, LAMA UMR8050, francois.bouchut, robert.eymard, alain.prignet@univ-mlv.fr

Note that (1) and (2) imply

$$\int f(x)\mathrm{d}\mu_s(x) = \int f(X(s,t,x))\mathrm{d}\mu_t(x), \quad f \in C_c(\mathbb{R}^N),\ (s,t) \in D. \tag{4}$$

Such a family $(\mu_t)_{t\geq 0}$ arises for example in probabilistic studies involved by dynamic reliability problems, or population dynamics. In these situations, μ_0 and μ_t are probability measures related to the marginal distribution of Markov processes, transported for each state of a given system according to deterministic laws, describing the evolution of the state of the system $x \in \mathbb{R}^N$ with respect to the time variable. In full problems, the system can change of state at probabilistic times, and then a system of such equations, coupled by the right-hand side, must be solved: this problem is then modeled by the Chapman-Kolmogorov equation (see [10] for a complete description of piecewise deterministic Markov processes). The approximation of this problem is then useful in the framework of engineering studies (see e.g. [1, 2, 9, 11]). In some regular cases, one can assume that the deterministic evolution of the system is prescribed by a regular ODE under the following form

$$\partial_s X(s,t,x) = \boldsymbol{a}(s, X(s,t,x)),\ \forall s \geq t, \tag{5}$$

with the initial condition (3), assuming that $\boldsymbol{a}$ is bounded and Lipschitz continuous. In such a case, according to the Cauchy-Lipschitz framework, the function X, which necessarily satisfies (2), has more regularity properties than that assumed in Hypothesis 1, and the measure μ_t is the solution to the transport equation

$$\partial_t \mu_t + \mathrm{div}(\mu_t \boldsymbol{a}) = 0, \tag{6}$$

with initial condition μ_0. Note that many recent studies have been performed on the resolution of equation (6), exploring in particular the existence and uniqueness of a solution, depending on the regularity of $\boldsymbol{a}$ [3, 4]. Works that are related to the framework of measures are [7, 8, 6, 14, 5]. In [9], the approximation of (6) has been done by an explicit finite volume scheme, in the case where the functions X are C^1 solutions to (5); this scheme is identical, in the simplified case of a pure transport (1), to that presented in Section 2. The convergence proof of this scheme is provided in [9], under the condition that the space step tends to zero faster than the time step (this convergence result is again provided here under more general hypotheses made on X, in Section 2.2).

Besides, in [11], the problem (6) is approximated by an implicit finite volume method. The convergence proof is provided letting both discretization steps tend to zero independently. This proof is based on the numerical viscosity introduced in the scheme, in addition to the one coming from upstream weighting, and on discrete Sobolev inequalities (a similar idea is used in the modified scheme with viscosity presented in Section 3). The results obtained in [11] again rely on the regularity of the function $\boldsymbol{a}$.

In this paper, we only assume that, considering a system at the state x at time t, its state for times $s \geq t$ is given by $X(s,t,x)$, where the function X satisfies (2) and (3). The typical situation we would like to consider is the following. Assume that $x \in \mathbb{R}$ is the level of water in a reservoir, and that the outward rate of the water is equal to 1, as long as $x > 0$, but that this rate becomes equal to 0 if $x = 0$. Then the function $X(s,t,x)$ is given by $X(s,t,x) = \max(x-(s-t),0)$ for $x \geq 0$ and $X(s,t,x) = x$ for $x < 0$, which is only Lipschitz continuous. It corresponds to the coefficient $\boldsymbol{a}(t,x) = -\mathbf{1}_{x>0}$ in (5).

The aim of this paper is to study the convergence of an approximation of $(\mu_t)_{t\geq 0}$ defined by (1). The numerical schemes used here only involve the function X, and the minimal regularity hypotheses are made on X. Indeed, we do not use the vector field $\boldsymbol{a}$ in the definition of the schemes, nor in the proofs. This point of view comes from the probabilistic interpretation, where the set of trajectories is the natural object to consider, rather than the vector field of point velocities. The only assumption that X is Lipschitz continuous enables to cover many situations where mass concentration occurs in finite time. Then, the difficulty of uniqueness of solutions to (6) is somehow avoided. Indeed we shall see in our assumptions and in the counterexamples that the difficulty comes back in stability properties of the flow X, that were handled by the notion of *reversibility* in [7, 8]. The stability issue is here solved via an assumption of some bound on the compression of the flow.

The paper is organized as follows. As written above, we study in Section 2 a simple finite volume scheme, based on the transport by the flow of a piecewise constant function. We show that this method provides a convergent approximation to $(\mu_t)_{t\geq 0}$ defined by (1), provided that the time step and the ratio between the space step and the time step together tend to zero; examples show that this condition is necessary in the context of this scheme. In Section 3, we study a proof based on the numerical viscosity introduced in the scheme. A convergence result is then proved without the latter condition on the ratio between the discretization steps, but only under further restrictions:

- on the data (some boundedness assumption on the compression, i.e. on the determinant of $\partial X/\partial x$, is necessary, as well as the assumption that μ_0 has a bounded support),

- on the mesh (uniform regularity),

- on the scheme (the natural viscous part introduced by the scheme given in Section 2 is not sufficient for the proof using discrete Sobolev inequalities, and additional viscosity has to be introduced, as in [11]).

In Section 4, some technical results are given, that are used in Section 3. A conclusion is finally given in Section 5.

2 A finite volume scheme defined by characteristics

2.1 Definition of the scheme

Let us first give the definition of admissible meshes of $\mathbb{R}^N$, $N \geq 1$.

Definition 1. *An admissible mesh of $\mathbb{R}^N$ is a locally finite partition $\mathcal{M}$ of $\mathbb{R}^N$ such that:*

1. *for all $K \in \mathcal{M}$, K is bounded, the interior of K is an open convex subset of $\mathbb{R}^N$ and the N dimensional measure of K, denoted by $|K|$, is strictly positive,*
2. *for all $K \in \mathcal{M}$, denoting by ∂K the boundary of K, and, for all $L \in \mathcal{M}$, denoting by $\sigma_{KL} = \partial K \cap \partial L$, there exists $\mathcal{N}_K \subset \mathcal{M}$ such that $K \notin \mathcal{N}_K$ and $\partial K = \bigcup_{L\in\mathcal{N}_K} \sigma_{KL}$, and, for all $L \in \mathcal{N}_K$, σ_{KL}, called an edge of K, is included in a hyperplane of $\mathbb{R}^N$, with a strictly positive $N-1$ dimensional measure equal to $|\sigma_{KL}|$; we then denote by $\mathbf{n}_{KL}$ the unit normal vector to σ_{KL} oriented from K to L,*
3. *the size of the mesh, defined by $h_{\mathcal{M}} = \sup_{K\in\mathcal{M}} \operatorname{diam}(K)$, is finite.*

Note that all classical regular meshes of $\mathbb{R}^N$ are admissible in the sense of the previous definition.

Now, let $\mathcal{M}$ be an admissible mesh of $\mathbb{R}^N$ in the above sense. For such a mesh, we set:

$$u_K^0 = \frac{1}{|K|}\int_K \mathrm{d}\mu_0(x),\ \forall K \in \mathcal{M}. \tag{7}$$

For a given $\tau > 0$, defining $t_n = n\tau$, we set

$$V_{L,K}^n = \{x \in L, X(t_{n+1}, t_n, x) \in K\},\ \forall n \in \mathbb{N},\ \forall K, L \in \mathcal{M}, \tag{8}$$

and we set

$$v_{L,K}^n = |V_{L,K}^n|,\ \forall n \in \mathbb{N},\ \forall K, L \in \mathcal{M}. \tag{9}$$

Observe that

$$\bigcup_{K\in\mathcal{M}} V_{L,K}^n = L \quad \text{and} \quad \sum_{K\in\mathcal{M}} v_{L,K}^n = |L|,\ \forall n \in \mathbb{N},\ \forall L \in \mathcal{M}, \tag{10}$$

but, setting

$$w_K^n = \sum_{L\in\mathcal{M}} v_{L,K}^n = \left|\left\{x \in \mathbb{R}^N, X(t_{n+1}, t_n, x) \in K\right\}\right|,\ \forall n \in \mathbb{N},\ \forall K \in \mathcal{M}, \tag{11}$$

w_K^n is not equal to $|K|$ in the general case (this property would only hold in the case of incompressible flows). We then define a finite volume scheme for the approximation of μ_t defined by (1), by (7)-(9) and

$$|K|\ u_K^{n+1} = \sum_{L\in\mathcal{M}} v_{L,K}^n\ u_L^n,\ \forall n\in\mathbb{N},\ \forall K\in\mathcal{M}. \tag{12}$$

This formula can be interpreted as follows. Consider the piecewise constant functions $u_{\mathcal{M}}^n(x) = u_K^n$, for $x\in K$, $K\in\mathcal{M}$, i.e.

$$u_{\mathcal{M}}^n(x) = \sum_{K\in\mathcal{M}} \mathbf{1}_{x\in K} u_K^n,\ \forall n\in\mathbb{N},\ \forall x\in\mathbb{R}^N. \tag{13}$$

Then starting from the measure $u_{\mathcal{M}}^n(x)dx$ at time t_n and applying the transport formula (4) between times t_n and t_{n+1}, one gets a measure $\widetilde{\mu}^{n+1}$ such that

$$\int f(x)\mathrm{d}\widetilde{\mu}^{n+1}(x) = \int f(X(t_{n+1},t_n,x))u_{\mathcal{M}}^n(x)dx,\quad f\in C_c(\mathbb{R}^N). \tag{14}$$

This formula is still valid for f measurable, bounded with compact support (see the last paragraph of the proof of Lemma 10 for the argument). Taking $f(x) = \mathbf{1}_K(x)$ and using (13), one gets $\int_K \mathrm{d}\widetilde{\mu}^{n+1} = \sum_{L\in\mathcal{M}} v_{L,K}^n\ u_L^n$, and thus with (12), $u_K^{n+1} = (1/|K|)\int_K \mathrm{d}\widetilde{\mu}^{n+1}$. Therefore, $u_{\mathcal{M}}^{n+1}$ is nothing else than the piecewise constant projection over the mesh of the exact transport by the flow of $u_{\mathcal{M}}^n(x)dx$ during the time interval (t_n,t_{n+1}).

The scheme (12) can also be rewritten, using (11), as

$$\Big(|K| - w_K^n\Big)u_K^{n+1} + \sum_{L\in\mathcal{M}} v_{L,K}^n(u_K^{n+1} - u_L^n) = 0,\ \forall n\in\mathbb{N},\ \forall K\in\mathcal{M}, \tag{15}$$

or as

$$|K|(u_K^{n+1} - u_K^n) + \Big(|K| - w_K^n\Big)u_K^n + \sum_{L\in\mathcal{M}} v_{L,K}^n(u_K^n - u_L^n) = 0,\ \forall n\in\mathbb{N},\ \forall K\in\mathcal{M}, \tag{16}$$

which mimics a discretization of the equation $\partial_t u + u\,\mathrm{div}\,\boldsymbol{a} + \boldsymbol{a}\cdot\nabla u = 0$ (see (6)).

We finally define the fully discrete solution by

$$\begin{aligned} &u_{\mathcal{M},\tau}(0,x) = u_K^0,\ x\in K,\ K\in\mathcal{M},\\ &u_{\mathcal{M},\tau}(t,x) = u_K^{n+1},\ t\in\,]t_n,t_{n+1}],\ n\in\mathbb{N},\ x\in K,\ K\in\mathcal{M}. \end{aligned} \tag{17}$$

Example 1. We consider the following data:

1. $N=1$, and $\mu_0 = \mathbf{1}_{x\in(0,1)}dx$,
2. the function X is defined by

$$\begin{aligned} &X(s,t,x) = x-(s-t),\quad \text{for all } s\geq t \text{ and } x\geq s-t,\\ &X(s,t,x) = 0,\quad \text{for all } s\geq t \text{ and } 0\leq x\leq s-t,\\ &X(s,t,x) = x,\quad \text{for all } s\geq t \text{ and } x\leq 0. \end{aligned}$$

Then the measure μ_t is given, for $t\in(0,1)$, by $\mu_t = t\delta_0(x) + \mathbf{1}_{x\in(0,1-t)}\mathrm{d}x$, and for $t\in[1,+\infty)$, by $\mu_t = \delta_0$. The scheme, obtained with the mesh $\mathcal{M} = \{[ih,(i+1)h),\ i\in\mathbb{Z}\}$ with $h = 1/M$ for an integer $M>0$, and $\tau = h$, reads (see Figure 1)

$$\begin{aligned} &u_{[0,h)}^{n+1} = u_{[0,h)}^n + u_{[h,2h)}^n,\\ &u_{[ih,(i+1)h)}^{n+1} = u_{[(i+1)h,(i+2)h)}^n,\ i\geq 1,\\ &u_{[ih,(i+1)h)}^{n+1} = u_{[ih,(i+1)h)}^n,\ i<0. \end{aligned}$$

The convergence of the scheme is then obvious.

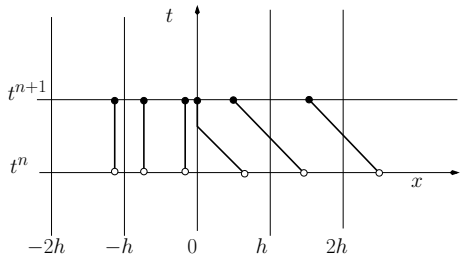

Figure 1: Construction of the scheme for Example 1

It happens that in one dimension, the monotonicity criterion following from the notion of reversible solutions of [7] enables to prove the convergence of the scheme, see [13], without any special condition relating $h_{\mathcal{M}}$ and τ. Let us consider a second example, showing that in two dimensions for the scheme defined as above, we cannot let $h_{\mathcal{M}} \to 0$ and $\tau \to 0$ with $h_{\mathcal{M}}/\tau$ constant without trouble.

Example 2. We consider the following data, taken from [8]: $N = 2$, $x = (x_1, x_2) \in \mathbb{R}^2$,

$$\begin{aligned} &X_\lambda(s,t,x_1,x_2) = \Big(\operatorname{sgn}(x_1)\max(0,|x_1| - (s-t)),\quad x_2 + \lambda \max(0, s-t-|x_1|)\Big), \\ &0 \le t \le s,\ x_1, x_2 \in \mathbb{R}, \end{aligned} \tag{18}$$

where $\lambda \in \mathbb{R}$ is a parameter. Then X_λ is globally Lipschitz continuous, and satisfies (2)-(3). Indeed, one can check that X_λ satisfies (5) with the vector field

$$\boldsymbol{a}_\lambda(t,x_1,x_2) = \begin{cases} (-\operatorname{sgn} x_1, 0) & \text{for } x_1 \neq 0, \\ (0,\lambda) & \text{for } x_1 = 0. \end{cases} \tag{19}$$

We consider the case where $\lambda = 1$, the mesh is defined by squares of side $h > 0$:

$$K_{ij} = [(i-1)h, ih) \times [(j-1)h, jh), \quad i,j \in \mathbb{Z}. \tag{20}$$

and the time step is defined by $\tau = h$. We notice (see Figure 2) that, for all $j \in \mathbb{Z}$,

1. $V^n_{K_{0,j-1},K_{1,j}} = \{(x_1,x_2) \in K_{0,j-1}, x_1 + x_2 - (j-2)h \ge 0\}$, which provides $v^n_{K_{0,j-1},K_{1,j}} = h^2/2$,
2. $V^n_{K_{1,j-1},K_{1,j}} = \{(x_1,x_2) \in K_{1,j-1}, x_1 - x_2 + (j-2)h \le 0\}$, which provides $v^n_{K_{1,j-1},K_{1,j}} = h^2/2$,
3. $V^n_{K_{0,j},K_{1,j}} = \{(x_1,x_2) \in K_{0,j}, x_1 + x_2 - (j-1)h < 0\}$, which provides $v^n_{K_{0,j},K_{1,j}} = h^2/2$,
4. $V^n_{K_{1,j},K_{1,j}} = \{(x_1,x_2) \in K_{1,j}, x_1 - x_2 + (j-1)h > 0\}$, which provides $v^n_{K_{1,j},K_{1,j}} = h^2/2$,
5. $V^n_{K_{i+1,j},K_{i,j}} = K_{i+1,j}$, for all $i \in \mathbb{N}$,
6. $V^n_{K_{-i-1,j},K_{-i,j}} = K_{-i-1,j}$, for all $i \in \mathbb{N}$.

Therefore Scheme (12) can be written as

$$u^{n+1}_{ij} = \begin{cases} u^n_{i+1,j} & \text{if } i \ge 2, \\ u^n_{i-1,j} & \text{if } i \le 0, \end{cases} \tag{21}$$

and

$$u^{n+1}_{1,j} = \frac{1}{2}\left(u^n_{1,j-1} + u^n_{1,j}\right) + u^n_{2,j} + \frac{1}{2}\left(u^n_{0,j-1} + u^n_{0,j}\right). \tag{22}$$

In order to evaluate what happens to the mass once it arrives in the cells for which $i = 1$, we write down (22) when $u^n_{i,j} = 0$ for all $i \neq 1$, which gives the one-dimensional scheme

$$u^{n+1}_j = \frac{1}{2}\left(u^n_{j-1} + u^n_j\right) = u^n_j - \frac{1}{2}\frac{\tau}{h}\left(u^n_j - u^n_{j-1}\right), \tag{23}$$

with $\tau/h = 1$. Since this scheme converges to the solution of the transport by the velocity $1/2$ as $h \to 0$ with $\tau = h$, we get that, instead of converging to the continuous solution resulting from the transport by the flow X_1

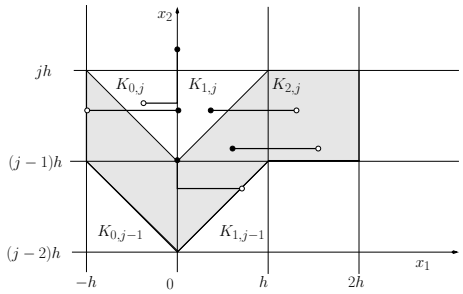

Figure 2: Set of all points (x_1, x_2) such that $X(t_{n+1}, t_n, x_1, x_2) \in K_{1,j}$

($\lambda = 1$) as $h \to 0$ with $\tau = h$, Scheme (12) converges to the continuous solution resulting from the transport by the flow $X_{1/2}$ ($\lambda = 1/2$). This is due to the fact that during a time-step, the mass contained in a vertical segment joining the points $(0, (j-1)h)$ and $(0, jh)$ is first spread to the cell $K_{1,j}$ (because we replace the measure by a piecewise constant function), and then only half of this mass is transported to the segment joining the points $(0, jh)$ and $(0, (j+1)h)$, the other half coming back to the segment joining the points $(0, (j-1)h)$ and $(0, jh)$.

More generally, one can check that Scheme (12) converges, as $h \to 0$ with $\tau = \alpha h$ such that $\alpha \in (0, 2]$, to the continuous solution resulting from the transport by the flow X_λ, with $\lambda = \alpha/2$ if $\alpha \in (0, 1]$ and $\lambda = 1 - 1/(2\alpha)$ if $\alpha \in [1, 2]$.

2.2 Convergence of the scheme under the conditions $\tau \to 0$ and $h_{\mathcal{M}}/\tau \to 0$

Theorem 1 (Convergence of Scheme (12) as $\tau \to 0$ and $h_{\mathcal{M}}/\tau \to 0$)**.** *Under Hypothesis 1, let, for any admissible mesh $\mathcal{M}$ of $\mathbb{R}^N$ in the sense of Definition 1 and any $\tau > 0$, $u_{\mathcal{M},\tau}$ be given by* (7)-(12) *and* (17)*. Then, for all $t \geq 0$, $u_{\mathcal{M},\tau}(t, \cdot)$ converges for the weak-$\star$ topology of $\mathcal{R}^+(\mathbb{R}^N)$ to μ_t defined by* (1)*, as $\tau \to 0$ and $h_{\mathcal{M}}/\tau \to 0$.*

Proof. We first notice that, using (12), the following relation holds,

$$\sum_{K\in\mathcal{M}} |K|\ u_K^{n+1} = \sum_{K\in\mathcal{M}} \sum_{L\in\mathcal{M}} v_{L,K}^n\ u_L^n = \sum_{L\in\mathcal{M}} u_L^n \sum_{K\in\mathcal{M}} v_{L,K}^n = \sum_{L\in\mathcal{M}} |L|\ u_L^n = \sum_{L\in\mathcal{M}} |L|\ u_L^0 = M_0. \tag{24}$$

Let $g \in \mathrm{Lip}_c(\mathbb{R}^N)$ be a Lipschitz continuous function with compact support, with $\mathrm{Lip}(g) \geq 0$ its Lipschitz constant. According also to (12), we have, for $n \in \mathbb{N}$,

$$\int_{\mathbb{R}^N} g(x) u_{\mathcal{M}}^{n+1}(x)\mathrm{d}x = \sum_{K\in\mathcal{M}} |K| u_K^{n+1} \frac{1}{|K|}\int_K g(x)\mathrm{d}x = \sum_{L\in\mathcal{M}} u_L^n \sum_{K\in\mathcal{M}} v_{L,K}^n \frac{1}{|K|}\int_K g(x)\mathrm{d}x.$$

Therefore we get

$$\begin{aligned}
&\int_{\mathbb{R}^N} g(x) u_{\mathcal{M}}^{n+1}(x)\mathrm{d}x - \int_{\mathbb{R}^N} g(X(t_{n+1}, t_n, x)) u_{\mathcal{M}}^n(x)\mathrm{d}x \\
&\qquad = \sum_{L\in\mathcal{M}} u_L^n \sum_{K\in\mathcal{M}} v_{L,K}^n \left(\frac{1}{|K|}\int_K g(x)\mathrm{d}x - \frac{1}{v_{L,K}^n}\int_{V_{L,K}^n} g(X(t_{n+1}, t_n, x))\mathrm{d}x \right).
\end{aligned} \tag{25}$$

Since we have for $x \in K$ and for $y \in V_{L,K}^n$, $X(t_{n+1}, t_n, y) \in K$, we can estimate

$$\begin{aligned}
&\left| \frac{1}{|K|}\int_K g(x)\mathrm{d}x - \frac{1}{v_{L,K}^n}\int_{V_{L,K}^n} g(X(t_{n+1}, t_n, x))\mathrm{d}x \right| \\
\leq &\frac{1}{|K| v_{L,K}^n}\int_K \int_{V_{L,K}^n} \Big| g(x) - g(X(t_{n+1}, t_n, y)) \Big| \,\mathrm{d}y\mathrm{d}x \\
\leq &\,\mathrm{Lip}(g) h_{\mathcal{M}}.
\end{aligned}$$

We therefore have, using (25),

$$\left|\int_{\mathbb{R}^N} g(x)u_{\mathcal{M}}^{n+1}(x)\mathrm{d}x - \int_{\mathbb{R}^N} g(X(t_{n+1},t_n,x))u_{\mathcal{M}}^{n}(x)\mathrm{d}x\right| \leq M_0\,\mathrm{Lip}(g)h_{\mathcal{M}}. \tag{26}$$

Let $t>0$, and let $n_t\in\mathbb{N}$ be such that $t\in\,]t_{n_t},t_{n_t+1}]$, which means that $n_t\tau<t\leq(n_t+1)\tau$. We now consider, for a given Lipschitz continuous function $f\in\mathrm{Lip}_c(\mathbb{R}^N)$ with constant $\mathrm{Lip}(f)$, the functions $f_n\ :\ x\mapsto f(X(t_{n_t+1},t_{n+1},x))$, for $n=0,1,2,\ldots,n_t$. They satisfy

$$|f_n(x)-f_n(y)|\leq\mathrm{Lip}(f)|X(t_{n_t+1},t_{n+1},x)-X(t_{n_t+1},t_{n+1},y)|\leq\xi\,\mathrm{Lip}(f)|x-y|,\ \forall x,y\in\mathbb{R}^N.$$

Therefore they are all Lipschitz continuous, and have compact support since $|X(t_{n_t+1},t_{n+1},x)-x|\leq\xi(t_{n_t+1}-t_{n+1})\leq\xi t$. We can apply the inequality (26), replacing g by f_n and $\mathrm{Lip}(g)$ by $\xi\,\mathrm{Lip}(f)$. We get, according to (2) that implies $X(t_{n_t+1},t_{n+1},X(t_{n+1},t_n,x))=X(t_{n_t+1},t_n,x)$,

$$\left|\int_{\mathbb{R}^N} f(X(t_{n_t+1},t_{n+1},x))u_{\mathcal{M}}^{n+1}(x)\mathrm{d}x - \int_{\mathbb{R}^N} f(X(t_{n_t+1},t_n,x))u_{\mathcal{M}}^{n}(x)\mathrm{d}x\right|\leq M_0\xi\,\mathrm{Lip}(f)h_{\mathcal{M}}.$$

Summing the above relations for $n=0,\ldots,n_t$, we obtain

$$\left|\int_{\mathbb{R}^N} f(x)u_{\mathcal{M}}^{n_t+1}(x)\mathrm{d}x - \int_{\mathbb{R}^N} f(X(t_{n_t+1},0,x))u_{\mathcal{M}}^{0}(x)\mathrm{d}x\right|\leq(n_t+1)M_0\xi\,\mathrm{Lip}(f)h_{\mathcal{M}}. \tag{27}$$

Moreover, writing

$$\begin{aligned}
&\int_{\mathbb{R}^N} f(X(t_{n_t+1},0,x))u_{\mathcal{M}}^{0}(x)\mathrm{d}x - \int_{\mathbb{R}^N} f(X(t,0,x))\mathrm{d}\mu_0(x)\\
&\quad=\int_{\mathbb{R}^N} f(X(t_{n_t+1},0,x))u_{\mathcal{M}}^{0}(x)\mathrm{d}x - \int_{\mathbb{R}^N} f(X(t_{n_t+1},0,x))\mathrm{d}\mu_0(x)\\
&\qquad+\int_{\mathbb{R}^N} f(X(t_{n_t+1},0,x))\mathrm{d}\mu_0(x) - \int_{\mathbb{R}^N} f(X(t,0,x))\mathrm{d}\mu_0(x),
\end{aligned}$$

we have, using that $t\in\,]t_{n_t},t_{n_t+1}]$,

$$\left|\int_{\mathbb{R}^N} f(X(t_{n_t+1},0,x))u_{\mathcal{M}}^{0}(x)\mathrm{d}x - \int_{\mathbb{R}^N} f(X(t,0,x))\mathrm{d}\mu_0(x)\right|\leq M_0\xi\,\mathrm{Lip}(f)(h_{\mathcal{M}}+\tau), \tag{28}$$

because $\xi\,\mathrm{Lip}(f)$ is a Lipschitz constant of the function $(t,x)\mapsto f(X(t,0,x))$. We thus get from (27) and (28)

$$\left|\int_{\mathbb{R}^N} f(x)u_{\mathcal{M},\tau}(t,x)\mathrm{d}x - \int_{\mathbb{R}^N} f(X(t,0,x))\mathrm{d}\mu_0(x)\right|\leq M_0\xi\,\mathrm{Lip}(f)\left((t+\tau)\frac{h_{\mathcal{M}}}{\tau}+h_{\mathcal{M}}+\tau\right),$$

which tends to zero as $\tau\to0$ and $\frac{h_{\mathcal{M}}}{\tau}\to0$. Using (1), this proves the theorem, since we can extend the convergence to any $f\in C_c(\mathbb{R}^N)$ by approximation, using the bound (24).

□

3 Modified scheme with viscosity

3.1 Definition of the modified scheme and first properties

Assuming Hypothesis 1, let $J(s,t,x)$ denote, for all $(s,t)\in D$ and a.e. $x\in\mathbb{R}^N$, the Jacobian determinant with respect to x of $X(s,t,x)$ (that is the determinant of the matrix $(\partial X_i/\partial x_j)$, $i,j=1,\ldots,N$). We have $J(s,t,x)\geq0$ according to Lemma 9 (we shall not use this property), and $J\in L^\infty$ ($|J(s,t,x)|\leq\xi^N$, where ξ is given in Hypothesis 1). Let us introduce the additional hypotheses and notations:

Hypothesis 2 (Estimate of the compression)**.** *We assume that :*

1. *there exists $p_0 > N$ such that for all $T > 0$ there exists some $\zeta_T \geq 0$ satisfying*

$$\|J(s,t,\cdot) - 1\|_{L^{p_0}(\mathbb{R}^N)} \leq \zeta_T(s-t),\ \forall (s,t) \in D_T, \tag{29}$$

   where $D_T = \{(s,t) \in \mathbb{R}^2, 0 \leq t \leq s \leq T\}$,

2. *the support of μ_0 is bounded.*

Remark 1. *We could make the slightly weaker assumption that there exists some $p_0 > N$ such that for all $T > 0$ there exists C_T such that for all finite increasing sequence $(s_n)_n$ of non-negative values bounded by T,*

$$\sum_n \|J(s_{n+1}, s_n, \cdot) - 1\|_{L^{p_0}_{\mathrm{loc}}(\mathbb{R}^N)} \leq C_T. \tag{30}$$

These assumptions (29) *and* (30) *can be understood by writing the following formula, valid when X solves the ODE* (5) *for some smooth coefficient $\boldsymbol{a}$:*

$$J(s,t,x) = \exp\left(\int_t^s (\operatorname{div} \boldsymbol{a})(\sigma, X(\sigma,t,x))d\sigma\right). \tag{31}$$

Thus, (29) *means more or less that $\operatorname{div} \boldsymbol{a} \in L^\infty(0,T,L^{p_0}(\mathbb{R}^N))$, while* (30) *means more or less that $\operatorname{div} \boldsymbol{a} \in L^1(0,T,L^{p_0}_{loc}(\mathbb{R}^N))$. However,* (29) *and* (30) *are formulated in terms of X only, and do not involve any vector field $\boldsymbol{a}$, according to the spirit of this work.*

Remark 2. *In the case of Example 1, one has $J(s,t,x) = 1$ for a.e. $x < 0$ and $x > s - t$ and $J(s,t,x) = 0$ for a.e. $0 < x < s - t$. Thus, for any $p_0 > 1$, $\|J(s,t,\cdot) - 1\|_{L^{p_0}(\mathbb{R})} = (s-t)^{1/p_0}$ and* (29) *is not satisfied. It is nevertheless possible to prove the convergence of the schemes presented in this paper in the one-dimensional case, by the method of [13].*

Example 3. This example is dedicated to show that it is not possible to prove the convergence of the approximate solution provided by Scheme (12) to the continuous solution, assuming Hypothesis 1, τ/h constant, and assuming (29) with $p_0 < N - 1$ (this means that the condition $p_0 > N$ in (29) is in some way optimal). Consider the generalization of Example 2, given, for $x = (x', y) \in \mathbb{R}^{N-1} \times \mathbb{R}$ ($N \geq 2$) by

$$\begin{aligned} X_\lambda(s,t,x',y) &= \left(\frac{x'}{|x'|}\max(0, |x'| - (s-t)),\ \ y + \lambda \max(0, s - t - |x'|)\right), \\ &0 \leq t \leq s,\ (x', y) \in \mathbb{R}^{N-1} \times \mathbb{R}, \end{aligned} \tag{32}$$

where $\lambda \in \mathbb{R}$ is a parameter. Then X_λ is globally Lipschitz continuous, and satisfies (2)-(3). Indeed, X_λ satisfies (5) with the vector field

$$\boldsymbol{a}_\lambda(t,x',y) = \begin{cases} \left(-\dfrac{x'}{|x'|}, 0\right) & \text{for } x' \neq 0, \\ (0,\lambda) & \text{for } x' = 0. \end{cases} \tag{33}$$

Denoting the first components of (32) by X'_λ, one has

$$\nabla_{x'} X'_\lambda = \left(\mathrm{Id}\Big(1 - \frac{s-t}{|x'|}\Big) + (s-t)\frac{x' \otimes x'}{|x'|^3}\right)\mathbf{1}_{|x'|>s-t}. \tag{34}$$

Therefore, one gets

$$J_\lambda(s,t,x',y) = \left(1 - \frac{s-t}{|x'|}\right)^{N-2} \mathbf{1}_{|x'|>s-t}, \tag{35}$$

and (29) is satisfied (locally in space) if and only if $p_0 < N - 1$ (except for $N = 2$ where the limit case $p_0 = 1$ is satisfied).

We consider a mesh of cubes of side $h > 0$, but differently from in Example 2, we take one cube centered around the origin. We take $\lambda = 1$, $\tau \leq h/2$, and we apply the scheme (12) with $\tau/h = cst$. We see that the points

that lie in the cube centered at the origin at time t_n and that travel to the neighboring cube in the direction of increasing y during the time-step, are those for which $|x'| + y \leq \tau$. Hence the corresponding volume is

$$\int_0^\tau |B_{N-1}(0, \tau - y)|\, dy = |B_{N-1}(0,1)| \frac{\tau^N}{N}. \tag{36}$$

Therefore, dividing by h^N, we find that the ratio of mass that travels to the neighboring cube by the mass originally in the cube is $\lambda_{eff}\tau/h$, with

$$\lambda_{eff} = \frac{|B_{N-1}(0,1)|}{N} \frac{\tau^{N-1}}{h^{N-1}}, \tag{37}$$

and instead of converging to the transport by the flow X_1 ($\lambda = 1$), the scheme converges to the transport by the flow $X_{\lambda_{eff}}$.

Example 4. This example is to show that, although the compression condition (29) does not allow $\operatorname{div} \boldsymbol{a}$ to be a measure, i.e. $p_0 = 1$, Hypothesis (2) nevertheless covers cases where, for some $t > 0$, the solution μ_t of (1) is a Dirac measure, whereas μ_0 is defined by a density function. Let us take $N = 1$ and

$$X(s,t,x) = \begin{cases} x \quad \text{if } x \leq 0, \\ (\max(0, \sqrt{x} - (s-t)/2))^2 \quad \text{if } 0 \leq x \leq 1, \\ x - (s-t) \quad \text{if } x \geq 1 \text{ and } s - t \leq x - 1, \\ (\max(0, 1 - (s-t-x+1)/2))^2 \quad \text{if } x \geq 1 \text{ and } x - 1 \leq s - t. \end{cases} \tag{38}$$

Then X corresponds to the vector field $\boldsymbol{a}(x) = 0$ for $x < 0$, $\boldsymbol{a}(x) = -\sqrt{x}$ for $0 < x < 1$, and $\boldsymbol{a}(x) = -1$ for $x > 1$

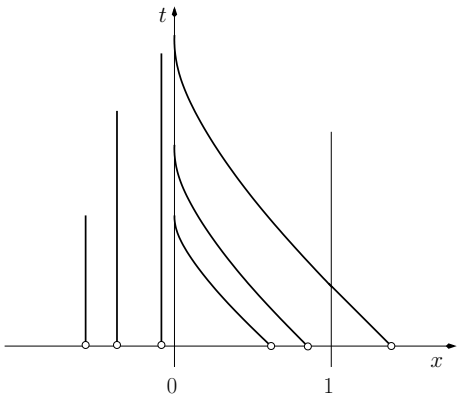

Figure 3: Characteristics in the case of Example 4

(see Figure 4). We observe that any compact set in the positive half space is sent to 0 after a finite time. One can also check that (29) is satisfied for all $p_0 < 2$.

In this section, we shall use the regularity of admissible meshes, defined as follows.

Definition 2. *Let $\mathcal{M}$ be an admissible mesh of $\mathbb{R}^N$ (with $N \in \mathbb{N}^\star$) in the sense of Definition 1. For a given $\rho > 0$, we say that $\mathcal{M}$ is ρ-regular if*

$$\frac{1}{\rho} h_{\mathcal{M}} \sum_{L \in \mathcal{N}_K} |\sigma_{KL}| \leq |K| \leq \rho h_{\mathcal{M}} \sum_{L \in \mathcal{N}_K} |\sigma_{KL}|, \ \forall K \in \mathcal{M}, \tag{39}$$

$$\frac{1}{\rho} h_{\mathcal{M}} \leq \operatorname{diam}(K) \leq h_{\mathcal{M}}, \ \forall K \in \mathcal{M}, \tag{40}$$

and

$$\frac{1}{\rho} h_{\mathcal{M}}^N \leq |K| \leq \rho h_{\mathcal{M}}^N, \ \forall K \in \mathcal{M}. \tag{41}$$

Remark 3. *A cubic mesh with constant side is ρ-regular for some constant ρ only depending on the dimension N.*

We again define a finite volume scheme for the approximation of μ_t defined by (1). Let $\mathcal{M}$ be an admissible mesh of $\mathbb{R}^N$ in the sense of Definition 1 and $\tau > 0$ be given. We still denote $t_n = n\tau$ for $n \in \mathbb{N}$, and u_K^0, $V_{L,K}^n$ and $v_{L,K}^n$ are again defined by (7)-(8)-(9). For a given $\nu > 0$, we define

$$\begin{gathered}\widehat{v}_{L,K}^n = v_{L,K}^n + \nu\ \tau|\sigma_{KL}|,\ \forall n \in \mathbb{N},\ \forall K \in \mathcal{M},\ \forall L \in \mathcal{M} \setminus \{K\},\\ \widehat{v}_{K,K}^n = v_{K,K}^n - \nu\ \tau \sum_{L \in \mathcal{M}\setminus\{K\}} |\sigma_{KL}|,\ \forall n \in \mathbb{N},\ \forall K \in \mathcal{M}.\end{gathered} \tag{42}$$

Let us observe that for all $L, K \in \mathcal{M}$, again defining w_K^n by (11), we have

$$\sum_{K \in \mathcal{M}} \widehat{v}_{L,K}^n = |L| \quad \text{and} \quad \sum_{L \in \mathcal{M}} \widehat{v}_{L,K}^n = w_K^n. \tag{43}$$

We may now define the modification of Scheme (12) by the additional viscosity term tuned by the value of ν:

$$|K|\ u_K^{n+1} = \sum_{L \in \mathcal{M}} \widehat{v}_{L,K}^n\ u_L^n,\ \forall n \in \mathbb{N},\ \forall K \in \mathcal{M}. \tag{44}$$

It results that, thanks to (43), the following conservation property, similar to (24), holds:

$$\sum_{K \in \mathcal{M}} |K|\ u_K^{n+1} = \sum_{K \in \mathcal{M}} \sum_{L \in \mathcal{M}} \widehat{v}_{L,K}^n\ u_L^n = \sum_{L \in \mathcal{M}} u_L^n \sum_{K \in \mathcal{M}} \widehat{v}_{L,K}^n = \sum_{L \in \mathcal{M}} |L| u_L^n = \sum_{L \in \mathcal{M}} |L| u_L^0 = M_0. \tag{45}$$

Remark 4 (Finite number of non zero terms)**.** *Since the support of μ_0 is bounded, for each $n \in \mathbb{N}$, there is only a finite number of $K \in \mathcal{M}$ for which $u_K^n \neq 0$. Indeed, according to the Lipschitz continuity of X one has $|X(t_{n+1}, t_n, x) - x| \leq \xi\ \tau$, thus $V_{L,K}^n$ in (8) is non empty only if $\mathrm{d}(L, K) \leq \xi\ \tau$. We deduce that $\widehat{v}_{L,K}^n$ vanishes for $\mathrm{d}(L, K) > \xi\ \tau$, and that at each time-step, the support extends only to the cells that are at distance at most $\xi\ \tau$ from the support at the previous time.*

Lemma 1 (Nonnegativity of the scheme)**.** *Under Hypothesis 1, let $\rho > 0$, $\lambda \in (0, 1)$ be given. Let $\mathcal{M}$ be an admissible mesh of $\mathbb{R}^N$ in the sense of Definition 1, which is ρ-regular in the sense of Definition 2. Let $\tau > 0$, $\nu > 0$ be given such that the CFL condition*

$$\tau \leq \frac{1 - \lambda}{\rho(\xi + \nu)} h_{\mathcal{M}} \tag{46}$$

holds. Then the coefficients $\widehat{v}_{K,K}^n$ given by (7)-(9) and (42) satisfy

$$\widehat{v}_{K,K}^n \geq \lambda |K|,\ \forall K \in \mathcal{M},\ \forall n \in \mathbb{N}. \tag{47}$$

Moreover, letting $(u_K^n)_{K \in \mathcal{M}, n \in \mathbb{N}}$ be given by (7)-(9) and (42)-(44), we have

$$u_K^n \geq 0,\ \forall K \in \mathcal{M},\ \forall n \in \mathbb{N}. \tag{48}$$

Finally, there exists β, only depending on ρ and N, such that

$$w_K^n \leq \beta |K| \leq \frac{\beta}{\lambda}\ \widehat{v}_{K,K}^n,\ \forall K \in \mathcal{M},\ \forall n \in \mathbb{N}. \tag{49}$$

Proof. According to the regularity condition (39), we have, for all $K \in \mathcal{M}$,

$$\nu\ \tau \sum_{L \in \mathcal{M}\setminus\{K\}} |\sigma_{KL}| \leq \frac{\nu(1 - \lambda)}{\rho(\xi + \nu)} \sum_{L \in \mathcal{M}\setminus\{K\}} |\sigma_{KL}| h_{\mathcal{M}} \leq \frac{\nu(1 - \lambda)}{\xi + \nu} |K|.$$

Recall that $V^n_{K,K}$ is defined by $V^n_{K,K} = \{x \in K, X(t_{n+1}, t_n, x) \in K\}$. Let $x \in K$ such that $\mathrm{d}(x, \partial K) > \frac{\xi(1-\lambda)}{\rho(\xi+\nu)} h_{\mathcal{M}}$. Then, under the CFL condition (46), we have

$$|X(t_{n+1}, t_n, x) - x| = |X(t_{n+1}, t_n, x) - X(t_n, t_n, x)| \leq \xi\tau \leq \frac{\xi(1-\lambda)}{\rho(\xi+\nu)} h_{\mathcal{M}},$$

and thus $X(t_{n+1}, t_n, x) \in K$, that means $x \in V^n_{K,K}$. Therefore, since the measure of all $x \in K$ such that $\mathrm{d}(x, \partial K) \leq \alpha$ is lower than $\alpha\,|\partial K|$ (using the convexity of K), we get, using (39),

$$v^n_{K,K} \geq \left| \left\{ x \in K,\ \mathrm{d}(x, \partial K) > \frac{\xi(1-\lambda)}{\rho(\xi+\nu)} h_{\mathcal{M}} \right\} \right| \geq |K| - \frac{\xi(1-\lambda)}{\rho(\xi+\nu)} h_{\mathcal{M}} \sum_{L \in \mathcal{M}\setminus\{K\}} |\sigma_{KL}| \geq |K| - \frac{\xi(1-\lambda)}{\xi+\nu}|K|.$$

Thus

$$\widehat{v}^n_{K,K} \geq |K| - \frac{\xi(1-\lambda)}{\xi+\nu}|K| - \frac{\nu(1-\lambda)}{\xi+\nu}|K| = \lambda|K|.$$

This proves (47), as well as its immediate consequence (48).

We now remark that, since $|X(t_{n+1}, t_n, x) - x| \leq \xi\ \tau$,

$$w^n_K \leq |B(0, h_{\mathcal{M}} + \tau\xi)| \leq |B(0, h_{\mathcal{M}} + \frac{h_{\mathcal{M}}}{\rho})|. \tag{50}$$

Using the regularity condition (41), we get the existence of β, only depending on ρ and N, such that (49) holds (the last inequality resulting from (47)). □

The following lemma completes Remark 4.

Lemma 2 (Support of the approximate solution)**.** *Under Hypotheses 1 and 2, let $\alpha > 0$, $\nu > 0$ be given. Let $\mathcal{M}$ be an admissible mesh of $\mathbb{R}^N$ in the sense of Definition 1, and let $\tau > 0$ be given such that the inverse CFL condition*

$$h_{\mathcal{M}} \leq \alpha\tau \tag{51}$$

holds. Let $(u^n_K)_{K\in\mathcal{M}, n\in\mathbb{N}}$ be given by (7)-(8) *and* (42)-(44)*. Then, for all $T \geq \tau$, there exists $R > 0$, only depending on T, supp(μ_0), ξ, and α such that*

$$u^n_K \neq 0 \Rightarrow K \subset \bar{B}(0, R),\ \forall n \in \mathbb{N}, n\tau \leq T,\ \forall K \in \mathcal{M},$$

denoting by $\bar{B}(0, R)$ the closed ball with center 0 and radius R.

Proof. Let R_0 be such that the support of μ_0 is included in $\bar{B}(0, R_0)$. Then, considering $\widetilde{R}_0 = R_0 + h_{\mathcal{M}}$, we have

$$u^0_K \neq 0 \Rightarrow K \subset \bar{B}(0, \widetilde{R}_0),\ \forall K \in \mathcal{M}.$$

Let us argue by induction, and assume that for a given $n \in \mathbb{N}$, $\widetilde{R}_n$ is such that

$$u^n_K \neq 0 \Rightarrow K \subset \bar{B}(0, \widetilde{R}_n),\ \forall K \in \mathcal{M}. \tag{52}$$

Let $K \in \mathcal{M}$ such that $u^{n+1}_K \neq 0$. Then there exists $L \in \mathcal{M}$ (the case $L = K$ is possible) such that $u^n_L \neq 0$ and $\widehat{v}^n_{L,K} \neq 0$, which means that $v^n_{L,K} \neq 0$ or $|\sigma_{KL}| > 0$. In the first case, then $\mathrm{d}(K, L) \leq \xi\ \tau$, in the second one, $\mathrm{d}(K, L) = 0$. Therefore, we have $K \subset \bar{B}(0, \widetilde{R}_{n+1})$, with $\widetilde{R}_{n+1} = \widetilde{R}_n + \xi\ \tau + h_{\mathcal{M}}$, and this proves (52) at level $n + 1$.

We find finally that $\widetilde{R}_n = R_0 + h_{\mathcal{M}} + n(h_{\mathcal{M}} + \xi\ \tau)$. Therefore, for all $n \in \mathbb{N}$ such that $n\tau \leq T$, one has $\widetilde{R}_n \leq R_0 + \alpha\tau + n\tau(\alpha + \xi) \leq R_0 + (2\alpha + \xi)T$. This concludes the proof with $R = R_0 + (2\alpha + \xi)T$.
□

3.2 Convergence

Theorem 2 (Convergence of the Scheme (44) as $h_{\mathcal{M}} \to 0$ and $\tau \simeq h_{\mathcal{M}}$.)**.** *Under Hypotheses 1 and 2, let $\rho > 0$, $\lambda \in (0,1)$, $\nu > 0$ and $\alpha > 0$ be fixed given values. Let $\mathcal{M}$ be an admissible mesh of $\mathbb{R}^N$ in the sense of Definition 1, which is ρ-regular in the sense of Definition 2, and let τ be given such that the CFL condition* (46) *and the inverse CFL condition* (51) *hold. Let $(u_K^n)_{K\in\mathcal{M},n\in\mathbb{N}}$ be given by* (7)-(8) *and* (42)-(44) *and let $u_{\mathcal{M},\tau}$ be defined by* (17)*. Then, for all $t \geq 0$, $u_{\mathcal{M},\tau}(t,\cdot)$ converges for the weak-$\star$ topology in $\mathcal{R}^+(\mathbb{R}^N)$ to μ_t defined by* (1)*, as $h_{\mathcal{M}} \to 0$ (which is equivalent to $\tau \to 0$ under conditions* (46) *and* (51)*).*

Proof. Let $g \in \mathrm{Lip}_c(\mathbb{R}^N)$ be a Lipschitz continuous function with compact support. Let us multiply (44), for a given $n \in \mathbb{N}$, by $\frac{1}{|K|}\int_K g(x)\mathrm{d}x$, and sum over $K \in \mathcal{M}$. We get with the notation (13)

$$\int_{\mathbb{R}^N} g(x)u_{\mathcal{M}}^{n+1}(x)\mathrm{d}x - \sum_{K\in\mathcal{M}} \frac{1}{|K|}\int_K g(x)\mathrm{d}x \sum_{L\in\mathcal{M}} \widehat{v}_{L,K}^n \; u_L^n = 0,$$

which provides

$$\int_{\mathbb{R}^N} g(x)u_{\mathcal{M}}^{n+1}(x)\mathrm{d}x - \int_{\mathbb{R}^N} g(X(t_{n+1},t_n,x))u_{\mathcal{M}}^n(x)\mathrm{d}x - R^n(g) = 0, \tag{53}$$

with

$$R^n(g) = \sum_{K\in\mathcal{M}} \frac{1}{|K|}\int_K g(x)\mathrm{d}x \sum_{L\in\mathcal{M}} \widehat{v}_{L,K}^n \; u_L^n - \int_{\mathbb{R}^N} g(X(t_{n+1},t_n,x))u_{\mathcal{M}}^n(x)\mathrm{d}x. \tag{54}$$

We can then write

$$R^n(g) = \sum_{L\in\mathcal{M}} u_L^n \sum_{K\in\mathcal{M}} \left(\frac{\widehat{v}_{L,K}^n}{|K|}\int_K g(x)\mathrm{d}x - \int_{V_{L,K}^n} g(X(t_{n+1},t_n,x))\mathrm{d}x \right).$$

Now, setting

$$\widetilde{R}^n(g) = \sum_{L\in\mathcal{M}}\sum_{K\in\mathcal{M}} u_K^n \left(\frac{\widehat{v}_{L,K}^n}{|K|}\int_K g(x)\mathrm{d}x - \int_{V_{L,K}^n} g(X(t_{n+1},t_n,x))\mathrm{d}x \right) \tag{55}$$

and using that $w_K^n = |\{x, X(t_{n+1},t_n,x) \in K\}|$, we get

$$\begin{aligned}\widetilde{R}^n(g) &= \sum_{K\in\mathcal{M}} u_K^n \left(\frac{w_K^n}{|K|}\int_K g(x)\mathrm{d}x - \int_{\{x,X(t_{n+1},t_n,x)\in K\}} g(X(t_{n+1},t_n,x))\mathrm{d}x \right) \\ &= -\sum_{K\in\mathcal{M}} u_K^n \int_{\{x,X(t_{n+1},t_n,x)\in K\}} \left(g(X(t_{n+1},t_n,x)) - \frac{1}{|K|}\int_K g(y)\mathrm{d}y \right)\mathrm{d}x.\end{aligned}$$

Applying Lemma 10, we may use the change of variable $y = X(t_{n+1},t_n,x)$ for the computation of $\int_{\mathbb{R}^N} \mathbb{1}_K(y)\mathrm{d}y$ and $\int_{\mathbb{R}^N} g(y)\mathbf{1}_K(y)\mathrm{d}y$. We thus get

$$|K| = \int_{\{x,X(t_{n+1},t_n,x)\in K\}} J(t_{n+1},t_n,x)\mathrm{d}x,$$

and

$$\int_K g(y)\mathrm{d}y = \int_{\{x,X(t_{n+1},t_n,x)\in K\}} g(X(t_{n+1},t_n,x))J(t_{n+1},t_n,x)\mathrm{d}x.$$

This leads to

$$\widetilde{R}^n(g) = \sum_{K\in\mathcal{M}} u_K^n \int_{\{x,X(t_{n+1},t_n,x)\in K\}} \left(g(X(t_{n+1},t_n,x)) - \frac{1}{|K|}\int_K g(y)\mathrm{d}y \right)(J(t_{n+1},t_n,x)-1)\,\mathrm{d}x.$$

Since $X(t_{n+1},t_n,x)\in K$ and $y\in K$, we have $|g(X(t_{n+1},t_n,x))-\frac{1}{|K|}\int_K g(y)\mathrm{d}y|\le \mathrm{Lip}(g)h_{\mathcal{M}}$. Hence, defining

$$|K|\,\Upsilon_K^n=\int_{\{x,X(t_{n+1},t_n,x)\in K\}}|J(t_{n+1},t_n,x)-1|\mathrm{d}x,$$

this allows to write

$$|\widetilde{R}^n(g)|\le \mathrm{Lip}(g)h_{\mathcal{M}}\sum_{K\in\mathcal{M}}|K|\,u_K^n\Upsilon_K^n. \tag{56}$$

We can now observe that

$$R^n(g)-\widetilde{R}^n(g)=\sum_{L\in\mathcal{M}}\sum_{K\in\mathcal{M}}(u_L^n-u_K^n)\left(\frac{\widehat{v}_{L,K}^n}{|K|}\int_K g(x)\mathrm{d}x-\int_{V_{L,K}^n}g(X(t_{n+1},t_n,x))\mathrm{d}x\right).$$

We then have $R^n(g)-\widetilde{R}^n(g)=T_1^n+T_2^n$ with

$$T_1^n=\sum_{L\in\mathcal{M}}\sum_{K\in\mathcal{M}}(u_L^n-u_K^n)\left(\frac{v_{L,K}^n}{|K|}\int_K g(x)\mathrm{d}x-\int_{V_{L,K}^n}g(X(t_{n+1},t_n,x))\mathrm{d}x\right),$$

and

$$T_2^n=\sum_{L\in\mathcal{M}}\sum_{K\in\mathcal{M}}(u_L^n-u_K^n)\frac{\widehat{v}_{L,K}^n-v_{L,K}^n}{|K|}\int_K g(x)\mathrm{d}x.$$

We then have

$$|T_1^n|\le\sum_{L\in\mathcal{M}}\sum_{K\in\mathcal{M}}|u_L^n-u_K^n|v_{L,K}^n\,\mathrm{Lip}(g)h_{\mathcal{M}}. \tag{57}$$

We now turn to the study of T_2^n. Since the term for $L=K$ vanishes in the above expressions, we can write

$$T_2^n=\sum_{L\in\mathcal{M}}\sum_{K\in\mathcal{M}}(u_L^n-u_K^n)\frac{\nu\,\tau|\sigma_{KL}|}{|K|}\int_K g(x)\mathrm{d}x.$$

We then get, gathering by pairs of control volumes,

$$T_2^n=\sum_{\{K,L\}}(u_L^n-u_K^n)\nu\,\tau|\sigma_{KL}|\left(\frac{1}{|K|}\int_K g(x)\mathrm{d}x-\frac{1}{|L|}\int_L g(x)\mathrm{d}x\right).$$

Since for all $\{K,L\}$ such that $|\sigma_{KL}|>0$, we have

$$\left|\frac{1}{|K|}\int_K g(x)\mathrm{d}x-\frac{1}{|L|}\int_L g(x)\mathrm{d}x\right|\le 2\,\mathrm{Lip}(g)h_{\mathcal{M}},$$

we get

$$|T_2^n|\le 2\,\mathrm{Lip}(g)h_{\mathcal{M}}\sum_{\{K,L\}}|u_L^n-u_K^n|\;\nu\;\tau|\sigma_{KL}|=\mathrm{Lip}(g)h_{\mathcal{M}}\sum_{L\in\mathcal{M}}\sum_{K\in\mathcal{M}}\nu\;\tau|\sigma_{KL}||u_L^n-u_K^n|. \tag{58}$$

Hence, from (57) and (58), we get

$$|R^n(g)-\widetilde{R}^n(g)|\le|T_1^n|+|T_2^n|\le\mathrm{Lip}(g)h_{\mathcal{M}}\sum_{L\in\mathcal{M}}\sum_{K\in\mathcal{M}}\widehat{v}_{L,K}^n|u_L^n-u_K^n|. \tag{59}$$

We now proceed as in the proof of Theorem 1: for a given $t>0$, we consider $n_t\in\mathbb{N}$ such that $n_t\tau<t\le(n_t+1)\tau$, and we again define for a given $f\in\mathrm{Lip}_c(\mathbb{R}^N)$ the functions $f_n:\;x\mapsto f(X(t_{n_t+1},t_{n+1},x))$, for $n=0,\dots,n_t$. We then define, for $n=0,\dots,n_t$, the terms $R^n(f_n)$ and $\widetilde{R}^n(f_n)$ by replacing, in (53),

(54) and (55), the function g by the function f_n, which is Lipschitz continuous with constant $\xi\,\mathrm{Lip}(f)$. Then (53) yields

$$\int_{\mathbb{R}^N} f(X(t_{n_t+1}, t_{n+1}, x))u_{\mathcal{M}}^{n+1}(x)\mathrm{d}x - \int_{\mathbb{R}^N} f(X(t_{n_t+1}, t_n, x))u_{\mathcal{M}}^{n}(x)\mathrm{d}x - R^n(f_n) = 0.$$

We sum all these equations for $n = 0, \ldots, n_t$. Since (28) again holds, we obtain

$$\left|\int_{\mathbb{R}^N} f(x)u_{\mathcal{M},\tau}(t,x)\mathrm{d}x - \int_{\mathbb{R}^N} f(X(t,0,x))\mathrm{d}\mu_0(x)\right| \le M_0\xi\,\mathrm{Lip}(f)(h_{\mathcal{M}} + \tau) + T_3^{\mathcal{M}} + T_4^{\mathcal{M}},$$

with, according to (59),

$$T_3^{\mathcal{M}} = \sum_{n=0}^{n_t} |R^n(f_n) - \widetilde{R}^n(f_n)| \le \xi\,\mathrm{Lip}(f)h_{\mathcal{M}} \sum_{n=0}^{n_t}\sum_{L\in\mathcal{M}}\sum_{K\in\mathcal{M}} \widehat{v}_{L,K}^n |u_L^n - u_K^n|, \tag{60}$$

and according to (56),

$$T_4^{\mathcal{M}} = \sum_{n=0}^{n_t} |\widetilde{R}^n(f_n)| \le \xi\,\mathrm{Lip}(f)h_{\mathcal{M}} \sum_{n=0}^{n_t}\sum_{K\in\mathcal{M}} |K|\, u_K^n \Upsilon_K^n. \tag{61}$$

The application of Lemma 4 with $T \ge t + \tau$ shows that

$$\lim_{h_{\mathcal{M}}\to 0} T_3^{\mathcal{M}} = 0,$$

and the application of Lemma 5 with $T \ge t + \tau$ provides

$$\lim_{h_{\mathcal{M}}\to 0} T_4^{\mathcal{M}} = 0.$$

This concludes the convergence proof.

□

It now remains to state and prove Lemmas 4 and 5, both of them being consequences of Lemma 3. This is the aim of the following sections.

3.3 A sharp weak BV estimate

Lemma 3 (Renormalized energy dissipation bound)**.** *Under Hypotheses 1 and 2, let $\rho > 0$, $\lambda \in (0,1)$, $\nu > 0$ and $\alpha > 0$ be given values. Let $\mathcal{M}$ be an admissible mesh of $\mathbb{R}^N$ in the sense of Definition 1, which is ρ-regular in the sense of Definition 2, and let τ be given such that the CFL condition* (46) *and the inverse CFL condition* (51) *hold. Let $(u_K^n)_{K\in\mathcal{M},n\in\mathbb{N}}$ be given by* (7)-(8) *and* (42)-(44)*, and let $T > 0$. Then there exist $C_1 \ge 0$, only depending on N, α, ξ, ζ_T, p_0, ρ, μ_0 and T, such that, denoting n_T the largest $n \in \mathbb{N}$ such that $n\tau \le T$,*

$$\sum_{n=0}^{n_T-1} \Lambda_n \le C_1\,, \tag{62}$$

where

$$\Lambda_n = \sum_{K\in\mathcal{M}}\sum_{L\in\mathcal{M}} \widehat{v}_{L,K}^n\ \mathrm{d}(u_K^{n+1}, u_L^n, \tfrac{m+1}{2})^2, \tag{63}$$

defining $m = 1/p_0$, and where, for any $\theta \in [0,1]$, we define the "renormalized distance" $\mathrm{d}(x,y,\theta)$ by

$$\mathrm{d}(x,y,\theta) = \frac{|y-x|}{\max(|x|,|y|)^\theta},\ \forall (x,y) \in \mathbb{R}^2 \setminus \{(0,0)\},$$
$$\mathrm{d}(0,0,\theta) = 0.$$

Proof. The scheme (44) may equivalently be written

$$\Big(|K| - w_K^n\Big)u_K^{n+1} + \sum_{L\in\mathcal{M}} \widehat{v}_{L,K}^n(u_K^{n+1} - u_L^n) = 0,\ \forall n \in \mathbb{N},\ \forall K \in \mathcal{M}. \tag{64}$$

Let $\Psi \in C^2(\mathbb{R})$ be a given convex function such that $\Psi(0) = 0$. Let us multiply (64) by $\Psi'(u_K^{n+1})$, for $n \in \mathbb{N}$ and $K \in \mathcal{M}$. We thus obtain

$$\Big(|K| - w_K^n\Big)u_K^{n+1}\Psi'(u_K^{n+1}) + \sum_{L\in\mathcal{M}} \widehat{v}_{L,K}^n\Psi'(u_K^{n+1})(u_K^{n+1} - u_L^n) = 0,$$

which provides, summing on $K \in \mathcal{M}$ (recall that the number of $K \in \mathcal{M}$ such that $u_K^n \neq 0$ is finite),

$$\sum_{K\in\mathcal{M}}\sum_{L\in\mathcal{M}} \widehat{v}_{L,K}^n\Psi'(u_K^{n+1})(u_K^{n+1} - u_L^n) = \sum_{K\in\mathcal{M}} \Big(w_K^n - |K|\Big)u_K^{n+1}\Psi'(u_K^{n+1}).$$

But since

$$\sum_{K\in\mathcal{M}}\sum_{L\in\mathcal{M}} \widehat{v}_{L,K}^n\Psi(u_L^n) = \sum_{L\in\mathcal{M}} |L|\Psi(u_L^n)$$

and

$$\sum_{K\in\mathcal{M}}\sum_{L\in\mathcal{M}} \widehat{v}_{L,K}^n\Psi(u_K^{n+1}) = \sum_{K\in\mathcal{M}} \Psi(u_K^{n+1})w_K^n,$$

we get

$$\begin{aligned}
&\sum_{K\in\mathcal{M}} |K|\Psi(u_K^{n+1}) - \sum_{L\in\mathcal{M}} |L|\Psi(u_L^n) \\
&\qquad + \sum_{K\in\mathcal{M}}\sum_{L\in\mathcal{M}} \widehat{v}_{L,K}^n(\Psi(u_L^n) - \Psi(u_K^{n+1}) + \Psi'(u_K^{n+1})(u_K^{n+1} - u_L^n)) \\
&\quad = \sum_{K\in\mathcal{M}} \Big(\Psi'(u_K^{n+1})u_K^{n+1} - \Psi(u_K^{n+1})\Big)(w_K^n - |K|).
\end{aligned}$$

This leads to

$$\begin{aligned}
&\sum_{K\in\mathcal{M}} |K|(\Psi(u_K^{n+1}) - \Psi(u_K^n)) \\
&\qquad + \frac{1}{2}\sum_{K\in\mathcal{M}}\sum_{L\in\mathcal{M}} \widehat{v}_{L,K}^n \min_{s\in[\min(u_K^{n+1},u_L^n),\max(u_K^{n+1},u_L^n)]} \Psi''(s)\ \ (u_K^{n+1} - u_L^n)^2 \\
&\quad \leq \sum_{K\in\mathcal{M}} \Big(\Psi'(u_K^{n+1})u_K^{n+1} - \Psi(u_K^{n+1})\Big)(w_K^n - |K|).
\end{aligned}$$

According to Lemma 1, that ensures the non-negativity of u_K^n, we can consider functions Ψ only defined on $[0,\infty)$. Hence we choose the function Ψ defined for given $m \in [0,1)$ and $\varepsilon > 0$, by

$$\Psi(s) = \frac{\varepsilon^{1-m} - (s+\varepsilon)^{1-m}}{1-m},\ \forall s \geq 0.$$

Then, for $s \geq 0$, we have $\Psi'(s) = -(s+\varepsilon)^{-m}$ and $\Psi''(s) = m/(s+\varepsilon)^{m+1}$, and we obtain, using that Ψ'' is non-increasing,

$$\begin{aligned}
&\frac{1}{1-m}\sum_{K\in\mathcal{M}} |K|((u_K^n + \varepsilon)^{1-m} - (u_K^{n+1} + \varepsilon)^{1-m}) \\
&\qquad + \frac{m}{2}\sum_{K\in\mathcal{M}}\sum_{L\in\mathcal{M}} \widehat{v}_{L,K}^n \frac{(u_K^{n+1} - u_L^n)^2}{(\max(u_K^{n+1}, u_L^n) + \varepsilon)^{1+m}} \\
&\quad \leq \sum_{K\in\mathcal{M}} (w_K^n - |K|)\left(-(u_K^{n+1} + \varepsilon)^{-m}\ u_K^{n+1} - \frac{1}{1-m}\big(\varepsilon^{1-m} - (u_K^{n+1} + \varepsilon)^{1-m}\big)\right).
\end{aligned}$$

It is then possible to let $\varepsilon \to 0$, thus obtaining

$$\frac{1}{1-m}\sum_{K\in\mathcal{M}}|K|(u_K^n)^{1-m}+\frac{m}{2}\Lambda_n \\ \leq \frac{1}{1-m}\sum_{K\in\mathcal{M}}|K|(u_K^{n+1})^{1-m}+\frac{m}{1-m}\sum_{K\in\mathcal{M}}(w_K^n-|K|)(u_K^{n+1})^{1-m}.$$

Therefore, summing over $n=0,\dots,n_T-1$ gives

$$\frac{m}{2}\sum_{n=0}^{n_T-1}\Lambda_n \leq \frac{1}{1-m}\sum_{K\in\mathcal{M}}|K|(u_K^{n_T})^{1-m}+\frac{m}{1-m}\sum_{n=0}^{n_T-1}\sum_{K\in\mathcal{M}}(w_K^n-|K|)(u_K^{n+1})^{1-m}. \tag{65}$$

Now, in order to estimate the right-hand side of (65), we write for all $n=0,\dots,n_T-1$,

$$\begin{aligned}\sum_{K\in\mathcal{M}}|w_K^n-|K||\,(u_K^{n+1})^{1-m} &\leq \left(\sum_{K\in\mathcal{M}}|K|\left|\frac{w_K^n-|K|}{|K|}\right|^{1/m}\right)^m\left(\sum_{K\in\mathcal{M}}|K|\;u_K^{n+1}\right)^{1-m}\\ &= \left(\sum_{K\in\mathcal{M}}\frac{1}{|K|^{1/m-1}}\,|w_K^n-|K||^{1/m}\right)^m {M_0}^{1-m}.\end{aligned} \tag{66}$$

We can write, according to the change of variable formula given in Lemma 10 with $f(x)=\mathbf{1}_K(x)$,

$$w_K^n-|K|=\int_{\{x,X(t_{n+1},t_n,x)\in K\}}(1-J(t_{n+1},t_n,x))\,\mathrm{d}x.$$

Thus

$$|w_K^n-|K||\leq (w_K^n)^{1-m}\left(\int_{\{x,X(t_{n+1},t_n,x)\in K\}}|J(t_{n+1},t_n,x)-1|^{1/m}\mathrm{d}x\right)^m.$$

Therefore, using Lemma 1, we have

$$\frac{1}{|K|^{1/m-1}}\,|w_K^n-|K||^{1/m}\leq \beta^{1/m-1}\int_{\{x,X(t_{n+1},t_n,x)\in K\}}|J(t_{n+1},t_n,x)-1|^{1/m}\mathrm{d}x.$$

This gives

$$\sum_{K\in\mathcal{M}}\frac{1}{|K|^{1/m-1}}\,|w_K^n-|K||^{1/m}\leq \beta^{1/m-1}\int_{\mathbb{R}^N}|J(t_{n+1},t_n,x)-1|^{1/m}\mathrm{d}x.$$

Therefore we get from (66)

$$\sum_{K\in\mathcal{M}}|w_K^n-|K||\,(u_K^{n+1})^{1-m}\leq(\beta\,M_0)^{1-m}\|J(t_{n+1},t_n,x)-1\|_{L^{1/m}(\mathbb{R}^N)}.$$

We now use the value of $m=\frac{1}{p_0}$ given in the statement of the lemma. Then, using Hypothesis 2,

$$\sum_{n=0}^{n_T-1}\sum_{K\in\mathcal{M}}|w_K^n-|K||\,(u_K^{n+1})^{1-1/p_0}\leq(\beta\,M_0)^{(1-1/p_0)}\zeta_T T. \tag{67}$$

Finally, we remark that, applying Lemma 2,

$$\begin{aligned}\sum_{K\in\mathcal{M}}|K|(u_K^{n_T})^{1-1/p_0} &\leq \sum_{K\in\mathcal{M},K\subset\bar{B}(0,R)}|K|(u_K^{n_T})^{1-1/p_0}\\ &\leq \left(\sum_{K\in\mathcal{M},K\subset\bar{B}(0,R)}|K|\right)^{1/p_0}\left(\sum_{K\in\mathcal{M},K\subset\bar{B}(0,R)}|K|u_K^{n_T}\right)^{1-1/p_0}\\ &\leq |\bar{B}(0,R)|^{1/p_0}M_0^{1-1/p_0}.\end{aligned}$$

Introducing Inequality (67) and the above inequality in (65) concludes the proof of the lemma.

□

Lemma 4 (Weak BV estimate)**.** *Under the hypotheses of Lemma 3, there exist $C_2 \geq 0$ and $r > 0$, only depending on N, λ, ν, α, ξ, ζ_T, p_0, ρ, μ_0 and T such that*

$$h_{\mathcal{M}} \sum_{n=0}^{n_T-1} \sum_{K\in\mathcal{M}} \sum_{L\in\mathcal{M}} \widehat{v}_{L,K}^n |u_K^n - u_L^n| \leq C_2 \, (h_{\mathcal{M}})^r. \tag{68}$$

Proof. In this proof, we denote by C_i various non-negative values, only depending on N, λ, ν, α, ξ, ζ_T, p_0, ρ, μ_0 and T. Since

$$|u_K^n - u_L^n| = \max(u_K^n, u_L^n)^{(m+1)/2} \mathrm{d}(u_K^n, u_L^n, \tfrac{m+1}{2}),$$

we have

$$\begin{aligned} &\sum_{K\in\mathcal{M}} \sum_{L\in\mathcal{M}} \widehat{v}_{L,K}^n |u_K^n - u_L^n| \\ &\qquad \leq \left(\sum_{K\in\mathcal{M}} \sum_{L\in\mathcal{M}} \widehat{v}_{L,K}^n \max(u_K^n, u_L^n)^{m+1} \right)^{1/2} \left(\sum_{K\in\mathcal{M}} \sum_{L\in\mathcal{M}} \widehat{v}_{L,K}^n \mathrm{d}(u_K^n, u_L^n, \tfrac{m+1}{2})^2 \right)^{1/2}. \end{aligned} \tag{69}$$

Then, using Lemma 1,

$$\begin{aligned} \sum_{K\in\mathcal{M}} \sum_{L\in\mathcal{M}} \widehat{v}_{L,K}^n \max(u_K^n, u_L^n)^{m+1} &\leq \sum_{K\in\mathcal{M}} \sum_{L\in\mathcal{M}} \widehat{v}_{L,K}^n ((u_K^n)^{m+1} + (u_L^n)^{m+1}) \\ &= \sum_{K\in\mathcal{M}} w_K^n (u_K^n)^{m+1} + \sum_{L\in\mathcal{M}} |L| (u_L^n)^{m+1} \\ &\leq (\beta+1) \sum_{K\in\mathcal{M}} |K| (u_K^n)^{m+1}. \end{aligned} \tag{70}$$

According to Lemma 8, we can write, for a given $n \in \mathbb{N}$,

$$\begin{aligned} \sum_{K\in\mathcal{M}} \sum_{L\in\mathcal{M}} \widehat{v}_{L,K}^n \mathrm{d}(u_K^n, u_L^n, \tfrac{m+1}{2})^2 &\leq 2 \sum_{K\in\mathcal{M}} \sum_{L\in\mathcal{M}} \widehat{v}_{L,K}^n \left(\mathrm{d}(u_K^n, u_K^{n+1}, \tfrac{m+1}{2})^2 + \mathrm{d}(u_K^{n+1}, u_L^n, \tfrac{m+1}{2})^2 \right), \\ &= 2 \sum_{K\in\mathcal{M}} w_K^n \, \mathrm{d}(u_K^{n+1}, u_K^n, \tfrac{m+1}{2})^2 + 2\Lambda_n. \end{aligned} \tag{71}$$

Again using Lemma 1, we get

$$\sum_{K\in\mathcal{M}} w_K^n \, \mathrm{d}(u_K^{n+1}, u_K^n, \tfrac{m+1}{2})^2 \leq \frac{\beta}{\lambda} \sum_{K\in\mathcal{M}} \widehat{v}_{K,K}^n \, \mathrm{d}(u_K^{n+1}, u_K^n, \tfrac{m+1}{2})^2 \leq \frac{\beta}{\lambda} \Lambda_n.$$

Therefore, we obtain the existence of C_3, only depending on λ, ρ and N, such that

$$\sum_{K\in\mathcal{M}} \sum_{L\in\mathcal{M}} \widehat{v}_{L,K}^n \mathrm{d}(u_K^n, u_L^n, \tfrac{m+1}{2})^2 \leq C_3 \, \Lambda_n. \tag{72}$$

Thus, gathering (69), (70) and (72), we get

$$\sum_{K\in\mathcal{M}} \sum_{L\in\mathcal{M}} \widehat{v}_{L,K}^n |u_K^n - u_L^n| \leq \left((\beta+1) \sum_{K\in\mathcal{M}} |K| (u_K^n)^{m+1} \right)^{1/2} (C_3 \, \Lambda_n)^{1/2}. \tag{73}$$

We let again $m = 1/p_0$. Since $p_0 > N$, it satisfies

$$0 < m < 1, \qquad m < \frac{2}{N}, \qquad m \leq \frac{1}{N-1}, \tag{74}$$

and we may apply Lemma 6 with $\widehat{u}_K = u_K^n$, $q = 2$, $p = 1 + m = 1 + 1/p_0 > 1$. Indeed, in the case $N > 2$, the last inequality of (74) yields the condition in Lemma 6 that $p \le (1-m)/(1-\frac{2}{N})$. We thus get

$$\left(\sum_{K\in\mathcal{M}} |K|(u_K^n)^{1+m}\right)^{1/(1+m)} \le C_4 \left(\sum_{K\in\mathcal{M}} |K|u_K^n\right)^{1-\theta} \left(\sum_{K\in\mathcal{M}}\sum_{L\in\mathcal{N}_K} |\sigma_{KL}|\, h_{\mathcal{M}}^{-1}\, \mathrm{d}\big(u_K^n, u_L^n, \tfrac{m+1}{2}\big)^2\right)^{\frac{\theta}{1-m}}.$$

Since $\nu\,\tau|\sigma_{KL}| \le \widehat{v}_{L,K}^n$ for $K \ne L$, and $\mathrm{d}\big(u_K^n, u_L^n, \frac{m+1}{2}\big) = 0$ for $K = L$, we have

$$\left(\sum_{K\in\mathcal{M}} |K|(u_K^n)^{1+m}\right)^{1/(1+m)} \le C_4\; {M_0}^{1-\theta} \left(\sum_{K\in\mathcal{M}}\sum_{L\in\mathcal{N}_K} \widehat{v}_{L,K}^n\,(h_{\mathcal{M}}\nu\tau)^{-1}\, \mathrm{d}\big(u_K^n, u_L^n, \tfrac{m+1}{2}\big)^2\right)^{\frac{\theta}{1-m}}.$$

Using (73) and (72), we obtain

$$\sum_{K\in\mathcal{M}}\sum_{L\in\mathcal{M}} \widehat{v}_{L,K}^n |u_K^n - u_L^n| \le C_5\, {M_0}^{1-\frac{1-m}{2-mN}}\, (\Lambda_n)^{1/2} \left(\frac{\Lambda_n}{h_{\mathcal{M}}\tau}\right)^{\frac{1}{2-mN}-\frac{1}{2}}.$$

Then, since $m \le \frac{1}{N}$, we have $2 - mN \ge 1$. According to Hölder's inequality, we can write

$$\sum_{n=0}^{n_T-1}\sum_{K\in\mathcal{M}}\sum_{L\in\mathcal{M}} \widehat{v}_{L,K}^n |u_K^n - u_L^n| \le C_6 \left(\frac{T}{\tau}\right)^{1-\frac{1}{2-mN}} \left(\sum_{n=0}^{n_T-1}\frac{\Lambda_n}{h_{\mathcal{M}}}\right)^{\frac{1}{2-mN}} {h_{\mathcal{M}}}^{\frac{1}{2}} \left(\frac{1}{\tau}\right)^{\frac{1}{2-mN}-\frac{1}{2}},$$

which gives, using the inverse CFL condition (51),

$$h_{\mathcal{M}} \sum_{n=0}^{n_T-1}\sum_{K\in\mathcal{M}}\sum_{L\in\mathcal{M}} \widehat{v}_{L,K}^n |u_K^n - u_L^n| \le C_7 \left(\sum_{n=0}^{n_T-1}\Lambda_n\right)^{\frac{1}{2-mN}} (h_{\mathcal{M}})^{1-\frac{1}{2-mN}}. \tag{75}$$

According to the estimate (62) proved in Lemma 3, the above inequality provides the estimate (68) with $r = 1 - \frac{1}{2-N/p_0} > 0$. □

Remark 5. *In order to get* (75)*, the inverse CFL condition* (51) *can be avoided if one makes some suitable assumptions so that* $\widehat{v}_{L,K}^n \le C\tau|\sigma_{KL}|$ *for* $K \ne L$*, that allows to improve* (70) *by a factor* $\tau/h_{\mathcal{M}}$*.*

Lemma 5 (Estimate of the term due to compression effects)**.** *Under the hypotheses of Lemma 3, there exists* $C_8 \ge 0$ *and* $r > 0$*, only depending on* N*,* λ*,* ν*,* α*,* ξ*,* ζ_T*,* p_0*,* ρ*,* μ_0 *and* T *such that*

$$h_{\mathcal{M}} \sum_{n=0}^{n_T-1}\sum_{K\in\mathcal{M}} |K|\, u_K^n \Upsilon_K^n \le C_8\,(h_{\mathcal{M}})^r, \tag{76}$$

with

$$|K|\Upsilon_K^n = \int_{\{x, X(t_{n+1},t_n,x)\in K\}} |J(t_{n+1},t_n,x) - 1|\,\mathrm{d}x.$$

Proof. In this proof, we again denote by C_i various non-negative values, only depending on N, λ, ν, α, ξ, ζ_T, p_0, ρ, μ_0 and T. Let $\overline{p} \ge p_0$ be given (it will be chosen later), and let again $m = 1/p_0$. Since J is bounded by ξ^N and $\overline{p} \ge p_0$, we can write

$$\|J(t_{n+1},t_n,\cdot) - 1\|_{L^{\overline{p}}(\mathbb{R}^N)} \le (\xi^N + 1)^{1-p_0/\overline{p}}\, \|J(t_{n+1},t_n,\cdot) - 1\|_{L^{p_0}(\mathbb{R}^N)}^{p_0/\overline{p}}. \tag{77}$$

For any $K \in \mathcal{M}$, we have

$$|K|^{\overline{p}}(\Upsilon_K^n)^{\overline{p}} \le (w_K^n)^{\overline{p}-1} \int_{\{x, X(t_{n+1},t_n,x)\in K\}} |J(t_{n+1},t_n,x) - 1|^{\overline{p}}\mathrm{d}x.$$

Thus, using (49),

$$|K|(\Upsilon_K^n)^{\overline{p}} \le \beta^{\overline{p}-1} \int_{\{x, X(t_{n+1},t_n,x)\in K\}} |J(t_{n+1},t_n,x)-1|^{\overline{p}} \mathrm{d}x,$$

which enables to estimate

$$\begin{aligned}
\sum_{K\in\mathcal{M}} |K|\, u_K^n \Upsilon_K^n &\le \left(\sum_{K\in\mathcal{M}} |K|(\Upsilon_K^n)^{\overline{p}}\right)^{1/\overline{p}} \left(\sum_{K\in\mathcal{M}} |K|(u_K^n)^{\frac{\overline{p}}{\overline{p}-1}}\right)^{1-1/\overline{p}} \\
&\le C_9 \left\|J(t_{n+1},t_n,\cdot)-1\right\|_{L^{\overline{p}}(\mathbb{R}^N)} \left(\sum_{K\in\mathcal{M}} |K|(u_K^n)^{\frac{\overline{p}}{\overline{p}-1}}\right)^{1-1/\overline{p}} \\
&\le C_{10} \left\|J(t_{n+1},t_n,.)-1\right\|_{L^{p_0}(\mathbb{R}^N)}^{p_0/\overline{p}} \left(\sum_{K\in\mathcal{M}} |K|(u_K^n)^{\frac{\overline{p}}{\overline{p}-1}}\right)^{1-1/\overline{p}}.
\end{aligned} \tag{78}$$

In the case $N>2$, since $\overline{p}\ge p_0>N$, the real $\overline{p}'$ such that $1/\overline{p}+1/\overline{p}'=1$ satisfies $\overline{p}'<1+1/(N-1)<1+1/(N-2)$. Hence for $m=1/p_0$,

$$\overline{p}' < 1+\frac{1}{N-2} < \left(1-\frac{1}{p_0}\right)\frac{N}{N-2} = (1-m)/(1-2/N).$$

Therefore, we can apply Lemma 6 with $q=2$, $m=1/p_0$, $p=\overline{p}'$, $\widehat{u}_K=u_K^n$, thus obtaining

$$\left(\sum_{K\in\mathcal{M}} |K|(u_K^n)^{\frac{\overline{p}}{\overline{p}-1}}\right)^{1-1/\overline{p}} \le C_{14} \left(\sum_{K\in\mathcal{M}} |K|u_K^n\right)^{1-\theta} \left(\sum_{K\in\mathcal{M}}\sum_{L\in\mathcal{N}_K} |\sigma_{KL}|\, h_{\mathcal{M}}^{-1}\, \mathrm{d}\big(u_K^n, u_L^n, \tfrac{m+1}{2}\big)^2\right)^{\frac{\theta}{1-m}},$$

with

$$\theta = \frac{1-m}{\overline{p}}\frac{1}{2/N-m}. \tag{79}$$

As in the proof of Lemma 4 and (72), it gives

$$\left(\sum_{K\in\mathcal{M}} |K|(u_K^n)^{\frac{\overline{p}}{\overline{p}-1}}\right)^{1-1/\overline{p}} \le C_{11}\, {M_0}^{1-\theta} \left(\frac{\Lambda_n}{h_{\mathcal{M}}\tau}\right)^{\frac{\theta}{1-m}},$$

which yields

$$\sum_{K\in\mathcal{M}} |K|\, u_K^n \Upsilon_K^n \le C_{12} \left\|J(t_{n+1},t_n,\cdot)-1\right\|_{L^{p_0}(\mathbb{R}^N)}^{p_0/\overline{p}} \left(\frac{\Lambda_n}{h_{\mathcal{M}}\tau}\right)^{\frac{\theta}{1-m}}.$$

Recalling (79), we now set the value of $\overline{p}\ge p_0$ by

$$\frac{\theta}{1-m}+\frac{p_0}{\overline{p}}=1,$$

which provides

$$\overline{p} = p_0 + \frac{1}{2/N-m} = \frac{p_0}{1-N/2p_0} \quad \text{and} \quad \frac{\theta}{1-m} = \frac{N}{2p_0}.$$

We then get

$$\begin{aligned}
&h_{\mathcal{M}} \sum_{n=0}^{n_T-1} \sum_{K\in\mathcal{M}} |K|\, u_K^n \Upsilon_K^n \\
&\le C_{12}\, h_{\mathcal{M}} \left(\sum_{n=0}^{n_T-1} \left\|J(t_{n+1},t_n,\cdot)-1\right\|_{L^{p_0}(\mathbb{R}^N)}\right)^{p_0/\overline{p}} \left(\sum_{n=0}^{n_T-1} \frac{\Lambda_n}{h_{\mathcal{M}}\tau}\right)^{\frac{\theta}{1-m}} \\
&\le C_{13}\, h_{\mathcal{M}} \left(\sum_{n=0}^{n_T-1} \frac{\Lambda_n}{h_{\mathcal{M}}\tau}\right)^{N/2p_0},
\end{aligned} \tag{80}$$

which gives, according to the inverse CFL condition (51),

$$h_{\mathcal{M}} \sum_{n=0}^{n_T-1} \sum_{K\in\mathcal{M}} |K|\, u_K^n \Upsilon_K^n \le C_{13} \left(\alpha \sum_{n=0}^{n_T-1} \Lambda_n\right)^{N/2p_0} {h_{\mathcal{M}}}^{1-N/p_0}.$$

Using the bound (62) proved in Lemma 3, the above inequality provides (76) with $r = 1 - N/p_0 > 0$. □

4 Technical results

In this section we gather several technical lemmas.

Lemma 6 (Improved discrete Gagliardo-Nirenberg-Sobolev inequality)**.** *Let $N \in \mathbb{N}^\star$, $q \in [1,\infty)$, m a real number such that $m < 1$, $m < q/N$ be given. Let $p \in [1,\infty)$ and assume in the case when $q < N$ that $p \le (1-m)/(1-q/N)$. Let θ be such that*

$$\frac{1}{p} = 1 - \theta + \theta\frac{1-q/N}{1-m}, \text{ or equivalently } \theta = (1-m)\frac{1-1/p}{q/N-m}, \tag{81}$$

which within our assumptions implies $\theta \in [0,1]$.

Let $\rho > 0$ be given and let $\mathcal{M}$ be an admissible mesh of $\mathbb{R}^N$ in the sense of Definition 1, which is ρ-regular in the sense of Definition 2. Let $(\widehat{u}_K)_{K\in\mathcal{M}}$ be a family of real numbers such that the number of $K \in \mathcal{M}$ such that $\widehat{u}_K \neq 0$ is finite. Then one has

$$\left(\sum_{K\in\mathcal{M}} |K||\widehat{u}_K|^p\right)^{1/p} \le C_{14} \left(\sum_{K\in\mathcal{M}} |K||\widehat{u}_K|\right)^{1-\theta} \left(\sum_{K\in\mathcal{M}} \sum_{L\in\mathcal{N}_K} |\sigma_{KL}|\, h_{\mathcal{M}}^{1-q}\, \mathrm{d}\Big(\widehat{u}_K, \widehat{u}_L, \tfrac{m+q-1}{q}\Big)^q\right)^{\frac{\theta}{1-m}}, \tag{82}$$

where C_{14} only depends on N, q, m, p, and on ρ.

Remark 6. *At the continuous level, the improved Gagliardo-Nirenberg-Sobolev inequality writes as follows, under the same assumptions on the exponents,*

$$\|u\|_{L^p(\mathbb{R}^N)} \le C_{15}\, \|u\|_{L^1(\mathbb{R}^N)}^{1-\theta} \left(\int_{\mathbb{R}^N} \frac{|\nabla u|^q}{|u|^{m+q-1}}\mathrm{d}x\right)^{\frac{\theta}{1-m}}, \quad \text{for all } u \in C_c^1(\mathbb{R}^N),$$

where C_{15} depends only on N, q, m, p (the proof is omitted).

Proof of Lemma 6. In this proof, we denote by C_i various non-negative constants which only depend on N, q, m, p and on ρ. Let $\gamma > 0$, that will be chosen later on. We define the function $\widehat{v}(x)$ for $x \in \mathbb{R}^N$ by $\widehat{v}(x) = \widehat{v}_K$ for almost every $x \in K$, all $K \in \mathcal{M}$, where $\widehat{v}_K = |\widehat{u}_K|^\gamma$ for all $K \in \mathcal{M}$. According to the Sobolev inequality, we have

$$\|\widehat{v}\|_{L^{N/(N-1)}(\mathbb{R}^N)} \le C_N |\widehat{v}|_{BV(\mathbb{R}^N)} \tag{83}$$

(which is also true for $N = 1$), where

$$|\widehat{v}|_{BV(\mathbb{R}^N)} = \sup\left\{\left|\int_{\mathbb{R}^N} \widehat{v}(x)\,\mathrm{div}\,\varphi(x)\,\mathrm{d}x\right|,\ \varphi \in C_c^1(\mathbb{R}^N),\ \|\varphi\|_\infty \le 1\right\}.$$

In the particular case of the piecewise constant function $\widehat{v}$, we have

$$|\widehat{v}|_{BV(\mathbb{R}^N)} = \frac{1}{2} \sum_{K\in\mathcal{M}} \sum_{L\in\mathcal{N}_K} |\sigma_{KL}|\, |\widehat{v}_K - \widehat{v}_L|, \tag{84}$$

the factor $1/2$ resulting from the fact that each interface appears twice in the double sum. According to Lemma 7, one has $\big||\widehat{u}_K|^\gamma - |\widehat{u}_L|^\gamma\big| \le \max(1,\gamma)\max(|\widehat{u}_K|,|\widehat{u}_L|)^{\gamma-1}|\widehat{u}_K - \widehat{u}_L|$, thus

$$|\widehat{v}|_{BV(\mathbb{R}^N)} \le \frac{\max(1,\gamma)}{2} \sum_{K\in\mathcal{M}} \sum_{L\in\mathcal{N}_K} |\sigma_{KL}|\, \max(|\widehat{u}_K|,|\widehat{u}_L|)^{\gamma-(1-m)/q} \mathrm{d}\Big(\widehat{u}_K, \widehat{u}_L, \tfrac{m+q-1}{q}\Big). \tag{85}$$

Let us first consider the case $q = 1$. We then take $\gamma = 1 - m$, and (85), together with (83), gives

$$\||\widehat{u}|^{1-m}\|_{L^{N/(N-1)}(\mathbb{R}^N)} \le \frac{C_N}{2}\max(1,\gamma)\left(\sum_{K\in\mathcal{M}} \sum_{L\in\mathcal{N}_K} |\sigma_{KL}|\, \mathrm{d}(\widehat{u}_K, \widehat{u}_L, m)\right). \tag{86}$$

If $N = 1$, it yields the existence of C_{16} such that

$$\sup_{K\in\mathcal{M}} |\widehat{u}_K| \le C_{16}\left(\sum_{K\in\mathcal{M}} \sum_{L\in\mathcal{N}_K} |\sigma_{KL}|\, \mathrm{d}(\widehat{u}_K, \widehat{u}_L, m)\right)^{1/(1-m)},$$

which means that (82) holds for $p = \infty$ (and $\theta = 1$). Since $1/p = (1-\theta)/1 + \theta/\infty$, one has

$$\|\widehat{u}\|_{L^p(\mathbb{R}^N)} \le \|\widehat{u}\|_{L^1(\mathbb{R}^N)}{}^{1-\theta} \|\widehat{u}\|_{L^\infty(\mathbb{R}^N)}{}^{\theta},$$

which implies that (82) holds for all $1 \le p \le \infty$.

Next, if $N > 1$ (still with $q = 1$), (86) means that (82) holds for $p = (1-m)N/(N-1)$ (and $\theta = 1$). Since $1/p = (1-\theta)/1 + \theta/((1-m)N/(N-1))$, one has

$$\|\widehat{u}\|_{L^p(\mathbb{R}^N)} \le \|\widehat{u}\|_{L^1(\mathbb{R}^N)}{}^{1-\theta} \|\widehat{u}\|_{L^{(1-m)N/(N-1)}(\mathbb{R}^N)}{}^{\theta},$$

which implies that (82) holds for all $1 \le p \le (1-m)N/(N-1)$, proving the result.

From now on, we assume that $q > 1$. Then, from (85) we get

$$\begin{aligned} |\widehat{v}|_{BV(\mathbb{R}^N)} \le \frac{\max(1,\gamma)}{2} &\left(\sum_{K\in\mathcal{M}} \sum_{L\in\mathcal{N}_K} |\sigma_{KL}| \left(\max(|\widehat{u}_K|,|\widehat{u}_L|)^{\gamma-(1-m)/q}\right)^{q/(q-1)}\right)^{1-1/q} \\ &\times \left(\sum_{K\in\mathcal{M}} \sum_{L\in\mathcal{N}_K} |\sigma_{KL}|\, \mathrm{d}\Big(\widehat{u}_K, \widehat{u}_L, \tfrac{m+q-1}{q}\Big)^q\right)^{1/q}. \end{aligned} \tag{87}$$

We write

$$\begin{aligned} &\sum_{K\in\mathcal{M}} \sum_{L\in\mathcal{N}_K} |\sigma_{KL}| \left(\max(|\widehat{u}_K|,|\widehat{u}_L|)^{\gamma-(1-m)/q}\right)^{q/(q-1)} \\ &\quad\le \sum_{K\in\mathcal{M}} \sum_{L\in\mathcal{N}_K} |\sigma_{KL}| \left(|\widehat{u}_K|^{(\gamma-(1-m)/q)q/(q-1)} + |\widehat{u}_L|^{(\gamma-(1-m)/q)q/(q-1)}\right) \\ &\quad= 2\sum_{K\in\mathcal{M}} \sum_{L\in\mathcal{N}_K} |\sigma_{KL}|\, |\widehat{u}_K|^{(\gamma-(1-m)/q)q/(q-1)} \\ &\quad\le 2\frac{\rho}{h_\mathcal{M}} \sum_{K\in\mathcal{M}} |K||\widehat{u}_K|^{(\gamma-(1-m)/q)q/(q-1)}, \end{aligned} \tag{88}$$

where we used (39). Therefore, (87), together with (83), (88), yield

$$\begin{aligned} \||\widehat{u}|^\gamma\|_{L^{N/(N-1)}(\mathbb{R}^N)} \le C_N\, \max(1,\gamma) &\left(\sum_{K\in\mathcal{M}} |K||\widehat{u}_K|^{(\gamma-(1-m)/q)q/(q-1)}\right)^{1-1/q} \\ &\times \left(\left(\frac{\rho}{h_\mathcal{M}}\right)^{q-1} \sum_{K\in\mathcal{M}} \sum_{L\in\mathcal{N}_K} |\sigma_{KL}|\, \mathrm{d}\Big(\widehat{u}_K, \widehat{u}_L, \tfrac{m+q-1}{q}\Big)^q\right)^{1/q}. \end{aligned} \tag{89}$$

Let us first consider the case $q < N$ (that implies $N > 1$). Then we choose $\gamma > 0$ by the relation (all the factors are positive)

$$\gamma\left(\frac{q}{q-1} - \frac{N}{N-1}\right) = \frac{1-m}{q-1}.$$

This choice gives

$$\left(\gamma - \frac{1-m}{q}\right)\frac{q}{q-1} = \gamma\frac{N}{N-1} = \frac{1-m}{1-q/N},$$

and (89) becomes

$$\begin{aligned}\left(\sum_{K\in\mathcal{M}} |K||\widehat{u}_K|^{\gamma N/(N-1)}\right)^{1-1/N} &\le C_N\,\max(1,\gamma)\left(\sum_{K\in\mathcal{M}} |K||\widehat{u}_K|^{(1-m)/(1-q/N)}\right)^{1-1/q}\\ &\quad\times\left(\left(\frac{\rho}{h_{\mathcal{M}}}\right)^{q-1}\sum_{K\in\mathcal{M}}\sum_{L\in\mathcal{N}_K}|\sigma_{KL}|\,\mathrm{d}\Big(\widehat{u}_K,\widehat{u}_L,\tfrac{m+q-1}{q}\Big)^q\right)^{1/q},\end{aligned} \tag{90}$$

that is

$$\begin{aligned}&\left(\sum_{K\in\mathcal{M}} |K||\widehat{u}_K|^{(1-m)/(1-q/N)}\right)^{(1-q/N)/(1-m)}\\ &\qquad\le (C_N\,\max(1,\gamma))^{q/(1-m)}\left(\left(\frac{\rho}{h_{\mathcal{M}}}\right)^{q-1}\sum_{K\in\mathcal{M}}\sum_{L\in\mathcal{N}_K}|\sigma_{KL}|\,\mathrm{d}\Big(\widehat{u}_K,\widehat{u}_L,\tfrac{m+q-1}{q}\Big)^q\right)^{1/(1-m)}.\end{aligned}$$

This inequality means that (82) holds for $p = (1-m)/(1-q/N)$ (and $\theta = 1$). Since $1/p = (1-\theta)/1 + \theta/[(1-m)/(1-q/N)]$, one has

$$\|\widehat{u}\|_{L^p(\mathbb{R}^N)} \le \|\widehat{u}\|_{L^1(\mathbb{R}^N)}{}^{1-\theta}\,\|\widehat{u}\|_{L^{(1-m)/(1-q/N)}(\mathbb{R}^N)}{}^{\theta},$$

that enables to conclude that (82) holds for all $1 \le p \le (1-m)/(1-q/N)$, proving the result.

Consider now the case $q \ge N$, still with $q > 1$. Then for all $\gamma > 0$

$$\left(\gamma - \frac{1-m}{q}\right)\frac{q}{q-1} < \gamma\frac{N}{N-1}. \tag{91}$$

We consider coefficients $\gamma > 0$ satisfying

$$r \equiv \left(\gamma - \frac{1-m}{q}\right)\frac{q}{q-1} \ge 1. \tag{92}$$

Assume first that $N = 1$. Then one has

$$\left(\sum_{K\in\mathcal{M}} |K||\widehat{u}_K|^r\right)^{1/r} \le \left(\sum_{K\in\mathcal{M}} |K||\widehat{u}_K|\right)^{1/r}\left(\sup_{K\in\mathcal{M}}|\widehat{u}_K|\right)^{1-1/r},$$

which yields in (89)

$$\begin{aligned}\left(\sup_{K\in\mathcal{M}}|\widehat{u}_K|\right)^{\gamma-r(1-1/q)(1-1/r)} &\le C_N\,\max(1,\gamma)\left(\sum_{K\in\mathcal{M}} |K||\widehat{u}_K|\right)^{1-1/q}\\ &\quad\times\left(\left(\frac{\rho}{h_{\mathcal{M}}}\right)^{q-1}\sum_{K\in\mathcal{M}}\sum_{L\in\mathcal{N}_K}|\sigma_{KL}|\,\mathrm{d}\Big(\widehat{u}_K,\widehat{u}_L,\tfrac{m+q-1}{q}\Big)^q\right)^{1/q}.\end{aligned}$$

Since $\gamma - r(1-1/q)(1-1/r) = 1 - m/q$, one gets

$$\sup_{K\in\mathcal{M}} |\widehat{u}_K| \le (C_N \max(1,\gamma))^{q/(q-m)} \left(\sum_{K\in\mathcal{M}} |K||\widehat{u}_K|\right)^{(q-1)/(q-m)} \times \left(\left(\frac{\rho}{h_{\mathcal{M}}}\right)^{q-1} \sum_{K\in\mathcal{M}} \sum_{L\in\mathcal{N}_K} |\sigma_{KL}|\, \mathrm{d}\Big(\widehat{u}_K, \widehat{u}_L, \tfrac{m+q-1}{q}\Big)^q\right)^{1/(q-m)}.$$

and one can verify, for $p=\infty$, that $1-\theta = (q-1)/(q-m)$ and $\theta/(1-m) = 1/(q-m)$ so that it means that (82) holds for $p=\infty$ (take any γ satisfying (92)). Since $1/p = (1/p)/1 + (1-1/p)/\infty$, we have

$$\|\widehat{u}\|_{L^p(\mathbb{R}^N)} \le \|\widehat{u}\|_{L^1(\mathbb{R}^N)}{}^{1/p} \|\widehat{u}\|_{L^\infty(\mathbb{R}^N)}{}^{1-1/p},$$

thus we get

$$\left(\sum_{K\in\mathcal{M}} |K||\widehat{u}_K|^p\right)^{1/p} \le (C_N \max(1,\gamma))^{(1-1/p)q/(q-m)} \left(\sum_{K\in\mathcal{M}} |K||\widehat{u}_K|\right)^{1/p+(1-1/p)(q-1)/(q-m)} \times \left(\left(\frac{\rho}{h_{\mathcal{M}}}\right)^{q-1} \sum_{K\in\mathcal{M}} \sum_{L\in\mathcal{N}_K} |\sigma_{KL}|\, \mathrm{d}\Big(\widehat{u}_K, \widehat{u}_L, \tfrac{m+q-1}{q}\Big)^q\right)^{(1-1/p)/(q-m)},$$

and one has

$$\frac{1}{p} = \left[\frac{1}{p} + \left(1-\frac{1}{p}\right)\frac{q-1}{q-m}\right] + (1-q)\left[\left(1-\frac{1}{p}\right)\frac{1}{q-m}\right],$$

hence we conclude again that (82) holds for all $1 \le p \le \infty$.

Consider finally the case $q \ge N > 1$. If $p=1$, (82) is obvious. So we assume that $\gamma N/(N-1) > 1$ (we will let $p = \gamma N/(N-1)$). According to (91) and (92), one can write

$$\frac{1}{r} = 1 - \alpha + \frac{\alpha}{\gamma}\left(1-\frac{1}{N}\right), \quad \text{for some } \alpha \in [0,1],$$

that enables to write

$$\left(\sum_{K\in\mathcal{M}} |K||\widehat{u}_K|^r\right)^{1/r} \le \left(\sum_{K\in\mathcal{M}} |K||\widehat{u}_K|\right)^{1-\alpha} \left(\sum_{K\in\mathcal{M}} |K||\widehat{u}_K|^{\gamma N/(N-1)}\right)^{\alpha(1-1/N)/\gamma}.$$

Inserting this in (89) gives

$$\left(\sum_{K\in\mathcal{M}} |K||\widehat{u}_K|^{\gamma N/(N-1)}\right)^{1-1/N-r(1-1/q)\alpha(1-1/N)/\gamma} \le C_N \max(1,\gamma) \left(\sum_{K\in\mathcal{M}} |K||\widehat{u}_K|\right)^{r(1-\alpha)(1-1/q)} \times \left(\left(\frac{\rho}{h_{\mathcal{M}}}\right)^{q-1} \sum_{K\in\mathcal{M}} \sum_{L\in\mathcal{N}_K} |\sigma_{KL}|\, \mathrm{d}\Big(\widehat{u}_K, \widehat{u}_L, \tfrac{m+q-1}{q}\Big)^q\right)^{1/q},$$

thus

$$\left(\sum_{K\in\mathcal{M}}|K||\widehat{u}_K|^{\gamma N/(N-1)}\right)^{(1-1/N)/\gamma}$$
$$\leq (C_N\max(1,\gamma))^{1/(\gamma-r(1-1/q)\alpha)}\left(\sum_{K\in\mathcal{M}}|K||\widehat{u}_K|\right)^{r(1-\alpha)(1-1/q)/(\gamma-r(1-1/q)\alpha)}$$
$$\times\left(\left(\frac{\rho}{h_{\mathcal{M}}}\right)^{q-1}\sum_{K\in\mathcal{M}}\sum_{L\in\mathcal{N}_K}|\sigma_{KL}|\,\mathrm{d}\Big(\widehat{u}_K,\widehat{u}_L,\tfrac{m+q-1}{q}\Big)^q\right)^{1/q(\gamma-r(1-1/q)\alpha)}.$$

Since $\alpha=(1-1/r)/(1-(1-1/N)/\gamma)$ one has

$$(1-1/N)/\gamma=\left[\frac{r(1-\alpha)(1-1/q)}{\gamma-r(1-1/q)\alpha}\right]+(1-q/N)\left[\frac{1}{q(\gamma-r(1-1/q)\alpha)}\right].$$

This means that (82) holds for $p=\gamma N/(N-1)$. Since there is no upper limitation in γ from (92), we can take arbitrarily large γ, and we conclude that (82) holds for all $1\leq p<\infty$.

□

Lemma 7. *Let $\gamma>0$ and $a,b\geq 0$. Then we have*

$$|a^\gamma-b^\gamma|\leq\max(1,\gamma)\max(a,b)^{\gamma-1}|a-b|,$$

where the product on the right-hand side is understood as 0 when $a=b=0$, even if $\gamma<1$.

Proof. We can assume that $a>0$ and $b>0$. In the case $\gamma\geq 1$, obviously

$$|a^\gamma-b^\gamma|\leq\sup_{x\in[a,b]}(\gamma x^{\gamma-1})|a-b|=\gamma\max(a,b)^{\gamma-1}|a-b|.$$

Assume now that $0<\gamma<1$. We can assume that $a\geq b$, and we set $x=a/b\geq 1$. Then we have to prove that

$$x^\gamma-1\leq x^{\gamma-1}(x-1),\quad\text{for all } x\geq 1, \tag{93}$$

or in other words $x^{\gamma-1}\leq 1$ for all $x\geq 1$, which is obvious since $\gamma<1$.

□

Lemma 8 (Renormalized distance)**.** *Let $0\leq\theta\leq 1$ be given, and define $\mathrm{d}(x,y,\theta)$ by*

$$\begin{aligned}&\mathrm{d}(x,y,\theta)=\frac{|y-x|}{\max(|x|,|y|)^\theta},\ \forall(x,y)\in\mathbb{R}^2\setminus\{(0,0)\},\\&\mathrm{d}(0,0,\theta)=0.\end{aligned} \tag{94}$$

Then the triangle inequality holds:

$$\mathrm{d}(x,y,\theta)\leq\mathrm{d}(x,z,\theta)+\mathrm{d}(z,y,\theta),\ \forall x,y,z\in\mathbb{R}. \tag{95}$$

In particular, $d(.,.,\theta)$ is a distance on $\mathbb{R}$ for all $\theta\in[0,1|$.

Proof. Let $x,y,z\in\mathbb{R}$ be given. For symmetry reasons, we can assume that $|x|\leq|y|$. Since the inequality (95) holds for $x=y=0$, let us assume that $0<|y|$.

In the case $|z|\leq|y|$, we have

$$\mathrm{d}(x,y,\theta)=\frac{|y-x|}{|y|^\theta}\leq\frac{|y-z|+|x-z|}{|y|^\theta}\leq\mathrm{d}(x,z,\theta)+\mathrm{d}(z,y,\theta),$$

since $\max(|x|,|z|)\leq|y|$.

We now assume $|z| \geq |y|$, and consider the function $g(x)$, defined for all $x \in [-|y|, |y|]$ by

$$g(x) = \mathrm{d}(x, z, \theta) + \mathrm{d}(z, y, \theta) - \mathrm{d}(x, y, \theta) = \frac{|z - x| + |y - z|}{|z|^\theta} - \frac{|y - x|}{|y|^\theta}.$$

Then g is linear over the interval $[-|y|, |y|]$, since nor y nor z belong to its interior. Thus the minimum of g on $[-|y|, |y|]$ is reached on the boundary, i.e. $x = y$ or $x = -y$. We have

$$g(y) = 2\frac{|z - y|}{|z|^\theta} \geq 0,$$

and since $|z| \geq |y|$,

$$g(-y) = \frac{|z + y| + |y - z|}{|z|^\theta} - 2\frac{|y|}{|y|^\theta} = 2(|z|^{1-\theta} - |y|^{1-\theta}) \geq 0.$$

Hence $g(x) \geq 0$ for all $x \in [-|y|, |y|]$. □

In the proof of the two next lemmas, we use Brouwer's topological degree in $\mathbb{R}^N$. In order to make these proofs clear, we first state a few basic properties of the topological degree, and we refer to [12] for the general theory.

The topological degree $d(\phi, \Omega, b)$ is defined for

- Ω non-empty bounded open set of $\mathbb{R}^N$,
- ϕ continuous from $\overline{\Omega}$ to $\mathbb{R}^N$,
- $b \in \mathbb{R}^N \backslash \phi(\partial\Omega)$.

It satisfies the following properties.
(0) $d(\phi, \Omega, b) \in \mathbb{Z}$,
(i) d is continuous with respect to ϕ for the topology of uniform convergence over $\overline{\Omega}$ (and with (0) it implies that the degree is invariant by homotopy),
(ii) for ϕ as above and additionally of class C^1 in Ω, denoting by $C(b)$ the connected component of b in $\mathbb{R}^N \backslash \phi(\partial\Omega)$, we have

$$\int_\Omega f(\phi(x)) \det \phi'(x)\, dx = d(\phi, \Omega, b) \int_{C(b)} f(y)\, dy, \quad \text{for all } f \in C_c(C(b), \mathbb{R}). \tag{96}$$

(iii) For ϕ as above and additionally of class C^1 in Ω, and if b is a regular value of ϕ, (i.e. $\forall x \in \phi^{-1}(b),\ \det \phi'(x) \neq 0$), then

$$d(\phi, \Omega, b) = \sum_{x \in \phi^{-1}(b)} \operatorname{sgn} \det \phi'(x). \tag{97}$$

We remark that if ϕ is not C^1 in property (ii), but only satisfies $\phi' \in L^N_{loc}(\mathbb{R}^N)$, then (96) still holds by simple approximation argument (use (i)). Property (ii) can be found in [12], Remark 1.14 p.16.

Lemma 9 (Nonnegativity of the Jacobian determinant for flows)**.** *Let $D_T = \{(s,t) \in \mathbb{R}^2, 0 \leq t \leq s \leq T\}$, and let X be a continuous map from $D_T \times \mathbb{R}^N$ to $\mathbb{R}^N$ (for an integer $N \geq 1$), such that*

$$X(s_3, s_2, X(s_2, s_1, x)) = X(s_3, s_1, x), \quad \textit{for all } 0 \leq s_1 \leq s_2 \leq s_3 \leq T,\ x \in \mathbb{R}^N, \tag{98}$$

$$X(s, s, x) = x, \quad \textit{for all } s \in [0, T],\ x \in \mathbb{R}^N. \tag{99}$$

Let $x_0 \in \mathbb{R}^N$ and $(s,t) \in D_T$ be given such that $X(s,t,\cdot)$ is differentiable at x_0. Let $DX(s,t,x_0)$ denote the differential of $X(s,t,\cdot)$ at the point x_0, and let $J(s,t,x_0)$ denote the determinant of $DX(s,t,x_0)$. Then $J(s,t,x_0)$ is positive or equal to zero.

Proof. If $DX(s,t,x_0)$ is non-invertible, then $J(s,t,x_0)=0$ and the conclusion holds. Let us now assume that $DX(s,t,x_0)$ is invertible, i.e. that $J(s,t,x_0)\neq 0$. Then there exists $M>0$ such that

$$|DX(s,t,x_0)\,z| \geq M\,|z| \quad \text{for all } z\in\mathbb{R}^N.$$

Since $X(s,t,\cdot)$ is differentiable at x_0, there exists $r_0>0$ such that

$$|X(s,t,y)-X(s,t,x_0)-DX(s,t,x_0)(y-x_0)| \leq \frac{M}{2}|y-x_0| \quad \text{for all } y\in\overline{B}(x_0,r_0).$$

Let us consider, for any $\theta\in[0,1]$, the map $\psi_\theta : \mathbb{R}^N\to\mathbb{R}^N$ defined by

$$\psi_\theta(y)=\theta X(s,t,y)+(1-\theta)\big(X(s,t,x_0)+DX(s,t,x_0)(y-x_0)\big), \quad \text{for } y\in\mathbb{R}^N.$$

We have

$$\psi_\theta(y)-\psi_\theta(x_0)=DX(s,t,x_0)(y-x_0)+\theta\big(X(s,t,y)-X(s,t,x_0)-DX(s,t,x_0)(y-x_0)\big),$$

which leads to

$$|\psi_\theta(y)-\psi_\theta(x_0)| \geq M\,|y-x_0|-\theta\frac{M}{2}|y-x_0| \geq \frac{M}{2}|y-x_0|, \quad \text{for all } y\in\overline{B}(x_0,r_0).$$

Therefore, for all $y\in\partial B(x_0,r_0)$, we have

$$|\psi_\theta(y)-\psi_\theta(x_0)| \geq \frac{M}{2}|y-x_0|=\frac{M}{2}r_0>0.$$

This proves that

$$\psi_\theta(x_0)=X(s,t,x_0)\notin\psi_\theta(\partial B(x_0,r_0)),\ \forall\theta\in[0,1]. \tag{100}$$

Hence the value of Brouwer's topological degree $\deg(\psi_\theta,B(x_0,r_0),X(s,t,x_0))$ (see the comments above on the topological degree) is defined and constant for $\theta\in[0,1]$. Since ψ_0 is an invertible affine function, we get that

$$\deg(\psi_\theta,B(x_0,r_0),X(s,t,x_0))=\deg(\psi_0,B(x_0,r_0),X(s,t,x_0))=\operatorname{sgn}(J(s,t,x_0)),$$

since $X(s,t,x_0)\in\psi_0(B(x_0,r_0))$. Setting $\theta=1$ in the previous equation, we obtain

$$\deg(X(s,t,\cdot),B(x_0,r_0),X(s,t,x_0))=\operatorname{sgn}(J(s,t,x_0)). \tag{101}$$

Now, for $\tau\in[t,s]$, one has according to (98),

$$X(s,\tau,X(\tau,t,y))=X(s,t,y). \tag{102}$$

Using (100) with $\theta=1$, we have that the map $X(s,t,.)$ takes values over $\partial B(x_0,r_0)$ that differ from its value at x_0. Thus, because of (102), the map $X(\tau,t,.)$ also satisfies this property. This proves that, for all $\tau\in[t,s]$, $X(\tau,t,x_0)\notin X(\tau,t,\cdot)(\partial B(x_0,r_0))$.

Therefore, for all $\tau\in[t,s]$, $\deg(X(\tau,t,\cdot),B(x_0,r_0),X(\tau,t,x_0))$ is well defined, and constant according to the continuity properties of X. Identifying the values for $\tau=s$ and $\tau=t$ yields

$$\deg\big(X(s,t,\cdot),B(x_0,r_0),X(s,t,x_0)\big)=\deg\big(X(t,t,\cdot),B(x_0,r_0),X(t,t,x_0)\big)=1,$$

because of (99) that states that $X(t,t,\cdot)$ is the identity operator, and since $x_0\in B(x_0,r_0)$. This proves, according to (101), that $\operatorname{sgn}(J(s,t,x_0))=1$, and concludes the proof. □

Lemma 10 (Change of variable formula)**.** *Let $X : \mathbb{R}^N \to \mathbb{R}^N$ (for an integer $N \geq 1$) be a Lipschitz continuous map. We assume that there exists $A \geq 0$ such that*

$$|X(x) - x| \leq A, \quad \text{for all } x \in \mathbb{R}^N. \tag{103}$$

Let $J(x)$ denote, for a.e. $x \in \mathbb{R}^N$, the determinant of the Jacobian matrix of X (recall that $J \in L^\infty(\mathbb{R}^N)$). Then the following change of variable formula holds:

$$\int_{\mathbb{R}^N} f(X(x))J(x)\,\mathrm{d}x = \int_{\mathbb{R}^N} f(y)\,\mathrm{d}y, \tag{104}$$

for all test functions $f : \mathbb{R}^N \to \mathbb{R}$ measurable, bounded, with compact support.

Proof. Note that J does not necessarily have a fixed sign in (104). For $0 \leq \theta \leq 1$, let us define $\psi_\theta(x) = (1-\theta)x + \theta X(x)$ for all $x \in \mathbb{R}^N$. Since by (103), $|\psi_\theta(x) - x| = \theta|X(x) - x| \leq A$, we deduce that

$$|\psi_\theta(x)| \geq |x| - A, \ \forall x \in \mathbb{R}^N. \tag{105}$$

For any radius $R > A$, let $B(0,R)$ be the open ball with center 0 and radius R. Let $b \in \mathbb{R}^N$ be given such that

$$|b| < R - A. \tag{106}$$

From (105) and (106), we get that $b \notin \psi_\theta(\partial B(0,R))$ for all $0 \leq \theta \leq 1$. Therefore the value of Brouwer's topological degree $\deg(\psi_\theta, B(0,R), b)$ (see the comments above about the topological degree) is independent of $0 \leq \theta \leq 1$. Since ψ_0 is the identity mapping and $b \in B(0,R)$, we have $\deg(\psi_0, B(0,R), b) = 1$, which leads for $\theta = 1$ that $\deg(X, B(0,R), b) = 1$. We then apply the relation (96) to X (X is not C^1 but see the comment after (iii)), which yields

$$\int_{B(0,R)} f(X(x))J(x)\,\mathrm{d}x = \int_{C(b)} f(y)\,\mathrm{d}y, \quad \text{for all } f \in C_c(C(b)), \tag{107}$$

where $C(b)$ denotes the greatest connected open subset in $\mathbb{R}^N \backslash X(\partial B(0,R))$ containing b. It now suffices to remark that $C(b) \supset B(0, R-A)$, which implies that (107) holds for all $f \in C_c(B(0,R-A))$, i.e.

$$\int_{\mathbb{R}^N} f(X(x))J(x)\,\mathrm{d}x = \int_{\mathbb{R}^N} f(y)\,\mathrm{d}y, \quad \text{for all } f \in C_c(B(0,R-A)), \tag{108}$$

where we replaced $B(0,R)$ by $\mathbb{R}^N$ for the set of integration in the left-hand side because according to (105), $|X(x)| \geq R - A$ for $x \notin B(0,R)$, which gives that $f(X(x)) = 0$ for $x \notin B(0,R)$. Now, since (108) holds for all $R > A$, letting $R \to \infty$, we get that (104) holds for all $f \in C_c(\mathbb{R}^N)$.

For extending the identity (104) to measurable test functions, we consider the local Radon measure defined by

$$\lambda(E) = \int_{X^{-1}(E)} J(x)\,\mathrm{d}x$$

for any bounded Borel set $E \subset \mathbb{R}^N$. Then one has

$$\int_{\mathbb{R}^N} f(y)\,d\lambda(y) = \int_{\mathbb{R}^N} f(X(x))J(x)\,dx, \quad \text{for all bounded measurable } f \text{ with compact support,} \tag{109}$$

which follows from uniform approximation of f by step functions with compact support. We have proved above that $\int f(y)d\lambda(y) = \int f(y)dy$ for all $f \in C_c(\mathbb{R}^N)$, and this implies that λ and the Lebesgue measure are identical. Thus, writing (109) with λ the Lebesgue measure yields (104) in the general case. □

5 Conclusion

In this paper, we have given sufficient conditions for the convergence of some finite volume schemes built via characteristics to the measure solution of a convection equation, in the case of irregular flow. These finite volume schemes are expressed, only using minimal hypotheses on the flow. It is worth noticing that the simplest schemes that are commonly used in the case of regular vector fields with bounded divergence and of bounded solutions, are no longer convergent in our context, without the surprising hypothesis that the ratio of the space step to the time-step tends to zero.

The ideas, appeared in the past decade, that enable to prove the convergence of transport schemes with additional numerical viscosity, using constant ratio of the space step to the time-step, can be used in the context of measures, but under strong restrictions on the concentration properties of the flow. Moreover, the convergence proofs involve the use of sharp weak BV-inequalities, leading to quite complex calculations.

Nevertheless, the results of this paper enhance the toolbox for the study of finite volume schemes, in cases of increasing importance in the probabilistic framework.

References

[1] L. M. Abia, O. Angulo, and J. C. López-Marcos. Size-structured population dynamics models and their numerical solutions. *Discrete Contin. Dyn. Syst. Ser. B*, 4(4):1203–1222, 2004.

[2] A. S. Ackleh and K. Ito. An implicit finite difference scheme for the nonlinear size-structured population model. *Numer. Funct. Anal. Optim.*, 18(9-10):865–884, 1997.

[3] L. Ambrosio. Transport equation and Cauchy problem for BV vector fields. *Invent. Math.*, 158(2):227–260, 2004.

[4] L. Ambrosio, G. Crippa, C. De Lellis, F. Otto, and M. Westdickenberg. *Transport equations and multi-D hyperbolic conservation laws*, volume 5 of *Lecture Notes of the Unione Matematica Italiana*. Springer-Verlag, Berlin, 2008. Lecture notes from the Winter School held in Bologna, January 2005, Edited by Fabio Ancona, Stefano Bianchini, Rinaldo M. Colombo, De Lellis, Andrea Marson and Annamaria Montanari.

[5] L. Ambrosio, A. Figalli, G. Friesecke, T. Paul, and J. Giannoulis. Semiclassical limit of quantum dynamics with rough potentials and well posedness of transport equations with measure initial data. 2010.

[6] S. Bianchini and M. Gloyer. An estimate on the flow generated by monotone operators. *preprint*, 2010.

[7] F. Bouchut and F. James. One-dimensional transport equations with discontinuous coefficients. *Nonlinear Anal.*, 32(7):891–933, 1998.

[8] F. Bouchut, F. James, and S. Mancini. Uniqueness and weak stability for multi-dimensional transport equations with one-sided Lipschitz coefficient. *Ann. Sc. Norm. Super. Pisa Cl. Sci. (5)*, 4(1):1–25, 2005.

[9] C. Cocozza-Thivent, R. Eymard, and S. Mercier. A finite-volume scheme for dynamic reliability models. *IMA J. Numer. Anal.*, 26(3):446–471, 2006.

[10] M. Davis. Piecewise-deterministic markov processes : a general class of non-diffusion stochastic models. *Journal of the Royal Statistical Society, Series B*, 46:353–388, 1984.

[11] R. Eymard, S. Mercier, and A. Prignet. An implicit finite volume scheme for a scalar hyperbolic problem with measure data related to piecewise deterministic markov processes. *Journal of Computational and Applied Mathematics*, 222(2):293–323, 2008.

[12] I. Fonseca and W. Gangbo. *Degree theory in analysis and applications*, volume 2 of *Oxford Lecture Series in Mathematics and its Applications*. The Clarendon Press Oxford University Press, New York, 1995. Oxford Science Publications.

[13] L. Gosse and F. James. Numerical approximations of one-dimensional linear conservation equations with discontinuous coefficients. *Math. Comp.*, 69(231):987–1015, 2000.

[14] S. Maniglia. Probabilistic representation and uniqueness results for measure-valued solutions of transport equations. *J. Math. Pures Appl. (9)*, 87(6):601–626, 2007.