



# Highly accurate monotonicity-preserving Semi-Lagrangian scheme for Vlasov-Poisson Simulations

C Yang, Michel Mehrenberger

## ► To cite this version:

C Yang, Michel Mehrenberger. Highly accurate monotonicity-preserving Semi-Lagrangian scheme for Vlasov-Poisson Simulations. *Journal of Computational Physics*, inPress, 10.1016/j.jcp.2021.110632 . hal-03126595v2

**HAL Id: hal-03126595**

**<https://hal.science/hal-03126595v2>**

Submitted on 17 Aug 2021

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

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

# Highly accurate monotonicity-preserving Semi-Lagrangian scheme for Vlasov-Poisson simulations.

Chang Yang<sup>a</sup>, Michel Mehrenberger<sup>b</sup>

<sup>a</sup>*School of Mathematics, Harbin Institute of Technology, No. 92 West Dazhi Street,  
Nangang District, Harbin, 150001, China*

<sup>b</sup>*Aix Marseille Univ, CNRS, I2M, Marseille, France*

---

## Abstract

In this paper, we study a highly accurate monotonicity-preserving (MP) Semi-Lagrangian scheme for Vlasov-Poisson simulations. The classical Semi-Lagrangian scheme is known to be highly accurate and free from CFL condition, but it does not satisfy local maximum principle. To remedy this drawback, using the conservative form of the Semi-Lagrangian scheme, we recast existing MP schemes for the numerical flux in a common framework, and then substitute the local extremum by some "better" guess, in order to avoid as much as possible loss of accuracy and clipping near extrema, while keeping the monotonicity on monotone portions. With the limiter, on the one hand, the scheme keeps the good properties of the unlimited scheme: it is conservative, free from CFL condition and highly accurate. On the other hand, for locally monotonic data, the monotonicity of the solution is preserved. Numerical tests are made on free transport equation and Vlasov-Poisson system illustrating that the limited scheme is more diffusive compared to cubic splines but has better  $L^1$  conservation, which is primarily an advantage for problems with sharp gradients.

*Keywords:* Monotonicity-preserving Scheme; Semi-Lagrangian scheme;  
Vlasov-Poisson Simulations

---

## 1. Introduction

The Vlasov-Poisson system describes the evolution of charged particles under a self-consistent electric field. One important application of the Vlasov-Poisson system is in the study of the controlled fusion. The Vlasov-Poisson system has a number of conservation properties that need special attention when developing numerical methods. Ideally, we want numerical method to retain the exact invariants in numerical methods. However, when it is not

possible to keep them all, they can be used to monitor the validity of the simulation by checking accuracy of these invariants. Many attempts have been made for solving Vlasov-Poisson system, including classical discretizations as finite difference methods [2], finite element method [60, 47, 1], finite volume method [32, 33, 18, 54], spectral method [42], discontinuous Galerkin methods [37, 22, 12, 11, 59, 50, 39], statistical based method as particle-in-cell method [14, 53, 24, 34, 23, 19]. There is also another important category named Semi-Lagrangian methods [46, 4, 20, 44, 43, 52, 9, 55, 5, 30, 31], which is popular thanks to its good precision and as it is free from time step limitation. In this paper, we will focus on designing a Semi-Lagrangian method for the Vlasov-Poisson system.

The Semi-Lagrangian methods use the characteristics of the scalar hyperbolic equation, along with an interpolation method, to update the unknown from one time step to the next. The classical Semi-Lagrangian method with high order interpolation can generate new extrema, violate the monotonicity and develop numerical oscillations.

Some remedies have been proposed. Conservative Semi-Lagrangian methods were introduced by using flux formulation [20], permitting to add filters to impose monotonicity or positivity, while keeping the conservativity, which turns out to be satisfied for the current applications in this paper (constant advection equation with periodic boundary conditions). Such design was essentially developed by changing the derivatives at the cell edges and a full monotonic or bounded preserving solution was obtained, in the framework of Hermite representation of the solution, that is locally a polynomial of degree  $\leq 3$ . The case of higher order reconstructions is more complex (see [27]), as one has to have conditions for a polynomial to be monotone or positive. A nearby solution is the positive and flux conservative (PFC) scheme that was earlier developed [35], and a local variant has been proposed by Umeda [51], and then a generalization has been performed for polynomials of degree  $\leq 4$  [52] instead of degree  $\leq 3$ . Another strategy is based on a weighted essentially non-oscillatory (WENO) type reconstruction [44, 55], typically with polynomials of degree  $\leq 5$ . However, WENO type reconstruction is too much dissipative for long term simulation as shown in Section 4. The key point of WENO type method is to find "optimal" non-linear combination weights. In contrast, we will position our design in the framework of monotonicity-preserving (MP) constraint, that is applied directly on classical interpolations and the modification is directly on the numerical flux [48], which removes the problem of dealing with high order polynomials and corresponding criteria to get monotonicity or positivity.

Suresh *et al.* [48] proposed such a limiter for numerical flux, which can retain high accuracy and in the meantime preserves monotonicity, so that the scheme can efficiently remove spurious oscillations. The idea is to distinguish automatically monotone portion and extrema portion of solution. On the one hand, for monotone solution, the limiter preserves monotonicity of the solution, on the other hand, it provides enough relaxation space to retain high accuracy of solution. However, this limiter has limitation of Courant-Friedrichs-Lewy (CFL) condition. It has been applied recently to the Semi-Lagrangian method

and for the Vlasov-Poisson system [49]. Note also that an extension work of [48] has been made [21] to give more relaxation space.

We refer also to the recent paper [45] for some complementary references on the numerical resolution of the Vlasov equation by Semi-Lagrangian type methods focussing precisely on removing spurious oscillations, which is the concern of our paper.

In this paper, we focus on developing a new limiter for Semi-Lagrangian method. We first reformulate the existing MP limiters in a common framework, that is the MP constraint mentioned in [28], then we substitute local extremum by some "better" guess extremum. In this framework, we can easily compare among the existing limiters and propose our best choice. In numerical tests, we use  $L^1$  norm invariant to estimate growing of spurious oscillation and  $L^2$  norm invariant for dissipation of the solution.

The outline of this paper is following. The highly accurate Semi-Lagrangian scheme is revised in Section 2. The construction for left flux and right flux is presented respectively. Some symmetry argument between left flux and right flux is also explained here. In Section 3, the new limiter is introduced. We first define the MP constraints for both left flux and right flux. Then the proposed relaxations to the MP constraints are explained in detail and a MP property is proven. Finally, a short analysis of comparison is given for different limiters. In Section 4, we give a portion of C code for the limiters. Then numerical results for linear advection equation and Vlasov-Poisson system are collected. At the end, we give a conclusion and perspectives.

## 2. Highly accurate Semi-Lagrangian scheme

In this section, we will introduce the highly accurate Semi-Lagrangian scheme for the free transport equation

$$\partial_t f + v \partial_x f = 0. \quad (1)$$

Let us first introduce uniform mesh in space as  $x_i = ih$ ,  $i \in \mathbb{Z}$ , where  $h$  is a fixed mesh size. Similarly, we give discrete time as  $t^n = n\Delta t$ ,  $n \in \mathbb{Z}^+$ , where  $\Delta t$  is time step. Hence the numerical solution at discrete grid is denoted by  $f_i^n$ .

The classical Semi-Lagrangian scheme is divided in two steps. The first step is devoted to finding foot of characteristic curve. For the free transport equation (1), the foot is simply  $x_i - v\Delta t$ , if we start from  $x_i$ . The second step consists to approximate the solution by an interpolation method thanks to the following relation

$$f(t^{n+1}, x_i) = f(t^n, x_i - v\Delta t).$$

Firstly, we notice that the foot of characteristic curve will locate between two successive nodes of the mesh. Once the foot is determined, we use Lagrange interpolation of odd degree  $2d + 1$ , with  $d \in \mathbb{N}$  to approximate  $f(t^n, x_i - v\Delta t)$ . The detailed construction steps are explained in [Appendix A](#).

Secondly, the Semi-Lagrangian scheme can be recast into a flux form, as we have in classical finite volume scheme for transport equation. It is a preparation

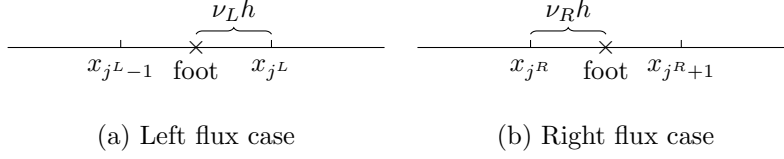


Figure 1: Foot of characteristic curve for Semi-Lagrangian method

$c_0^L(\nu)$   1		$c_{-1}^L(\nu)$	$\frac{1}{6}\nu^2 - \frac{1}{6}$
		$c_0^L(\nu)$	$-\frac{1}{3}\nu^2 + \frac{1}{2}\nu + \frac{5}{6}$
		$c_1^L(\nu)$	$\frac{1}{6}\nu^2 - \frac{1}{2}\nu + \frac{1}{3}$
$c_{-2}^L(\nu)$	$\frac{1}{120}\nu^4 - \frac{1}{24}\nu^2 + \frac{1}{30}$		
$c_{-1}^L(\nu)$	$-\frac{1}{30}\nu^4 + \frac{1}{24}\nu^3 + \frac{1}{4}\nu^2 - \frac{1}{24}\nu - \frac{13}{60}$		
$c_0^L(\nu)$	$\frac{1}{20}\nu^4 - \frac{1}{8}\nu^3 - \frac{1}{3}\nu^2 + \frac{5}{8}\nu + \frac{47}{60}$		
$c_1^L(\nu)$	$-\frac{1}{30}\nu^4 + \frac{1}{8}\nu^3 + \frac{1}{12}\nu^2 - \frac{5}{8}\nu + \frac{9}{20}$		
$c_2^L(\nu)$	$\frac{1}{120}\nu^4 - \frac{1}{24}\nu^3 + \frac{1}{24}\nu^2 + \frac{1}{24}\nu - \frac{1}{20}$		

Table 1: Values of  $c_\ell^L(\nu) = c_{-\ell}^R(\nu)$ ,  $\ell = -d, \dots, d$ , for  $d = 0, 1, 2$ , using (A.3).

phase for monotonicity-preserving limiter. There are two types of flux formulae, called left flux or right flux formula.

**Left flux formula (symbolized by  $L$ ).** Suppose that the foot  $x_i - v\Delta t$  locates between  $x_{j^L-1}$  and  $x_{j^L}$ , as shown in Fig 1(a). Then we define local CFL number  $\nu_L := (x_{j^L} - (x_i - v\Delta t))/h$ , it is clear that  $\nu_L \in (0, 1)$ . Now, the Semi-Lagrangian scheme in left flux formula can be expressed as

$$f_i^{n+1} = f_{j^L}^n - \nu_L \left( f_{j^L+1/2}^L - f_{j^L-1/2}^L \right), \quad (2)$$

defining linear formula

$$f_{j^L+1/2}^L = \sum_{\ell=-d}^d c_\ell^L(\nu_L) f_{j^L+\ell}^n, \quad (3)$$

where  $c_\ell^L(\nu)$ ,  $\ell \in \{-d, \dots, d\}$ , functions of variable  $\nu$ , are functional coefficients given in (A.3). We report on Table 1, the expressions of  $c_\ell^L(\nu)$  for some values of  $d$ .

**Right flux formula (symbolized by  $R$ ).** Suppose that the foot  $x_i - v\Delta t$  locates between  $x_{j^R}$  and  $x_{j^R+1}$ , as shown in Fig 1(b). Then we define local CFL number  $\nu_R := ((x_i - v\Delta t) - x_{j^R})/h$ , also  $\nu_R \in (0, 1)$ , and  $\nu_R = 1 - \nu_L$ . Now, the Semi-Lagrangian scheme in right flux formula can be expressed as

$$f_i^{n+1} = f_{j^R}^n + \nu_R \left( f_{j^R+1/2}^R - f_{j^R-1/2}^R \right), \quad (4)$$

defining linear formula

$$f_{j^R-1/2}^R = \sum_{\ell=-d}^d c_\ell^R(\nu_R) f_{j^R+\ell}^n, \quad (5)$$

where  $c_\ell^R(\nu)$ ,  $\ell \in \{-d, \dots, d\}$ , functions of variable  $\nu$ , are functional coefficients given in (A.6). Moreover, we can easily deduce that  $c_\ell^R(\nu) = c_{-\ell}^L(\nu)$ , for  $\ell = -d, \dots, d$  (see Appendix A.2).

**Remark 2.1.** *By following a classical finite volume scheme (not Semi-Lagrangian), we may think of using the left flux formula for  $v > 0$  and the right flux formula for  $v < 0$ , but it is not necessary. Indeed, the Lagrange interpolation depends on the location of foot of characteristic curve, not on the sign of  $v$ . The two flux formulae (2) and (4) are equivalent to the Lagrange interpolation approximation. Therefore, we can choose freely from the left flux formula (2) or the right flux formula (4) according to convenience.*

### 3. MP constraints and relaxation strategies

The proposed scheme in the previous section is simple and very accurate for regular solutions. However, it may provoke spurious oscillations for irregular solutions. In this section, we focus on proposing a flux limiter, such that on the one hand, the scheme has a **monotonicity-preserving** (MP) property, and on the other hand, the scheme can preserve high accuracy.

The strategy for deriving flux limiter consists of two steps: the first is to identify the MP constraints, the second is to relax the MP constraints near extrema.

#### 3.1. MP constraints

The MP constraints are related to the left flux formula or the right flux formula.

##### MP constraints related to the left flux formula (symbolized by $L$ ).

The MP constraints consist of two parts, as mentioned in [28]. The first one is that the flux  $f_{j^L+1/2}^n$  should be located between  $f_{j^L}^n$  and  $f_{j^L+1}^n$ , *i.e.*

$$m_{j^L+1/2} \leq f_{j^L+1/2}^L \leq M_{j^L+1/2}, \quad (6)$$

where  $m_{j^L+1/2} = \min(f_{j^L}^n, f_{j^L+1}^n)$  and  $M_{j^L+1/2} = \max(f_{j^L}^n, f_{j^L+1}^n)$ , for  $j^L \in \mathbb{Z}$ . The second one is the total variation diminishing (TVD) condition, *i.e.*

$$m_{j^L-1/2} \leq f_i^{n+1} \leq M_{j^L-1/2}.$$

Using the left flux formula (2), the above TVD condition is equivalent to

$$f_{j^L-1/2}^n + \frac{1}{\nu_L} (f_{j^L}^n - M_{j^L-1/2}) \leq f_{j^L+1/2}^L \leq f_{j^L-1/2}^n + \frac{1}{\nu_L} (f_{j^L}^n - m_{j^L-1/2}).$$

Consider that  $m_{j^L-1/2} \leq f_{j^L-1/2}^L \leq M_{j^L-1/2}$ , then a sufficient condition of TVD scheme is

$$\Phi_{\nu_L, j^L}(M_{j^L-1/2}) \leq f_{j^L+1/2}^L \leq \Phi_{\nu_L, j^L}(m_{j^L-1/2}), \quad (7)$$

where

$$\Phi_{\nu, j}(x) = x + \frac{1}{\nu}(f_j - x).$$

Thanks to the MP constraints (6) and (7), which can be rewritten as

$$\max(m_{j^L+1/2}, \Phi_{\nu_L, j^L}(M_{j^L-1/2})) \leq f_{j^L+1/2}^L \leq \min(M_{j^L+1/2}, \Phi_{\nu_L, j^L}(m_{j^L-1/2})), \quad (8)$$

we can easily find that if the data  $\{f_i^n\}$  is monotone, then  $\{f_i^{n+1}\}$  is also monotone for all  $i \in \mathbb{Z}$ , because  $f_i^{n+1}$  lies between  $f_{j^L-1}^n$  and  $f_{j^L}^n$ .

In the sequel, we will use same notations as in [48]:

$$\text{minmod}(x, y) = \frac{1}{2}(\text{sgn}(x) + \text{sgn}(y)) \min(|x|, |y|),$$

$$\text{median}(x, y, z) = x + \text{minmod}(y - x, z - x),$$

$$I[x_1, \dots, x_k] = [\min(x_1, \dots, x_k), \max(x_1, \dots, x_k)],$$

where  $\text{sgn}$  is the sign function.

As mentioned in the literature [56, 58], these MP constraints (6)-(7) will lead to first order accuracy for smooth solution, the so called clipping near extrema. For instance, in [48], two cases near extrema are identified:

Case 1: When  $f_{j^L}^n = f_{j^L+1}^n$ , the constraint (6) leads that the numerical flux is limited as  $f_{j^L+1/2}^L = f_{j^L}^n$ .

However, let us consider a second order Semi-Lagrangian scheme interpolating on  $x_{j^L-2}$ ,  $x_{j^L-1}$  and  $x_{j^L}$ , which is nothing but the Beam-Warming scheme (see e.g. [25])

$$f_i^{n+1} = f_{j^L}^n - \nu_L(f_{j^L+1/2}^L - f_{j^L-1/2}^L),$$

where the flux is defined by

$$f_{j^L+1/2}^L = f_{j^L}^n + \frac{1}{2}(1 - \nu_L)(f_{j^L}^n - f_{j^L-1}^n). \quad (9)$$

Due to the MP constraint (6), the flux will be limited to  $f_{j^L+1/2}^L = f_{j^L}^n$ , which is quite different to (9) if  $f_{j^L}^n \neq f_{j^L-1}^n$ , thus the Beam-Warming scheme degenerates to a first order scheme.

Case 2: When  $f_{j^L-1}^n = f_{j^L}^n$ , the constraint (7) leads also that  $f_{j^L+1/2}^L = f_{j^L}^n$ .

This time, let us consider a second order Semi-Lagrangian scheme interpolating on  $x_{j^L-1}$ ,  $x_{j^L}$  and  $x_{j^L+1}$ , which is nothing but the Lax-Wendroff scheme (see e.g. [25])

$$f_i^{n+1} = f_{j^L}^n - \nu_L(f_{j^L+1/2}^L - f_{j^L-1/2}^L),$$

where the flux is defined by

$$f_{j^L+1/2}^L = f_{j^L}^n + \frac{1}{2}(1 - \nu_L)(f_{j^L+1}^n - f_{j^L}^n).$$

Again the Lax-Wendroff scheme degenerates to a first order scheme.

We see in both Case 1 and Case 2, we lose accuracy near extrema.

To remedy this drawback of MP constraints, we should relax the constraints (6) and (7) near extrema. The rule of relaxation is on the one hand to provide relaxation space as much as possible near extrema, and on the other hand to preserve monotonicity for monotone solution.

We notice that the constraints (6) and (7) are defined by the local maximum  $M_{j+1/2}$  and local minimum  $m_{j+1/2}$ , for  $j \in \mathbb{Z}$ . Thus, one possible way to relax the MP constraints near extrema is to replace the local extremum by a "better" extremum guess. More precisely, we denote the "better" extremum guess by  $M_{j+1/2}^{(1,L)}/m_{j+1/2}^{(1,L)}$  and  $M_{j+1/2}^{(2,L)}/m_{j+1/2}^{(2,L)}$ , for  $j \in \mathbb{Z}$ . Then injecting them into the MP constraints (6)-(7) yields

$$m_{j^L+1/2}^{(1,L)} \leq f_{j^L+1/2}^L \leq M_{j^L+1/2}^{(1,L)}, \quad (10)$$

and

$$\Phi_{\nu_L, j^L}(M_{j^L-1/2}^{(2,L)}) \leq f_{j^L+1/2}^L \leq \Phi_{\nu_L, j^L}(m_{j^L-1/2}^{(2,L)}), \quad (11)$$

which lead to the limiting

$$f_{j^L+1/2}^L = \text{median} \left( f_{\min, j^L}, f_{j^L+1/2}^L, f_{\max, j^L} \right),$$

with

$$\begin{aligned} f_{\min, j^L} &= \max \left( m_{j^L+1/2}^{(1,L)}, \Phi_{\nu_L, j^L}(M_{j^L-1/2}^{(2,L)}) \right), \\ f_{\max, j^L} &= \min \left( M_{j^L+1/2}^{(1,L)}, \Phi_{\nu_L, j^L}(m_{j^L-1/2}^{(2,L)}) \right). \end{aligned}$$

Those "better" extremum guesses are the key point of our proposed limiter, which will be specified in Section 3.2 and Section 3.3. Moreover, looking at (8) a definition of new MP constraints is deduced as follows

**Definition 3.1.** *For monotonic data, that is  $f_{j^L-2}^n \leq f_{j^L-1}^n \leq f_{j^L}^n \leq f_{j^L+1}^n \leq f_{j^L+2}^n$  or  $f_{j^L-2}^n \geq f_{j^L-1}^n \geq f_{j^L}^n \geq f_{j^L+1}^n \geq f_{j^L+2}^n$ , if the following constraints are verified*

1.  $\max \left( m_{j^L+1/2}^{(1,L)}, \Phi_{\nu_L, j^L}(M_{j^L-1/2}^{(2,L)}) \right) \leq \max \left( m_{j^L+1/2}^{(1,L)}, \Phi_{\nu_L, j^L}(M_{j^L-1/2}^{(2,L)}) \right)$
2.  $\min \left( M_{j^L+1/2}^{(1,L)}, \Phi_{\nu_L, j^L}(m_{j^L-1/2}^{(2,L)}) \right) \geq \min \left( M_{j^L+1/2}^{(1,L)}, \Phi_{\nu_L, j^L}(m_{j^L-1/2}^{(2,L)}) \right)$

*then, the constraints defined in (10)-(11) are MP.*



We first consider a case where we take the same value for the extremum guess, that is  $m_{j+1/2}^{(1,L)} = m_{j+1/2}^{(2,L)}$  and  $M_{j+1/2}^{(1,L)} = M_{j+1/2}^{(2,L)}$ . Then the previous definition simplifies and we can check the following proposition.

**Proposition 3.1.** *We suppose that  $m_{j+1/2}^L := m_{j+1/2}^{(1,L)} = m_{j+1/2}^{(2,L)}$  and  $M_{j+1/2}^L := M_{j+1/2}^{(1,L)} = M_{j+1/2}^{(2,L)}$  for all  $j$ . Then, the constraints defined in (10)-(11) are MP if we have the following **strong MP property**:*

$$m_{j+1/2} \leq m_{j+1/2}^L, \quad M_{j+1/2} \geq M_{j+1/2}^L, \quad j \in \{j_L - 1, j_L\}, \quad (12)$$

for monotonic data, that is  $f_{j^L-2}^n \leq f_{j^L-1}^n \leq f_{j^L}^n \leq f_{j^L+1}^n \leq f_{j^L+2}^n$  or  $f_{j^L-2}^n \geq f_{j^L-1}^n \geq f_{j^L}^n \geq f_{j^L+1}^n \geq f_{j^L+2}^n$ .

**MP constraints related to the right flux formula (symbolized by  $R$ ).** Similarly to the modified constraints (10)-(11), we obtain MP constraints related to the right flux formula

$$m_{j^R-1/2}^{(1,R)} \leq f_{j^R-1/2}^R \leq M_{j^R-1/2}^{(1,R)} \quad (13)$$

together with

$$\Phi_{\nu_R, j^R}(M_{j^R+1/2}^{(2,R)}) \leq f_{j^R-1/2}^R \leq \Phi_{\nu_R, j^R}(m_{j^R+1/2}^{(2,R)}), \quad (14)$$

leading to the limiting

$$f_{j^R-1/2}^R = \text{median} \left( f_{\min, j^R}, f_{j^R-1/2}^R, f_{\max, j^R} \right)$$

with

$$\begin{aligned} f_{\min, j^R} &= \max \left( m_{j^R-1/2}^{(1,R)}, \Phi_{\nu_R, j^R}(M_{j^R+1/2}^{(2,R)}) \right), \\ f_{\max, j^R} &= \min \left( M_{j^R-1/2}^{(1,R)}, \Phi_{\nu_R, j^R}(m_{j^R+1/2}^{(2,R)}) \right). \end{aligned}$$

Here,  $M_{j+1/2}^{(1,R)}/m_{j+1/2}^{(1,R)}$  or  $M_{j+1/2}^{(2,R)}/m_{j+1/2}^{(2,R)}$  for  $j \in \mathbb{Z}$  are "better" extremum guesses for right flux formula.

**Remark 3.1.** *When we have  $m_{j^R+1/2}^{(1,R)} = m_{j^L-1/2}^{(2,L)}$ ,  $M_{j^R+1/2}^{(1,R)} = M_{j^L-1/2}^{(2,L)}$  and  $m_{j^R+1/2}^{(2,R)} = m_{j^L-1/2}^{(1,L)}$ ,  $M_{j^R+1/2}^{(2,R)} = M_{j^L-1/2}^{(1,L)}$ , then the limiter for right flux formula is equivalent to the one for left flux formula. The detailed explanation is given in [Appendix B.1](#).*

**3.2. Relaxation of the MP constraints (6)-(7) to satisfy the strong MP property (12)**

In [48], the authors look for relaxing the constraint (6) and then in a second step, the constraint (7). In our framework, we directly want to determine

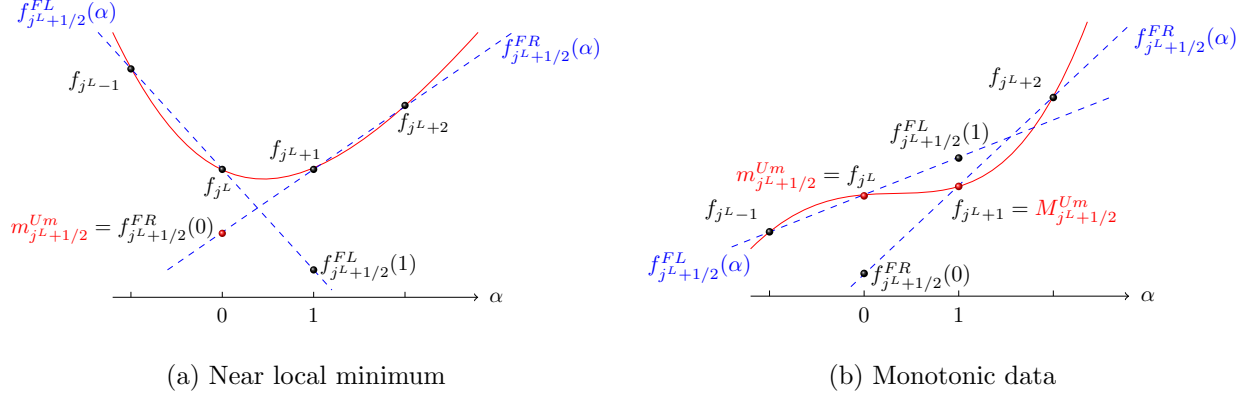


Figure 2: Searching extremum guess by linear extrapolation.

"better" extremum guess  $M_{j+1/2}^L/m_{j+1/2}^L$ , for  $j \in \mathbb{Z}$ , which then will relax both constraints (6) and (7). The idea is on the one hand to search enough relaxation space near extrema, on the other hand to simply preserve the constraints (6)-(7) for monotone solution. A first way to find "better" extremum guess is inspired from Umeda's method [51] (**Um** for short), that is to search extremum by extrapolating from  $f_{j^L-1}^n$  and  $f_{j^L}^n$  (or  $f_{j^L}^n$  and  $f_{j^L+1}^n$ ) to interval  $]x_{j^L}, x_{j^L+1}[$  (see Fig 2).

Let us define linear extrapolation from left (FL) or from right (FR) as a function of variable  $\alpha$ :

$$\begin{aligned} f_{j^L+1/2}^{FL}(\alpha) &= f_{j^L}^n + \alpha (f_{j^L}^n - f_{j^L-1}^n), \\ f_{j^L+1/2}^{FR}(\alpha) &= f_{j^L+1}^n + (1 - \alpha) (f_{j^L+1}^n - f_{j^L+2}^n). \end{aligned}$$

Thanks to these definitions, we propose the minimum guess, as shown in Fig 2(a), is  $\max(f_{j^L+1/2}^{FL}(1), f_{j^L+1/2}^{FR}(0))$ . Clearly, this minimum guess provides enough relaxation space.

On the other hand, for monotonic data (see Fig 2(b)), we shall restrict flux in  $I[f_{j^L}^n, f_{j^L+1}^n]$

$= I[f_{j^L+1/2}^{FL}(0), f_{j^L+1/2}^{FR}(1)]$ . So the minimum guess is  $\min(f_{j^L+1/2}^{FL}(0), f_{j^L+1/2}^{FR}(1))$ .

Combining the above two cases, the proposed minimum guess is

$$m_{j^L+1/2}^{Um} = \min(\min(f_{j^L+1/2}^{FL}(0), f_{j^L+1/2}^{FR}(1)), \max(f_{j^L+1/2}^{FL}(1), f_{j^L+1/2}^{FR}(0))). \quad (15a)$$

Similarly, the maximum guess is

$$M_{j^L+1/2}^{Um} = \max(\max(f_{j^L+1/2}^{FL}(0), f_{j^L+1/2}^{FR}(1)), \min(f_{j^L+1/2}^{FL}(1), f_{j^L+1/2}^{FR}(0))). \quad (15b)$$

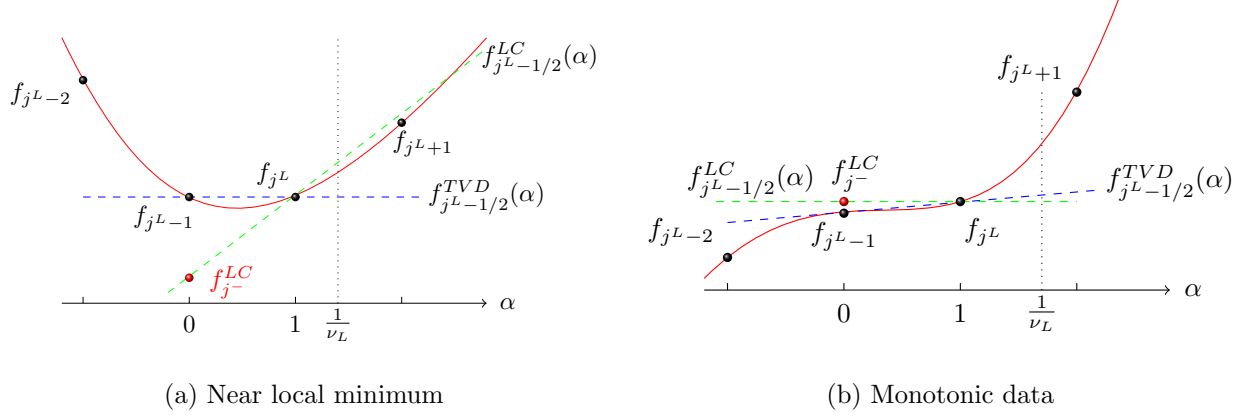


Figure 3: Searching extremum guess by using large curvature.

Let us now show property of MP constraints with the above extremum guess.

**Proposition 3.2.** *For monotonic data, the constraints defined in (10)-(11) is MP when substituting  $m_{j^L+1/2}^{(1,L)}/M_{j^L+1/2}^{(1,L)}$  by  $m_{j^L+1/2}^{Um}/M_{j^L+1/2}^{Um}$  and  $m_{j^L-1/2}^{(2,L)}/M_{j^L-1/2}^{(2,L)}$  by  $m_{j^L-1/2}^{Um}/M_{j^L-1/2}^{Um}$ .*

*Proof.* For increasing data, we have

$$\begin{aligned} m_{j^L+1/2}^{Um} &= \min \left( \min(f_{j^L}^n, f_{j^L+1}^n), \max(2f_{j^L}^n - f_{j^L-1}^n, 2f_{j^L+1}^n - f_{j^L+2}^n) \right) \\ &= \min \left( f_{j^L}^n, \max(f_{j^L}^n + (f_{j^L}^n - f_{j^L-1}^n), 2f_{j^L+1}^n - f_{j^L+2}^n) \right). \end{aligned}$$

It is clear that  $f_{j^L}^n + (f_{j^L}^n - f_{j^L-1}^n) \geq f_{j^L}^n$ , so that  $\max(f_{j^L}^n + (f_{j^L}^n - f_{j^L-1}^n), 2f_{j^L+1}^n - f_{j^L+2}^n) \geq f_{j^L}^n$ . Thus  $m_{j^L+1/2}^{Um} = f_{j^L}^n$ .

On the other hand, we have

$$\begin{aligned} M_{j^L+1/2}^{Um} &= \max \left( \max(f_{j^L}^n, f_{j^L+1}^n), \min(2f_{j^L}^n - f_{j^L-1}^n, 2f_{j^L+1}^n - f_{j^L+2}^n) \right) \\ &= \max \left( f_{j^L+1}^n, \min(2f_{j^L}^n - f_{j^L-1}^n, f_{j^L+1}^n + (f_{j^L+1}^n - f_{j^L+2}^n)) \right) \end{aligned}$$

It is clear that  $f_{j^L+1}^n + (f_{j^L+1}^n - f_{j^L+2}^n) \leq f_{j^L+1}^n$ , so that  $\min(2f_{j^L}^n - f_{j^L-1}^n, f_{j^L+1}^n + (f_{j^L+1}^n - f_{j^L+2}^n)) \leq f_{j^L+1}^n$ . Thus  $M_{j^L+1/2}^{Um} = f_{j^L+1}^n$ .

Similarly, we have  $m_{j^L-1/2}^{Um} = f_{j^L-1}^n$  and  $M_{j^L-1/2}^{Um} = f_{j^L}^n$ . Therefore, we have the result.

For decreasing data, the proof follows the same way.  $\square$

### 3.3. Further relaxation of the MP constraints (6)-(7)

We now allow for the scheme to take different values for  $m_{j+1/2}^{(1,L)}/M_{j+1/2}^{(1,L)}$  and  $m_{j+1/2}^{(2,L)}/M_{j+1/2}^{(2,L)}$ , in order to have even more relaxation space. For this, we follow

[48], where a technique using Large Curvature (as it will be explained later) was used for constraint (7). By a symmetry argument, we will get also a way to relax constraint (6). Let us consider the Case 2 of Section 3.1. When near local minimum, for example as shown in Fig 3(a), the constraint (7) becomes

$$f_{j^L}^n \leq f_{j^L+1/2}^n \leq f_{j^L-1}^n + \frac{1}{\nu_L}(f_{j^L}^n - f_{j^L-1}^n) = \Phi_{\nu_L, j^L}(f_{j^L-1}^n). \quad (16)$$

Since  $f_{j^L}^n = f_{j^L-1}^n$ , we deduce that  $f_{j^L+1/2}^n = f_{j^L}^n$ . One way to relax the constraint (16) is to replace  $f_{j^L-1}^n$  by a better minimum guess. Notice from Fig 3(a) that the curvature is positive. We thus propose the following minimum guess

$$f_{j^-}^{LC} = f_{j^L}^n - d_{j^L-1/2}, \quad (17)$$

where **LC** stands for **large curvature**, and  $d_{j-1/2} = \min\text{mod}(d_{j-1}, d_j)$ ,  $d_j = f_{j-1} - 2f_j + f_{j+1}$  for  $j \in \mathbb{Z}$ . Let us denote

$$f_{j^L-1/2}^{TVD}(\alpha) = f_{j^L-1}^n + \alpha(f_{j^L}^n - f_{j^L-1}^n) = \Phi_{1/\alpha, j^L}(f_{j^L-1}^n),$$

$$f_{j^L-1/2}^{LC}(\alpha) = f_{j^-}^{LC} + \alpha(f_{j^L}^n - f_{j^-}^{LC}) = \Phi_{1/\alpha, j^L}(f_{j^-}^{LC}).$$

It is clear that  $f_{j^L-1/2}^{TVD}(\frac{1}{\nu_L}) = \Phi_{\nu_L, j^L}(f_{j^L-1}^n) \leq \Phi_{\nu_L, j^L}(f_{j^-}^{LC}) = f_{j^L-1/2}^{LC}(\frac{1}{\nu_L})$  since  $\frac{1}{\nu_L} \geq 1$  and  $f_{j^-}^{LC} \leq f_{j^L}^n$  as  $d_{j^L-1/2} \geq 0$ , thus we gain more relaxation space.

On the other hand, for increasing data (see Fig 3(b)), we have  $d_{j^L-1/2} \leq f_{j^L}^n - f_{j^L-1}^n$  since  $d_{j^L-1/2} = 0$ , thus  $f_{j^-}^{LC} = f_{j^L}^n \geq f_{j^L-1}^n$ , so that  $\Phi_{\nu_L, j^L}(f_{j^-}^{LC}) = f_{j^L}^n \in [f_{j^L}^n, f_{j^L+1}^n]$ .

Combining with the limiter introduced in Section 3.2, we propose extremum guess  $M_{j^L-1/2}^{(2,L)}/m_{j^L-1/2}^{(2,L)}$  as  $\max(M_{j^L-1/2}^{Um}, f_{j^-}^{LC}) / \min(m_{j^L-1/2}^{Um}, f_{j^-}^{LC})$ . Therefore, we have

$$\max\left(M_{j^L+1/2}^{Um}, \Phi_{\nu_L, j^L}(\max(M_{j^L-1/2}^{Um}, f_{j^-}^{LC}))\right) \leq f_{j^L+1/2}^L \leq \min\left(M_{j^L+1/2}^{Um}, \Phi_{\nu_L, j^L}(\min(m_{j^L-1/2}^{Um}, f_{j^-}^{LC}))\right). \quad (18)$$

Notice that near local minimum, for instance the case in Fig 3(a),  $m_{j^L-1/2}^{Um}$  and  $f_{j^-}^{LC}$  can be simplified as

$$m_{j^L-1/2}^{Um} = \max(f_{j^L}^n - d_{j^L-1}, f_{j^L-1}^n - d_{j^L}),$$

$$f_{j^-}^{LC} = \max(f_{j^L}^n - d_{j^L-1}, f_{j^L}^n - d_{j^L}),$$

since  $m_{j^L-1/2}^{Um} = \max(2f_{j^L-1}^n - f_{j^L-2}^n, 2f_{j^L}^n - f_{j^L+1}^n) \leq \min(f_{j^L}^n, f_{j^L-1}^n)$ , using that  $d_{j^L} \geq 0$ ,  $d_{j^L-1} \geq 0$  and  $f_{j^L} \leq f_{j^L+1}$  and  $f_{j^L-1} \leq f_{j^L-2}$ . It is clear that  $f_{j^-}^{LC} \leq m_{j^L-1/2}^{Um}$  if  $f_{j^L}^n \leq f_{j^L-1}^n$ . Moreover, for constant curvature case, we always have  $f_{j^-}^{LC} \leq m_{j^L-1/2}^{Um}$ . This indicates that, compared with  $m_{j^L-1/2}^{Um}$ ,  $f_{j^-}^{LC}$  may lead more relaxation space.

Then, we consider the limiter for right flux formula, by symmetric argument, we have the constraint

$$\max \left( m_{j^R+1/2}^{Um}, \Phi_{\nu_R, j^R+1}(\max(M_{j^R+3/2}^{Um}, f_{j^+}^{LC})) \right) \leq f_{j^R+1/2}^R \leq \min \left( M_{j^R+1/2}^{Um}, \Phi_{\nu_R, j^R+1}(\min(m_{j^R+3/2}^{Um}, f_{j^+}^{LC})) \right),$$

with

$$f_{j^+}^{LC} := f_{j^R+1} - d_{j^R+3/2} = f_{j^L} - d_{j^L+1/2}. \quad (19)$$

Thanks to the equivalent relation between the left flux formula and the right flux formula shown in Remark 3.1, we have

$$\max \left( \min(m_{j^L+1/2}^{Um}, f_{j^+}^{LC}), \Phi_{\nu_L, j^L}(M_{j^L-1/2}^{Um}) \right) \leq f_{j^L+1/2}^L \leq \min \left( \max(M_{j^L+1/2}^{Um}, f_{j^+}^{LC}), \Phi_{\nu_L, j^L}(m_{j^L-1/2}^{Um}) \right). \quad (20)$$

Now, taking union of intervals defined in (18) and (20), we can propose the final limiting:

$$f_{j^L+1/2}^L = \text{median} \left( f_{\min, j^L}, f_{j^L+1/2}^L, f_{\max, j^L} \right), \quad j^L \in \mathbb{Z}, \quad (21a)$$

with

$$f_{\min, j^L} = \min \left( \max \left( m_{j^L+1/2}^{Um}, \Phi_{\nu_L, j^L}(\max(M_{j^L-1/2}^{Um}, f_{j^+}^{LC})) \right), \max \left( \min(m_{j^L+1/2}^{Um}, f_{j^+}^{LC}), \Phi_{\nu_L, j^L}(M_{j^L-1/2}^{Um}) \right) \right) \quad (21b)$$

and

$$f_{\max, j^L} = \max \left( \min \left( M_{j^L+1/2}^{Um}, \Phi_{\nu_L, j^L}(\min(m_{j^L-1/2}^{Um}, f_{j^+}^{LC})) \right), \min \left( \max(M_{j^L+1/2}^{Um}, f_{j^+}^{LC}), \Phi_{\nu_L, j^L}(m_{j^L-1/2}^{Um}) \right) \right). \quad (21c)$$

**Proposition 3.3.** *For monotonic data, the constraint defined in (21) is MP.*

Note that MP is defined for constraints (10)-(11), so, saying that "the constraint defined in (21) is MP" might not be clear. However, the constraint defined in (21) can be seen as constraints (10)-(11). In fact, by defining

- $(m_{j^L+1/2}^{(1,L)}, M_{j^L-1/2}^{(2,L)}) = (m_{j^L+1/2}^{Um}, \max(M_{j^L-1/2}^{Um}, f_{j^+}^{LC}))$ , if
$$\max \left( m_{j^L+1/2}^{Um}, \Phi_{\nu_L, j^L}(\max(M_{j^L-1/2}^{Um}, f_{j^+}^{LC})) \right) \leq \max \left( \min(m_{j^L+1/2}^{Um}, f_{j^+}^{LC}), \Phi_{\nu_L, j^L}(M_{j^L-1/2}^{Um}) \right)$$
- $(m_{j^L+1/2}^{(1,L)}, M_{j^L-1/2}^{(2,L)}) = (\min(m_{j^L+1/2}^{Um}, f_{j^+}^{LC}), M_{j^L-1/2}^{Um})$ , otherwise,

and by defining similarly for  $(M_{j^L+1/2}^{(1,L)}, m_{j^L-1/2}^{(2,L)})$ .

The proof of the proposition is given in Appendix B.2. Moreover, the MP constraint related to the right flux formula can be obtained by symmetric argument from (21).

### 3.4. Other possible choices

In this part, we explore other possible choices for extremum guess and make more precise links with the literature.

A first natural choice is the TVD limiting, that is to substitute  $m_{j^L+1/2}^{(1,L)}/M_{j^L+1/2}^{(1,L)}$  or  $m_{j^L-1/2}^{(2,L)}/M_{j^L-1/2}^{(2,L)}$  by

$$m_{j^L+1/2}^{TVD} = \min_{\alpha \in \{0,1\}} f_{j^L+1/2}^{MD}(\alpha), \quad M_{j^L+1/2}^{TVD} = \max_{\alpha \in \{0,1\}} f_{j^L+1/2}^{MD}(\alpha), \quad (22)$$

where  $f_{j^L+1/2}^{MD}(\alpha)$  (MD stands for median), coming from [48], reads

$$f_{j^L+1/2}^{MD}(\alpha) = (1-\alpha)f_{j^L}^n + \alpha f_{j^L+1}^n - \min\{0, \alpha d_{j^L} - (1-\alpha)d_{j^L+1}\}.$$

It is easy to find that TVD limiting is nothing but MP constraints (6)-(7) itself.

Second choice is inspired from the Suresh's limiting [48] or the Daru-Tenaud's limiting [21] (**DaTe** for short). Actually, the MP constraints (6)-(7) define two limiting intervals, *i.e.*

$$I[f_{j^L}^n, f_{j^L+1}^n], \quad I[f_{j^L}^n, f_{j^L}^{UL}], \quad (23)$$

where  $f_{j^L}^{UL} = f_{j^L}^n + \frac{1-\nu_L}{\nu_L}(f_{j^L}^n - f_{j^L-1}^n)$  (UL stands for upper limit). These two limiting intervals are enlarged by

$$I[f_{j^L}^n, f_{j^L+1}^n, f_{j^L}^{MD}], \quad I[f_{j^L}^n, f_{j^L}^{UL}, f_{j^L}^{LC}]. \quad (24)$$

For  $f_{j^L}^{MD}$ , in both the Suresh's limiting and the Daru-Tenaud's limiting, we have

$$f_{j^L}^{MD} = f_{j^L+1/2}^{MD}(1/2). \quad (25)$$

However, for  $f_{j^L}^{LC}$ , in the Suresh's limiting, we have

$$f_{j^L}^{LC} = f_{j^L}^n + \frac{1}{2}(f_{j^L}^n - f_{j^L-1}^n) + \frac{4}{3}d_{j^L-1/2}, \quad (26)$$

whereas in the Daru-Tenaud's limiting, we have

$$f_{j^L}^{LC} = f_{j^L}^n + \frac{1}{2} \frac{1-\nu_L}{\nu_L} (f_{j^L}^n - f_{j^L-1}^n + d_{j^L-1/2}). \quad (27)$$

In [48], the authors point out that (26) works for  $\nu_L \leq 0.2$ , while (27) is free from this CFL constraint.

We remark that our limiting (21), without taking LC strategy, can be regarded as the enlarged intervals of (23), *i.e.*

$$I[m_{j^L+1/2}^{Um}, M_{j^L+1/2}^{Um}], \quad I[\Phi_{\nu_L, j^L}(m_{j^L-1/2}^{Um}), \Phi_{\nu_L, j^L}(M_{j^L-1/2}^{Um})].$$

This means that our limiting (21) can relax (23) without using LC strategy, whereas the Suresh's limiting or the Daru-Tenaud's limiting cannot : there, a

different strategy is adopted for (6) and (7), which is maybe linked to the finite volume formulation which could favor the case of small CFL. Furthermore, the LC strategy in (21) can be recast as

$$\Phi_{\nu_L, j^L}(f_j^{LC}) = f_j^n + \frac{1 - \nu_L}{\nu_L} d_{j^L-1/2}. \quad (28)$$

Thus near local extremum, for instance as shown in Fig 2(a), compared with (27), our LC strategy provides more relaxation space, *i.e.*

$$f_j^{LC} \leq \Phi_{\nu_L, j^L}(f_j^{LC}).$$

Now thanks to the enlarged interval  $I[f_{j^L}^n, f_{j^L+1}^n, f_{j^L}^{MD}]$ , we define extremum guess as

$$m_{j^L+1/2}^{DaTe} = \min_{\alpha \in \{0, 1/2, 1\}} f_{j^L+1/2}^{MD}(\alpha), \quad M_{j^L+1/2}^{DaTe} = \max_{\alpha \in \{0, 1/2, 1\}} f_{j^L+1/2}^{MD}(\alpha). \quad (29)$$

Let us consider again the near local minimum, as shown in Fig 2(a), the choice DaTe implies

$$m_{j^L+1/2}^{DaTe} = \max \left( f_{j^L+1/2}^{FL}(1/2), f_{j^L+1/2}^{FR}(1/2) \right).$$

Evidently, we have

$$m_{j^L+1/2}^{Um} \leq m_{j^L+1/2}^{DaTe} \leq m_{j^L+1/2}^{TVD}.$$

This conclude that, compared with the choices TVD and DaTe, Um in Section 3.2 provides largest relaxation space. On the other hand, the limiting with the choices TVD and DaTe is also MP for monotonic data. The proof following the same arguments presented in Section 3.2.

At last, we remark that Tanaka [49] has adopted the original Suresh's MP limiting to the Semi-Lagrangian scheme.

We finish this subsection, by giving a remark about the Daru-Tenaud limiting. Indeed, we have two formulations: the historical one, which consists in using (24) with (27) and (25), and the formulation from subsection 3.2 using (29) (that is, without using the LC strategy). It turns out that both formulations are equivalent as shown in the following proposition. This gives also a new explanation why formula (27) is more natural than (26), as told in [21].

**Proposition 3.4.** *We have*

$$I[m_{j^L+1/2}^{DaTe}, M_{j^L+1/2}^{DaTe}] \cap I[\Phi_{\nu_L, j^L}(m_{j^L-1/2}^{DaTe}), \Phi_{\nu_L, j^L}(M_{j^L-1/2}^{DaTe})] = I[f_{j^L}^n, f_{j^L+1}^n, f_{j^L}^{MD}] \cap I[f_{j^L}^n, f_{j^L}^{UL}, f_{j^L}^{LC}],$$

with  $f_{j^L}^{MD}$  defined by (25) and  $f_{j^L}^{LC}$  defined by (27).

*Proof.* We have  $f_{j^L-1}^{MD} = f_{j^L-1/2}^{MD}(1/2) = \frac{f_{j^L}^n + f_{j^L-1}^n}{2} - \frac{d_{j^L-1/2}}{2}$ , so that

$$\begin{aligned} \Phi_{\nu_L, j^L}(f_{j^L-1}^{MD}) &= \frac{f_{j^L}^n + f_{j^L-1}^n}{2} - \frac{d_{j^L-1/2}}{2} + \frac{1}{\nu_L} \left( \frac{f_{j^L}^n - f_{j^L-1}^n}{2} + \frac{d_{j^L-1/2}}{2} \right) \\ &= f_{j^L}^n - \frac{f_{j^L}^n - f_{j^L-1}^n}{2} - \frac{d_{j^L-1/2}}{2} + \frac{1}{2\nu_L} (f_{j^L}^n - f_{j^L-1}^n + d_{j^L-1/2}) = f_{j^L}^{LC}, \end{aligned}$$

which leads to the result.  $\square$

So, concerning the Daru-Tenaud limiting, the LC initial strategy (27) can be interpreted with an extremum that satisfies the strong MP property. We can then wonder if it could be the same with the new LC strategy (28). In fact, it is not the case. We can not use  $f_{j-}^{LC}$  to define an extremum in replacement of  $m_{j^L-1/2}^{Um}/M_{j^L-1/2}^{Um}$ . As counterexample, we can take  $f_{j^L-2}^n = 0, f_{j^L-1}^n = 0, f_{j^L}^n = 1$  and  $f_{j^L+1}^n \geq 3$ ; we get  $\min(f_{j^L}^n, f_{j^L+1}^n, f_{j^L}^n - d_{j^L-1/2}) = -1$ , which violates the monotonicity property.

### 3.5. Summary of MP limiter

In this part, we give a summary of MP limiter for left flux formula, the one corresponding to right flux formula can be obtained similarly. First, compute flux  $f_{j^L+1/2}^L, j^L \in \mathbb{Z}$  by linear formula (3). Then compute local extremum guess  $m_{j^L+1/2}^{Um}/M_{j^L+1/2}^{Um}, m_{j^L+1/2}^{TVD}/M_{j^L+1/2}^{TVD}, m_{j^L+1/2}^{DaTe}/M_{j^L+1/2}^{DaTe}$  by (15), (22), (29) respectively, and compute  $f_{j-}^{LC}$  and  $f_{j+}^{LC}$  by (17) and (19) respectively. At last, the flux is limited by MP constraint as

$$f_{j^L+1/2}^L = \text{median}\left(f_{\min,j^L}, f_{j^L+1/2}^L, f_{\max,j^L}\right), j^L \in \mathbb{Z}, \quad (30a)$$

with

$$f_{\min,j^L} = \min\left(\max\left(m_{j^L+1/2}^X, \Phi_{\nu_L,j^L}(\max(M_{j^L-1/2}^Y, f_{j-}^{LC}))\right), \max\left(\min(m_{j^L+1/2}^X, f_{j+}^{LC}), \Phi_{\nu_L,j^L}(M_{j^L-1/2}^Y)\right)\right) \quad (30b)$$

and

$$f_{\max,j^L} = \max\left(\min\left(M_{j^L+1/2}^X, \Phi_{\nu_L,j^L}(\min(M_{j^L-1/2}^Y, f_{j-}^{LC}))\right), \min\left(\max(M_{j^L+1/2}^X, f_{j+}^{LC}), \Phi_{\nu_L,j^L}(M_{j^L-1/2}^Y)\right)\right) \quad (30c)$$

for  $X, Y \in \{Um, TVD, DaTe\}$ . When, we do not use the LC relaxation, the formula simplifies into

$$f_{\min,j^L} = \max\left(m_{j^L+1/2}^X, \Phi_{\nu_L,j^L}(M_{j^L-1/2}^Y)\right), f_{\max,j^L} = \min\left(M_{j^L+1/2}^X, \Phi_{\nu_L,j^L}(M_{j^L-1/2}^Y)\right). \quad (31)$$

## 4. Numerical results

In order to have an idea of the performance of the new developed schemes for 1d constant advection and Vlasov-Poisson simulations, we will focus mainly on two limiters. The first one is the default limiter, which will be called **lim** or "with limiter", corresponding to the Umeda strategy, together with Large Curvature (LC) relaxation (that is, **lim=Um+LC**: (30b),(30c), with  $X=Y=Um$ ). The second one is the Daru-Tenaud limiter without LC relaxation, which is called **DaTe**, that is (31), with  $X=Y=DaTe$ . The 2 other combinations: **Um** (Umeda strategy without LC relaxation: (31), with  $X=Y=Um$ ) and **DaTe+LC** (Daru-Tenaud limiter with LC relaxation: (30b),(30c), with  $X=Y=DaTe$ ) will be just



showed on one example, in order to see that they can lead to different numerical results. In general, the schemes that give the most different behavior are `lim` and `DaTe`. So we prefer to focus on these 2 schemes, in order not to make the presentation too heavy, and the two other combinations generally lead to results between these two schemes. Note also that `lim` is the scheme that leads to the most relaxation near extrema, while `DaTe` the least. In order to see the influence of the limiter, we will compare with the method without limiter, which will be called **no lim** or "without limiter". There is then the choice of  $d$ ; we will mainly focus on  $\mathbf{d} = \mathbf{2}$  (the degree of precision is 5) and  $\mathbf{d} = \mathbf{4}$  (degree of precision is 9), in order to see the influence of the degree in the numerical results. For  $1d$  constant advection, we will consider also higher degrees (typically  $d = 8$ ), but we experimented drop of efficiency for the limited schemes when going to higher degrees, in particular on Vlasov-Poisson simulations. So, we prefer to stick to  $d = 2$  and  $d = 4$ , where we already can see the drop of efficiency. In order to see how the scheme behaves on the different test cases in comparison to other schemes, we have selected the **cubic splines** method (as standard unlimited scheme, already used in [10]) and the fifth-order conservative semi-Lagrangian WENO (which will be called **SLWENO5**) scheme (as example of limited scheme developed in [44]; see [41] for a recent application of such scheme; we use here the implementation given in [40], with  $\varepsilon = 10^{-6}$  instead of  $\varepsilon = 10^{-10}$ ).

#### 4.1. Implementation issues

We give part of an implementation in the language `C` of the limiter in Figure 4; we have chosen here the right flux and take the example of the limiter with `lim=Um+LC`. We have not tried to fully optimize the code, but we remark that the limiter has a moderate overhead of computation, thanks to some reuse of computation, in this flux form. Note that some implementations can lead to large floating point errors and unsymmetric results; we have tried to limit this, in particular, by avoiding to have  $1/\nu$  factors; we also choose to switch to first order limiter if the difference with it is very small.

#### 4.2. Free transport equation

We first consider the classical constant advection equation  $\partial_t f + \partial_x f = 0$  with initial condition  $f(t = 0, x) = f_0(x)$  on the periodic domain  $[-1, 1]$  and for  $t \in [0, T]$ , with  $T \in \mathbb{R}^+$ , the final time. Spatial mesh is  $x_i = -1 + ih$ ,  $i = 0, \dots, N$ , with spatial step  $h = \frac{2}{N}$ , and  $N \in \mathbb{N}^*$  is the number of cells. The time step is  $\Delta t = T/M \geq 0$ , with  $M \in \mathbb{N}^*$ , the number of time steps. We first focus on a **square wave** test, in order to see the effect of the limiter. Then we look at smooth tests: first a **sinusoidal wave**, which is the canonical smooth function and a **quartic sine function**, whose solution should remain positive. Finally, we consider a test introduced in [38], called here **Shu test case**, to test how WENO type schemes are able to handle different types of oscillations. Note that the such tests were chosen in [49] and [57] for example.

For the constant advection, no error is introduced in time, as the characteristics are exactly solved in time. It is in fact better in this context to use semi-Lagrangian schemes than Eulerian schemes, the latter introducing an error in

```

flux2 = 0.;
for (ii=-d;ii<=d;ii++) flux2+=w[ii]*q[ii];
flux2 = nu*flux2;

//TVD
fmin[1] = DMIN(q[0],q[1]);
fmax[1] = DMAX(q[0],q[1]);
fmin[0] = DMIN(q[-1],q[0]);
fmax[0] = DMAX(q[-1],q[0]);
//Um
fmin[1] = DMIN(fmin[1],DMAX(2.*q[0]-q[-1],2.*q[1]-q[2]));
fmax[1] = DMAX(fmax[1],DMIN(2.*q[0]-q[-1],2.*q[1]-q[2]));
fmin[0] = DMIN(fmin[0],DMAX(2.*q[-1]-q[-2],2.*q[0]-q[1]));
fmax[0] = DMAX(fmax[0],DMIN(2.*q[-1]-q[-2],2.*q[0]-q[1]));
//fLC
dj = minmod(q[-1]-2.*q[0]+q[1],q[0]-2.*q[1]+q[2]);
fmin2[1] = DMIN(fmin[1],q[0]-dj);
fmax2[1] = DMAX(fmax[1],q[0]-dj);
fmin3[0] = DMIN(fmin[0],q[0]-minmod(q[-1]-2.*q[0]+q[1],q[-2]-2.*q[-1]+q[0]));
fmax3[0] = DMAX(fmax[0],q[0]-minmod(q[-1]-2.*q[0]+q[1],q[-2]-2.*q[-1]+q[0]));

bound0 = DMAX(nu*fmin3[0],q[0]-(1.-nu)*fmax[1]);
bound0 = DMIN(bound0,DMAX(nu*fmin[0],q[0]-(1.-nu)*fmax2[1]));
bound1 = DMIN(nu*fmax3[0],q[0]-(1.-nu)*fmin[1]);
bound1 = DMAX(bound1,DMIN(nu*fmax[0],q[0]-(1.-nu)*fmin2[1]));

flux2 = DMIN(flux2,bound1);
flux2 = DMAX(bound0,flux2);
for(i=0;i<N;i++){
    flux1 = flux2;
    q++;
    flux2 = 0.;
    for (ii=-d;ii<=d;ii++)flux2+=w[ii]*q[ii];
    flux2 = nu*flux2;

    //TVD
    fmin[0] = fmin[1];
    fmax[0] = fmax[1];
    fmin[1] = DMIN(q[0],q[1]);
    fmax[1] = DMAX(q[0],q[1]);
    //Um
    fmin[1] = DMIN(fmin[1],DMAX(2.*q[0]-q[-1],2.*q[1]-q[2]));
    fmax[1] = DMAX(fmax[1],DMIN(2.*q[0]-q[-1],2.*q[1]-q[2]));
    //fLC
    fmin3[0] = DMIN(fmin[0],q[0]-dj);
    fmax3[0] = DMAX(fmax[0],q[0]-dj);
    dj = minmod(q[-1]-2.*q[0]+q[1],q[0]-2.*q[1]+q[2]);
    fmin2[1] = DMIN(fmin[1],q[0]-dj);
    fmax2[1] = DMAX(fmax[1],q[0]-dj);

    bound0 = DMAX(nu*fmin3[0],q[0]-(1.-nu)*fmax[1]);
    bound0 = DMIN(bound0,DMAX(nu*fmin[0],q[0]-(1.-nu)*fmax2[1]));
    bound1 = DMIN(nu*fmax3[0],q[0]-(1.-nu)*fmin[1]);
    bound1 = DMAX(bound1,DMIN(nu*fmax[0],q[0]-(1.-nu)*fmin2[1]));

    flux2 = DMIN(flux2,bound1);
    flux2 = DMAX(bound0,flux2);
    if(fabs(flux2-nu*q[0])<1.e-16) flux2 = nu*q[0];
    p[i] = q[-1]-(flux1-flux2);
}

```

time (see [49] for such a study).

The error is governed by the CFL number (defined here by  $CFL = \Delta t/h$ ), the final time  $T$  and the number of cells  $N$ . We take as default parameters  $CFL = 2.5$ , a quite big final time  $T = 800$ , in order to better see the difference between the schemes. We make then a convergence study with respect to  $N$  (as in [57]) for the three first tests (for the last one, the convergence study is more difficult to interpret, as different numerical errors coming from different smoothness are mixed). As quantities of interest, we look at the error in  $L^1$  and  $L^2$  norm and also at the total variation (TV) error for the square wave (which is defined as the total variation minus 2). In order to see the influence of the CFL number, we change also this number for the square wave test (taking  $CFL = 0.25$  and  $CFL = 0.025$ ). Note that taking smaller CFL number leads generally to bigger error, as more interpolations are used, but only to a certain limit (as the interpolation gets more accurate when the foot of the characteristic approaches a grid point), which in fact corresponds to the Eulerian scheme (see e.g. [18]) where the time error becomes negligible (see [45] for a study where the number of interpolation is fixed and only the CFL number is changed). For the convergence study, we only consider the  $\text{lim}=\text{Um}+\text{LC}$  scheme and take different values of  $d$ , in order to see the influence of the degree. For completing the numerical results, we consider also plots of the numerical solution together with the exact solution for the square wave and for the Shu test case (where the errors are the biggest and thus the plots lead to visible difference between the schemes).

For the square wave, we can distinguish the behavior between the limiters at low resolution  $N = 25$  and  $CFL = 0.25$ , for different values of  $d$  ( $d = 3, 6, 8$ ). The Shu test case permits also to appreciate the qualitative behavior of the different schemes, taking the standard value  $N = 200$ .

The different numerical results are summarized here.

- Square wave:
  - Initial condition:  $f_0(x) = 1$ , if  $x \in [-0.75, 0.25]$  and  $f_0(x) = 0$ , if  $x \in [-1, -0.75[ \cup ]0.25, 1]$ .
  - On Table 2, representation of the error in  $L^1$ ,  $L^2$  norm, its order (in  $L^1$  and  $L^2$  norm), and the error in total variation (TV) (defined here as the total variation minus 2), for  $CFL = 2.5$ ,  $0.25$  and  $0.025$ , for SLWENO5, cubic splines,  $d = 2$  and  $d = 8$  with and without limiter.
  - On Table 3 (top), comparison of the limiter  $\text{lim}=\text{Um}+\text{LC}$ , with other combinations:  $\text{DaTe}, \text{Um}$  and  $\text{Date}+\text{LC}$   $d = 3, 6$  and  $8$ .
- Sinusoidal wave:
  - initial condition:  $f_0(x) = \sin(\pi x)$
  - On Table 3 (middle), representation of the error in  $L^1$ ,  $L^2$  norm, its order  $r$  (in  $L^1$  and  $L^2$  norm) for  $CFL = 2.5$ , for SLWENO5, cubic splines,  $d = 1, \dots, 4$  with limiter.

- Quartic sine function

- initial condition:  $f_0(x) = \sin^4(4\pi x)$
- Numerical results on Table 3 (bottom).

- Shu test

- Initial condition is given by

$$f_0(x) = \begin{cases} \frac{1}{6} (G(x, \beta, z - \delta) + G(x, \beta, z + \delta) + 4G(x, \beta, z)), & x \in [-0.8, -0.6] \\ 1, & x \in [-0.4, -0.2] \\ 1 - |10(x - 0.1)|, & x \in [0, 0.2] \\ \frac{1}{6} (F(x, \alpha, a - \delta) + F(x, \alpha, a + \delta) + 4F(x, \alpha, a)), & x \in [0.4, 0.6] \\ 0, & x \in [-1, -0.8] \cup [-0.2, 0] \cup [0.2, 0.4] \cup [0.6, 1] \end{cases}$$

with  $G(x, \beta, z) = e^{-\beta(x-z)^2}$  and  $F(x, \alpha, a) = \sqrt{\max(1 - \alpha^2(x - a)^2, 0)}$ ,  
and  $a = 0.5$ ,  $z = -0.7$ ,  $\delta = 0.005$ ,  $\alpha = 10$  and  $\beta = \frac{\ln(2)}{36\delta^2}$ .

- Numerical results on Figure 5

One first observation is that the new limiting scheme is converging with (almost) the expected order of the unlimited scheme (see [26, 57]) which is  $\frac{2d+1}{2d+2}$  in  $L^1$  and the half value in  $L^2$  for the square wave and  $2d + 1$  for the smooth functions. For the square wave, for  $d = 1, 3, \dots, 7$ , we have also found that the order in  $L^1$  is around 0.75, 0.86, 0.89, 0.90, 0.91, 0.92 and half for the  $L^2$  norm. We also observe that the limited and unlimited scheme give generally the same results on the smooth functions (we can however notice some differences, taking  $N \leq 10$  for the sinusoidal wave). For the quartic sine function, we note that the maximum principle is not satisfied (as the scheme follows the unlimited one and is not able to preserve the positivity, as it was already the case in [49]) and a further study enforcing global maximum principle would be worth to be added, but is not tackled here, in order to already see the sole influence of the monotonicity preserving property. On the contrary, the SLWENO5 scheme (which is known to be more diffusive) does not converge with the expected order for the square wave. Taking however a smaller final time  $T = 8$  and  $CFL = 0.25$ , we observe an order of convergence in  $L^1$  of 0.83 (in accordance to [57]), which is the one expected. For the sinusoidal wave, the 5th order of convergence is reached with however a bigger error than the corresponding unlimited scheme (that is  $d = 2$ , in other words, the WENO scheme with linear weights), but not for the quartic sine function, as the mesh is certainly not fine enough. So, with respect to the SLWENO5 scheme, the new scheme behaves favorably and is less diffusive as already mentioned in [49] and even [48]. We also remark that the schemes, being higher order, behave favorably with respect to cubic splines, except of course for  $d = 1$ . In that case ( $d = 1$ ), cubic splines have the same order of accuracy, but with a better constant: the accuracy is similar between cubic splines and  $d = 1$  with twice more points (as example, we can look at the error in  $L^2$  norm for the sinusoidal wave: 0.0413 for  $N = 25$  with cubic splines, which is comparable to 0.0456 for  $N = 50$  with the Lagrange method

using  $d = 1$ ).

As seen on the square wave, the total variation error (TV) is increasing (in absolute value) with the degree for the unlimited scheme. On the contrary, for the limited scheme, the total variation is very well preserved, but we can observe some little degradation for the small value of  $CFL = 0.025$  and when  $N$  is small. This can be due to the fact that the numerical scheme can diffuse the solution and then the solution becomes smooth and so the limiter do no more act. This can sometimes lead to negative values of TV, as already observed in [49]. Furthermore, we did in fact sometimes encounter situations where new extrema were found, but this was due to the propagation of round off errors, which has lead us to do a dedicated modification previously explained. For the SLWENO5 scheme, we observe that the total variation error is also small, but it is not as well preserved as for the limited scheme.

So, the limiter permits clearly to improve the total variation error. The introduction of numerical oscillations inherent to the high order unlimited scheme are removed. On the other hand, for smooth initial data and fine enough mesh, the limiter does not act and we recover the performance of the high order unlimited scheme. We can see some improvements for the limiter  $\text{lim} = \text{Um} + \text{LC}$ , and see the advantage of  $\text{Um}$  w.r.t DaTe and also the use of LC limitation for not too large degree. Furthermore, there are some other conclusions from observations. On the one hand, for the higher degree, the error does not always improve, which is not the case for the unlimited scheme which has proven  $L^p$ ,  $p \geq 1$  convergence even for non smooth data [26], with improvements when the degree get higher and higher, even if the total variation is of course not improved. On the other hand, the test for detecting smooth extrema might be improved when a large stencil is considered. In principle, we could use also the scheme for even degree interpolation or for cubic or higher order splines. But it seems that both the larger stencil and dispersion effects seem to lead to bad behavior of the scheme. Particularly, note that in [48, 21, 49],  $d = 2$  or  $d = 3$  are used and odd degree is recommended, because it favors diffusion over dispersion.

CFL=2.5											
SLWENO5						cubic splines					
$N$	$L^1$	r	$L^2$	r	TV	$N$	$L^1$	r	$L^2$	r	TV
25	0.339		0.321		0.129	25	0.329		0.302		0.415
50	0.216	0.65	0.245	0.39	0.000164	50	0.201	0.71	0.231	0.38	0.417
100	0.123	0.81	0.183	0.42	0.00065	100	0.124	0.69	0.179	0.37	0.486
200	0.0744	0.73	0.138	0.4	0.000256	200	0.0738	0.75	0.138	0.37	0.486
400	0.065	0.19	0.138	0.0063	0.00499	400	0.0437	0.75	0.107	0.37	0.483
800	0.0575	0.18	0.132	0.061	0.00117	800	0.026	0.75	0.0821	0.37	0.479
$d = 2$ with limiter						$d = 2$ without limiter					
$N$	$L^1$	r	$L^2$	r	TV	$N$	$L^1$	r	$L^2$	r	TV
25	0.26		0.279		0	25	0.326		0.298		0.493
50	0.146	0.83	0.21	0.41	0	50	0.194	0.75	0.214	0.48	0.766
100	0.0825	0.83	0.158	0.41	0	100	0.108	0.84	0.159	0.43	0.784
200	0.0464	0.83	0.119	0.41	0	200	0.0614	0.82	0.12	0.41	0.799
400	0.0261	0.83	0.0891	0.41	0	400	0.0344	0.84	0.0899	0.41	0.801
800	0.0147	0.83	0.0669	0.41	4.44e-16	800	0.0194	0.83	0.0674	0.41	0.8
$d = 8$ with limiter						$d = 8$ without limiter					
$N$	$L^1$	r	$L^2$	r	TV	$N$	$L^1$	r	$L^2$	r	TV
25	0.133		0.191		0	25	0.159		0.163		1.19
50	0.07	0.92	0.14	0.45	0	50	0.083	0.94	0.119	0.45	1.41
100	0.0371	0.92	0.102	0.45	0	100	0.0476	0.8	0.0878	0.44	1.63
200	0.0196	0.92	0.0743	0.46	0	200	0.0246	0.95	0.064	0.46	1.64
400	0.0103	0.92	0.0541	0.46	4.44e-16	400	0.0129	0.93	0.0466	0.46	1.66
800	0.00543	0.93	0.0393	0.46	0	800	0.0069	0.91	0.0339	0.46	1.66

CFL=0.25											
SLWENO5						cubic splines					
$N$	$L^1$	r	$L^2$	r	TV	$N$	$L^1$	r	$L^2$	r	TV
25	0.436		0.364		-0.233	25	0.364		0.334		0.00555
50	0.331	0.4	0.296	0.3	-0.201	50	0.318	0.19	0.297	0.17	0.41
100	0.437	-0.4	0.375	-0.34	-0.0354	100	0.189	0.75	0.223	0.41	0.482
200	0.323	0.43	0.315	0.25	0.00369	200	0.114	0.73	0.172	0.38	0.483
400	0.234	0.46	0.265	0.25	0.00564	400	0.0674	0.75	0.132	0.38	0.484
800	0.161	0.55	0.22	0.27	0.00806	800	0.04	0.75	0.102	0.38	0.481
$d = 2$ with limiter						$d = 2$ without limiter					
$N$	$L^1$	r	$L^2$	r	TV	$N$	$L^1$	r	$L^2$	r	TV
25	0.338		0.32		0.146	25	0.337		0.306		0.451
50	0.205	0.72	0.25	0.36	4.44e-16	50	0.237	0.51	0.25	0.29	0.256
100	0.116	0.83	0.188	0.41	4.44e-16	100	0.146	0.7	0.189	0.4	0.721
200	0.0649	0.83	0.141	0.41	0	200	0.0858	0.77	0.142	0.41	0.802
400	0.0365	0.83	0.106	0.42	4.44e-16	400	0.0482	0.83	0.106	0.42	0.803
800	0.0205	0.83	0.0791	0.42	-2.22e-16	800	0.0271	0.83	0.0798	0.42	0.805
$d = 8$ with limiter						$d = 8$ without limiter					
$N$	$L^1$	r	$L^2$	r	TV	$N$	$L^1$	r	$L^2$	r	TV
25	0.165		0.213		0.0389	25	0.145		0.173		0.95
50	0.0824	1	0.152	0.48	0	50	0.101	0.52	0.132	0.39	1.62
100	0.0433	0.93	0.111	0.46	4.44e-16	100	0.0544	0.89	0.0956	0.47	1.64
200	0.0227	0.93	0.0805	0.46	4.44e-16	200	0.0292	0.9	0.0695	0.46	1.69
400	0.0119	0.93	0.0584	0.46	0	400	0.0152	0.94	0.0505	0.46	1.65
800	0.00624	0.93	0.0424	0.46	-2.22e-16	800	0.00807	0.91	0.0366	0.46	1.56

CFL=0.025											
SLWENO5						cubic splines					
$N$	$L^1$	r	$L^2$	r	TV	$N$	$L^1$	r	$L^2$	r	TV
25	0.457		0.373		-0.296	25	0.341		0.311		0.456
50	0.351	0.38	0.307	0.28	-0.193	50	0.285	0.26	0.272	0.19	0.635
100	0.326	0.11	0.306	0.0029	-0.0316	100	0.183	0.64	0.208	0.39	1.23
200	0.279	0.23	0.303	0.016	0.00122	200	0.111	0.72	0.158	0.4	1.31
400	0.204	0.45	0.253	0.26	0.00482	400	0.0625	0.83	0.119	0.4	1.22
800	0.145	0.49	0.212	0.26	0.00805	800	0.035	0.84	0.0903	0.4	1.09
$d = 2$ with limiter						$d = 2$ without limiter					
$N$	$L^1$	r	$L^2$	r	TV	$N$	$L^1$	r	$L^2$	r	TV
25	0.342		0.323		0.12	25	0.337		0.306		0.437
50	0.211	0.69	0.253	0.35	-1.73e-06	50	0.249	0.44	0.256	0.26	0.266
100	0.119	0.83	0.19	0.41	-9.9e-13	100	0.149	0.74	0.191	0.42	0.681
200	0.0667	0.83	0.143	0.41	-4e-14	200	0.0883	0.76	0.144	0.41	0.806
400	0.0375	0.83	0.107	0.42	-4.06e-14	400	0.0496	0.83	0.108	0.42	0.81
800	0.021	0.83	0.0802	0.42	-4.15e-14	800	0.0278	0.83	0.0809	0.42	0.805
$d = 8$ with limiter						$d = 8$ without limiter					
$N$	$L^1$	r	$L^2$	r	TV	$N$	$L^1$	r	$L^2$	r	TV
25	0.216		0.22		0.165	25	0.149		0.175		0.958
50	0.107	1	0.165	0.42	-0.000688	50	0.103	0.53	0.133	0.39	1.66
100	0.0565	0.92	0.121	0.45	6.8e-10	100	0.0562	0.88	0.0965	0.47	1.72
200	0.0294	0.94	0.0872	0.47	1.27e-09	200	0.0303	0.89	0.0701	0.46	1.78
400	0.0154	0.94	0.0628	0.47	5.68e-09	400	0.0158	0.94	0.0509	0.46	1.72
800	0.00809	0.92	0.0458	0.46	3.41e-09	800	0.00839	0.91	0.0369	0.46	1.68

Table 2: Error for square wave at  $T = 800$  in  $L^1$  and  $L^2$  norm, with corresponding order (denoted by  $r$ ) and in TV (total variation)

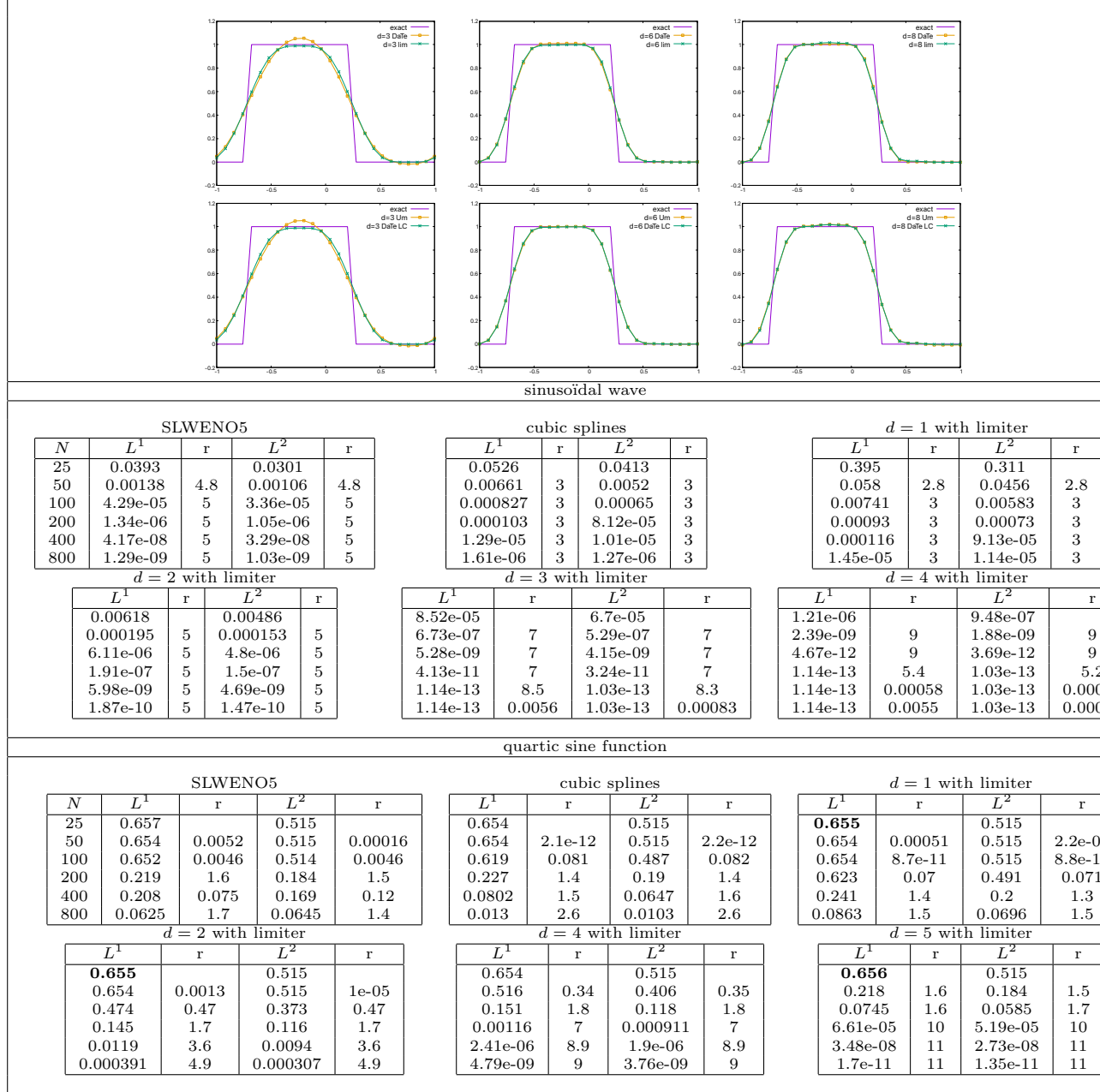


Table 3: On top: square wave,  $T = 800$ ,  $N = 25$ ,  $CFL = 0.25$  (left,  $d = 3$ , center,  $d = 6$  and right  $d = 8$ ); then, error for sinusoidal wave and quartic sine function,  $CFL = 2.5$ ,  $T = 800$  in  $L^1$  and  $L^2$  norm, with corresponding order denoted by  $r$ . The values in bold are the sole values that differ with the schemes without limiter; in that case the value without limiter is always 0.654

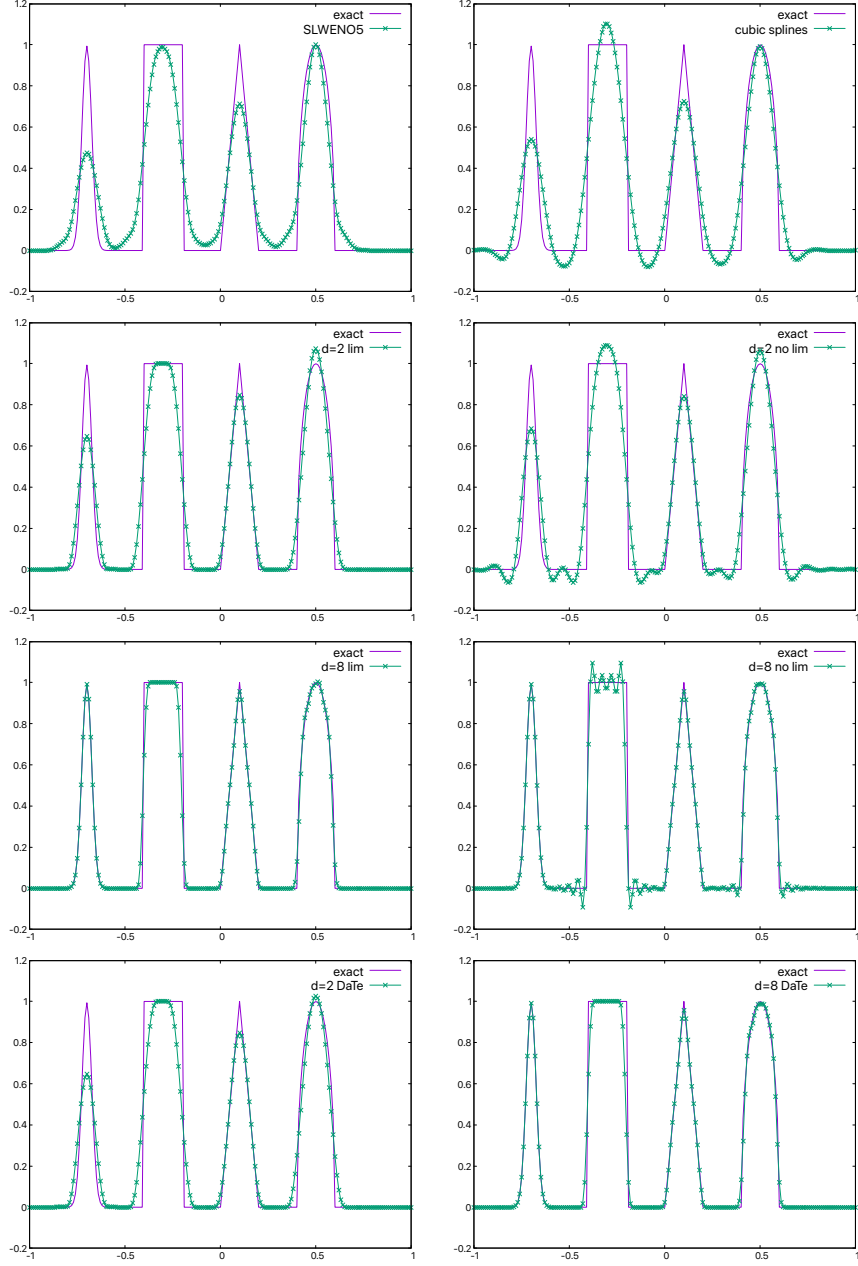


Figure 5: Shu test case,  $T = 800$ ,  $N = 200$ ,  $CFL = 2.5$



### 4.3. Vlasov-Poisson system

We solve the Vlasov equation  $\partial_t f + v \partial_x f + E \partial_v f = 0$ , coupled with the Poisson equation;  $E = -\partial_x \phi$  and  $-\partial_x^2 \phi = \rho - 1$ , using Strang splitting. We also have an initial condition  $f(t = 0, x, v) = f_0(x, v)$ . We can either first do the advection in  $x$ :  $\partial_t f + v \partial_x f = 0$  or the advection in  $v$ : Poisson equation and then  $\partial_t f + E \partial_v f = 0$ . In the numerical results we consider to do at each time step the advection in  $x$  for  $\Delta t/2$ , the advection in  $v$  for  $\Delta t$  and then again the advection in  $x$  for  $\Delta t/2$ . Then unknowns are  $f_{i,j}^n \simeq f(t_n, x_i, v_j)$ , with  $t_n = n\Delta t$ . The phase-space domain is  $[0, L] \times [-v_{\max}, v_{\max}]$ , and we have  $x_i = i\Delta x$ ,  $i = 0, \dots, N_x$ , with  $\Delta x = L/N_x$  and  $v_j = -v_{\max} + j\Delta v$ ,  $j = 0, \dots, N_v$ , with  $\Delta v = 2v_{\max}/N_v$ , with  $N_x, N_v \in \mathbb{N}^*$ . We also have  $t \in [0, T]$ , with  $T \in \mathbb{R}^+$ , the final time, and the time step is  $\Delta t = T/M \geq 0$ , with  $M \in \mathbb{N}^*$ , the number of time steps.

Our aim is to evaluate the possible benefit of using our limiting strategy in the context of Vlasov-Poisson simulations, by comparing with the SLWENO5 strategy (which has also the property of limiting the oscillations) and unlimited schemes (cubic splines, as reference scheme, and the methods without activating the limiting procedure). We have to check that the limiting strategy reasonably works well for Vlasov-Poisson simulations and that the damping induced by the limiter to prevent from numerical oscillations (inherent to the high order unlimited schemes) does not degrade the numerical solution. In order to show an interest of the method, we also have to find situations where the numerical oscillations become critical and the limiting strategy has a real beneficial effect. We restrict the study to  $1d \times 1d$  single species Vlasov-Poisson simulations using directional splitting with constant advection on uniform mesh and without collisions (we refer for example to [15] for multi-species applications with sharp gradients and non-uniform meshes, that could be further studied).

In the sequel, we consider first two standard test cases: the strong non linear Landau damping and the two stream instability. We have not considered the linear Landau damping, but we expect better conservation error than in the strong non linear case (see for example [45]). Then, we consider the more challenging bump on tail test case, with three vortices. After that, we consider again a two stream instability test case, but this time with 13 vortices. Finally, we consider a beam test case initialized with (almost) discontinuous data and a plasma expansion test.

We will compare always the same 8 methods: SLWENO5 and cubic splines (as reference schemes), lim=Um+LC, DaTe and "no lim" for  $d = 2$  and  $d = 4$  (for seeing the influence of the limiters and the degree). We will change the phase space resolution in order to see the influence of the grid (it is in particular important to see the results on coarse grids, since the refined grids are possible here in  $1d \times 1d$ , but maybe not in higher dimensions). The time step will be kept fixed to standard value  $\Delta t = 0.1$ , and we stick to Strang splitting (higher order may be envisaged [8, 3], but we focus here on the error made in phase space, which is by the way often the most significant).

#### 4.3.1. Strong non linear Landau damping [20]

Initial condition is  $f_0(x, v) = (1 + 0.5 \cos(x/2)) \frac{1}{\sqrt{2\pi}} e^{-v^2/2}$ . The domain is  $[0, 4\pi] \times [-v_{\max}, v_{\max}]$ , with  $v_{\max} = 6$ . We use  $\Delta t = 0.1$  and final time  $T = 100$ . Numerical results are given on Figure 6. We represent the relative total energy error, using initial analytical value. We remark that the limited schemes lead to an increase of the total energy error; this was indeed also already observed in [20]. Note that the results are kept in reasonable and comparable range as in the literature [44, 45]. Lim behaves favorably w.r.t DaTe; SLWENO5 has the worst conservation. The results improve on finer grids, but not with the degree. The  $L^1$  norm (not shown) is better conserved for limited schemes, which is expected. The behavior of DaTe is similar to that of lim, with slight improvement of  $L^1$  norm conservation. SLWENO5 behaves differently with better  $L^1$  norm conservation. Changing the degree does not change a lot the results and going to finer grids permits to have better results at longer times, but then there is no improvement over even longer time.

The limited schemes induce some damping and generate more decrease of the  $L^2$  norm (not shown) as already observed in [20]. The level of saturation of relative  $L^2$  norm error is around 10%, for all the methods, the same as the one observed in [20]. The SLWENO5 method diffuses much more and going to higher order methods and finer grids leads to better  $L^2$  norm conservation for longer times. The difference w.r.t the unlimited scheme is more pronounced for the high degree case, and lim is better than DaTe.

#### 4.3.2. Two stream instability I [45]

We consider initial condition as  $f_0(x, v) = \frac{v^2}{\sqrt{8\pi}} (2 - \cos(x/2)) e^{-v^2/2}$ . The domain is  $[-2\pi, 2\pi]^2$ . Numerical results are shown on Figure 7. We represent the time evolution of the  $L^2$  norm error. We remark that the behavior of the unlimited and limited schemes are very similar; there is less difference than for the strong non linear Landau damping test case, maybe because the solution is less oscillatory. For total energy and  $L^1$  norm (not shown), the results are quite similar with the Landau damping (with better total energy conservation: around  $10^{-3}$  on the finest grids for the limited schemes).

#### 4.3.3. Bump on tail [20]

As already mentioned in [51], the effect of limiters is not so much visible on the two previous test cases. We now move to more delicate cases, where we look also for longer time, where convergence is not reached. Initial condition is  $f_0(x, v) = (1 + \varepsilon \cos(kx)) (\frac{0.9}{\sqrt{2\pi}} e^{-v^2/2} + \frac{0.1}{v_{th} \sqrt{2\pi}} e^{-(v-u)^2/(2v_{th}^2)})$ , with  $u = 4.5$ ,  $\varepsilon = 0.04$ ,  $v_{th} = 0.5$  and  $k = 0.3$ . The domain is  $[0, 3\frac{2\pi}{k}] \times [-v_{\max}, v_{\max}]$ , with  $v_{\max} = 9$ . We use  $\Delta t = 0.1$  and final time  $T = 400$ . Numerical results are given on Figure 8.

In this test, we have three small vortices, and generally only one will remain if we wait long enough. In order to discretize them well in space, we use a number of points in  $x$  that is a multiple of 3; otherwise, the method with limiter is not able to keep the three vortices and some merging appears leading also to

a break in electric energy, as observed in [20] for some methods. It has been indeed observed in [39] that it is important that the number is a multiple of three to represent correctly the periodicity of the solution; see also [16] where a multiple of three is used.

The electric energy is shown on Figure 8. We observe that the SLWENO5 scheme damps much more, it is similar to Lagrange interpolation of degree 3 (see [20]). Other limited schemes show also more damping than unlimited schemes, but not as much as SLWENO5; they follow more the unlimited schemes, with lim better than DaTe. For  $d = 4$ , we see again more differences between limited and unlimited schemes, especially on coarse grids. The  $L^1$  norm (not shown) is already quite well preserved on fine grids for the unlimited schemes. On the coarsest grid, we see an improvement for limited schemes, in particular, for  $d = 4$ .

Concerning the  $L^2$  norm (not shown), apart from SLWENO5 which diffuses much more, the limited schemes show quite good conservation. There is an improvement with the degree of interpolation (going from  $d = 2$  to  $d = 4$ ) on the coarsest grid for long time, but no more on finer grids.

The total energy (not shown) is again less conserved for the limited schemes. The situation improves, when the grid gets finer (certainly the limiter acts less). On the other hand, for  $d = 4$ , the situation can be worse than for  $d = 2$ . The scheme has to deal with a less damped solution at the beginning and has then more difficulties to handle it for longer time.

#### 4.3.4. Two stream instability II [20]

Initial condition is  $f_0(x, v) = (1 + \varepsilon \cos(kx))(\frac{1}{2v_{th}\sqrt{2\pi}}e^{-(v-u)^2/(2v_{th}^2)} + \frac{1}{2v_{th}\sqrt{2\pi}}e^{-(v+u)^2/(2v_{th}^2)})$ , with  $\varepsilon = 0.05$ ,  $u = 0.99$ ,  $v_{th} = 0.3$  and  $k = 1$ . The domain is  $[0, 26\pi] \times [-5, 5]$ . Numerical results are given on Figures 9, 10, 11, 12. We use  $\Delta t = 0.1$ .

We have here 13 vortices and we wait long enough, so that the vortices merge, a priori from numerical errors (we have no reference solution for this test until  $T = 1000$ , but have run a reference solution valid until time  $T = 70$ , on a  $4096 \times 4096$  grid, with 6-th order time splitting [8]). On Figures 9, 10, we represent the electric energy. We can appreciate the convergence of the methods until time  $T = 70$  (the latter reference run is plotted in red). We clearly see the effect of the limiting, which permits to damp the solution. Solutions at time  $T = 70$  (for the reference solution [computed on  $4096 \times 4096$  grid with high order splitting in time [8] leading to same result than with Strang splitting], which is not shown to save place, we have still 13 vortices) are given on Figure 11. We see for all the methods, that the merging is already occurring. At final time  $T = 1000$  (Figure 12), we have generally 2 vortices that are remaining, but the results different depending on the methods and grids. We observe more oscillations for the unlimited schemes on the coarsest grids. SLWENO5 and DaTe with  $d = 2$  can lead to only one vortex at the final time. There is more diffusion, because the limiter acts more.

#### 4.3.5. Beam test case [57]

Initial condition is  $f_0(x, v) = \frac{4}{\sqrt{2\pi\alpha}}\chi(x)e^{-v^2/(2\alpha)}$ , with  $\chi(x) = \frac{1}{2}\text{erf}(\frac{x+1.2}{\beta}) - \frac{1}{2}\text{erf}(\frac{x-1.2}{\beta})$  and  $\alpha = 0.2$ . We choose  $\beta = 10^{-10}$  (instead of 0.3 in [57], so that the initial data is (almost) discontinuous). Also we do not solve exactly the previous Vlasov-Poisson system, but  $\partial_t f + \frac{v}{\varepsilon}\partial_x f + (qE - \frac{x}{\varepsilon})\partial_v f = 0$ , together with electric field  $E$  satisfying  $\frac{1}{x}\partial_x(xE) = \int f dv$ , on  $[0, L/2]$  and imposing that  $E$  is odd function.  $E$  is given by  $E = \frac{1}{x}\int_0^x s\rho(t, s)dx$  and we use a trapezoidal formula for the approximation of the integral. We choose  $\varepsilon = 0.7$  and the domain is  $[-L/2, L/2] \times [-v_{\max}, v_{\max}]$ , with  $L = 8$  and  $v_{\max} = 4$ . We choose  $q = 0.01$ , instead of  $q = 1$  in [57], in order to give a strong weight to the rotation. We take  $T = 1000$ , instead of  $T = 100$ , in [57], in order to look for time long enough. Note that we consider always the Strang splitting. We could however use more adequate splitting, as for example the recently developed exact splitting [3], but (as already said) as we consider here mainly the reconstruction in phase space, we do not change here the scheme in time and also do not change the time step. Numerical results are given on Figure 13, 14, 15. Initially, there are numerical oscillations due to the discontinuous profile of the initial condition, leading to much better  $L^1$  norm conservation for the limited schemes than for the unlimited ones. However, the oscillations of the unlimited schemes permit to handle more efficiently the solution in the long run, and unfortunately, the limited scheme with high order interpolation (here  $d = 4$ ) present bad qualitative behavior (Figure 13), in particular on coarse grids. Such phenomenon does not occur for  $d = 2$  and also for SLWENO5, as the solution is more diffused. Again, lim is better than DaTe. So we recommend to use the limiting strategy for  $d = 2$ ; for higher  $d$ , we can have some troubles; this can be due to dispersion effects. Note that the method with limiter nevertheless still converges as seen also on the  $L^2$  norm (Figure 15) and electric energy conservation (Figure 14), but the convergence is faster for  $d = 2$  than for  $d = 4$ .

#### 4.3.6. Plasma expansion [29]

Initial condition is  $f_0(x, v) = \frac{1}{\sqrt{2\pi}}\chi(x-2\pi)e^{-v^2/2}$ , with  $\chi(x) = \frac{1}{2}\text{erf}(\frac{x+0.5}{\beta}) - \frac{1}{2}\text{erf}(\frac{x-0.5}{\beta})$ . We choose  $\beta = 10^{-10}$ , so that the initial data is almost discontinuous (in [29], the profile was Maxwellian in  $x$  and  $v$ , here we have considered a sharper profile in  $x$ ). The domain is  $[0, 4\pi] \times [-v_{\max}, v_{\max}]$ , with  $v_{\max} = 6$ . We use as usually  $\Delta t = 0.1$ , with Strang splitting. Final time is  $T = 400$ . Numerical results are shown on Figures 16, 17, 18, 19, 20, 21. Note that for this test, we have no convergence at final time, as indicated by the electric energy (not shown); there the convergence is around  $T = 20$ , on the  $4096 \times 4096$  grid and around  $T = 10$  on the  $512 \times 512$  grid. On Figures 16, 17, we see the distribution function at time  $T = 400$ . On Figures 18, 19, the relative error for the  $L^1$  norm is given. On Figures 20, 21, the relative error for the  $L^2$  norm is given. We clearly see a difference for the cubic splines case: it better conserves the  $L^2$  norm, while having a very bad conservation of  $L^1$  norm; this was indeed already observed in [29]. We also see on Figures 16, 17, that the oscillations are kept for the splines, and there is more diffusion for the other methods. The

Lagrange methods without limiters have better  $L^1$  norm conservation (but still not good), at the price of more diffusion. Finally the Lagrange methods with limiters have really better  $L^1$  norm conservation, but with even more diffusion. Note that the methods go to high diffusion with time; we can only improve the results for a fixed final time, by refining the grid. Concerning the total energy (not shown), it is again always better preserved without limiters, but one nice feature is that it gets better for the limited schemes, while refining the grids; all the schemes go to a certain saturation around  $10^{-3}$ , which could be further enhanced by taking a smaller time step or considering a higher order in time method (see for example [8] for a study of the behavior of the error for high order time splitting schemes).

## Conclusion and perspectives

We have revisited the monotonicity preserving schemes for semi-Lagrangian schemes based on odd order Lagrange interpolation. A detailed numerical study is performed for  $1d$  constant advection and Vlasov-Poisson simulations. The new scheme has a proven monotonicity preserving property and controls in particular the  $L^1$  norm, with some limited degradation of the  $L^2$  norm. Comparison with cubic splines and SLWENO5 is made to show the accuracy of our method. It turns out that the method is not completely satisfactory for high order interpolation, and we recommend to use it for  $d = 2$ , which is still less diffusive than using SLWENO5. A better test for the detecting of smooth extrema, using a wider stencil (which fits well for  $d = 2$ ) might be considered for having better results. The use of recently WENO type strategies with higher order methods might also better work (see [45],[13]). We also illustrate here, as a known fact in the community, that the unlimited method is generally preferred; even if initially sharper gradients are discretized with oscillations, in the long run, the oscillations do not further develop and this is beneficial as the solution gets smoother. The plasma expansion test however presents a real better behavior for the limited schemes and thus shows a situation where the monotonicity limiting is useful.

One natural extension of this work is to add global maximum principle (in particular positivity), that we have here not added, permitting to measure the  $L^1$  norm conservation as indicator of the well behavior of the scheme (if positivity is ensured, in our conservative setting, the  $L^1$  norm automatically exactly satisfied). For this, we can follow the works [55, 49]. One other more demanding extension is to consider the non constant advection. For this, we can work as in [20] on the splitted conservative form, but this has the disadvantage of breaking the conservation of constant states at the level of the equations [36]. Another more common and popular method is to work on the  $2D$  unsplit advective form [46, 57, 36]. However, the conservation of mass is lost, and this is generally amplified when using the limiters [57, 36]. A dual way is to work with the more involved and technical  $2D$  unsplit conservative form [17], which has been later developed in a Semi-Lagrangian Discontinuous Galerkin context [6, 7]. Finally, extensions on non-uniform meshes can also be envisioned.

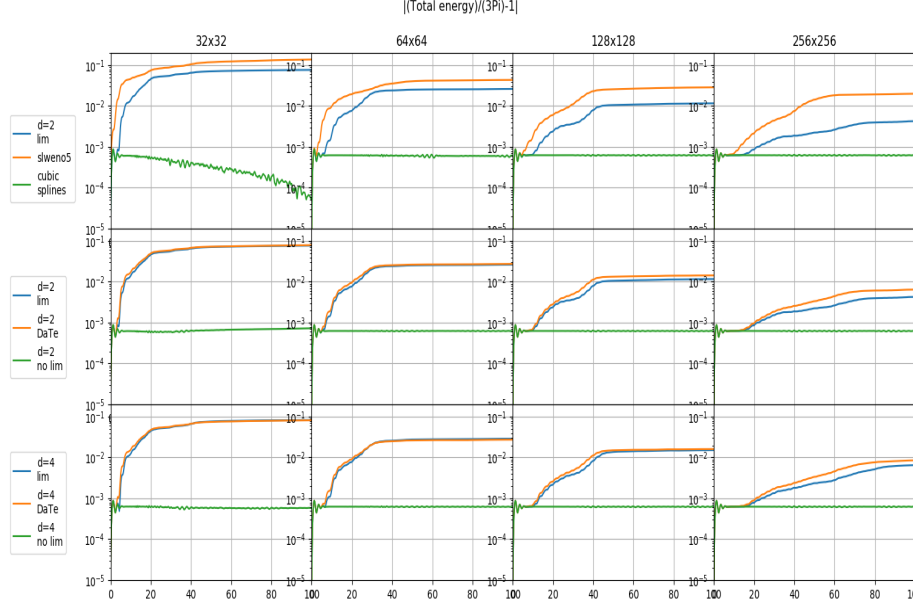


Figure 6: Time evolution of total energy error for non linear Landau damping test case

## Acknowledgements

This work has been supported by Heilongjiang Provincial Natural Science Foundation of China (LH2019A013).

MM acknowledges invitations as a scholar professor by Harbin Institute of Technology in 2018, 2019 and 2020. This work has been carried out within the framework of the EUROfusion consortium and has received funding from the Euratom research and training programme 2014-2018 and 2019-2020 under grant agreement No 633053. The views and opinions expressed herein do not necessarily reflect those of the European Commission. Centre de Calcul Intensif d'Aix-Marseille is acknowledged for granting access to its high performance computing resource.

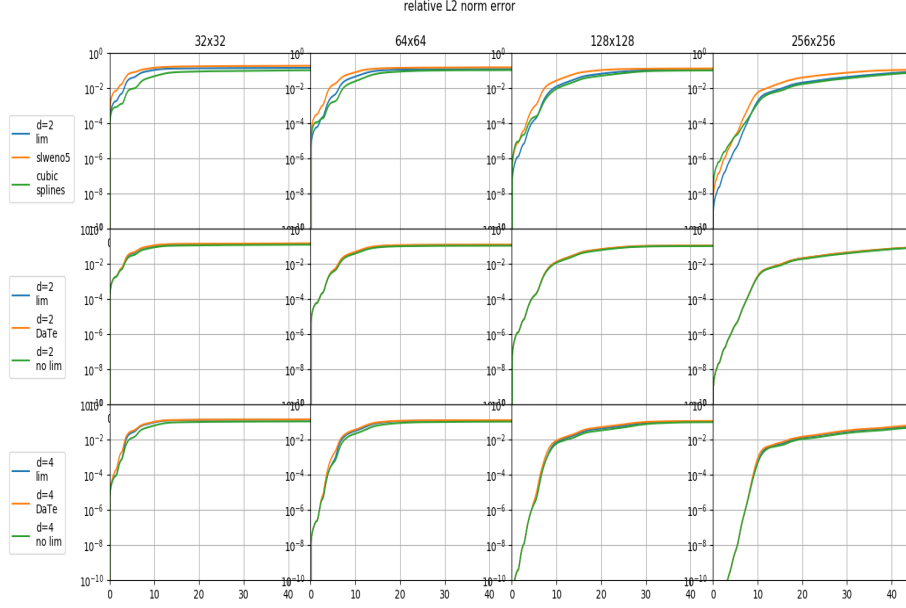


Figure 7: Time evolution of  $L^2$  norm error for two stream instability I test case

## Appendix A. Semi-Lagrangian scheme in flux formula

### Appendix A.1. Construction of left flux formula (symbolized by $L$ )

Suppose the foot of characteristic curve locates between  $x_{j^L-1}$  and  $x_{j^L}$ , and we write  $x_{j^L} - (x_i - v\Delta t) = \nu_L h$ , thus the numerical solution by interpolation of degree one is given by  $f_i^{n+1} = \nu_L f_{j^L-1}^n + (1 - \nu_L) f_{j^L}^n$ , and more generally of odd degree  $2d + 1$ , with  $d \in \mathbb{N}$ , is given by

$$f_i^{n+1} = \sum_{\ell=-d}^{d+1} \mathcal{L}_\ell(1 - \nu_L) f_{j^L-1+\ell}^n = \sum_{\ell=-d}^{d+1} \mathcal{L}_{1-\ell}(\nu_L) f_{j^L-1+\ell}^n = \sum_{\ell=-d}^{d+1} \mathcal{L}_\ell(\nu_L) f_{j^L-\ell}^n, \quad (\text{A.1})$$

where  $\mathcal{L}_\ell$  is Lagrange basis function defined by  $\mathcal{L}_\ell(\nu) = \prod_{k=-d, k \neq \ell}^{d+1} \frac{\nu-k}{\ell-k}$ .

The Semi-Lagrangian scheme (A.1) can be written under a conservative form. For  $d = 0$ , we have  $f_i^{n+1} = f_{j^L}^n - \nu_L (f_{j^L}^n - f_{j^L-1}^n)$ . We shall more generally write the Semi-Lagrangian scheme (A.1) with left flux formula, expressed as in (2)-(3). Thanks to this linear definition (3), the Semi-Lagrangian

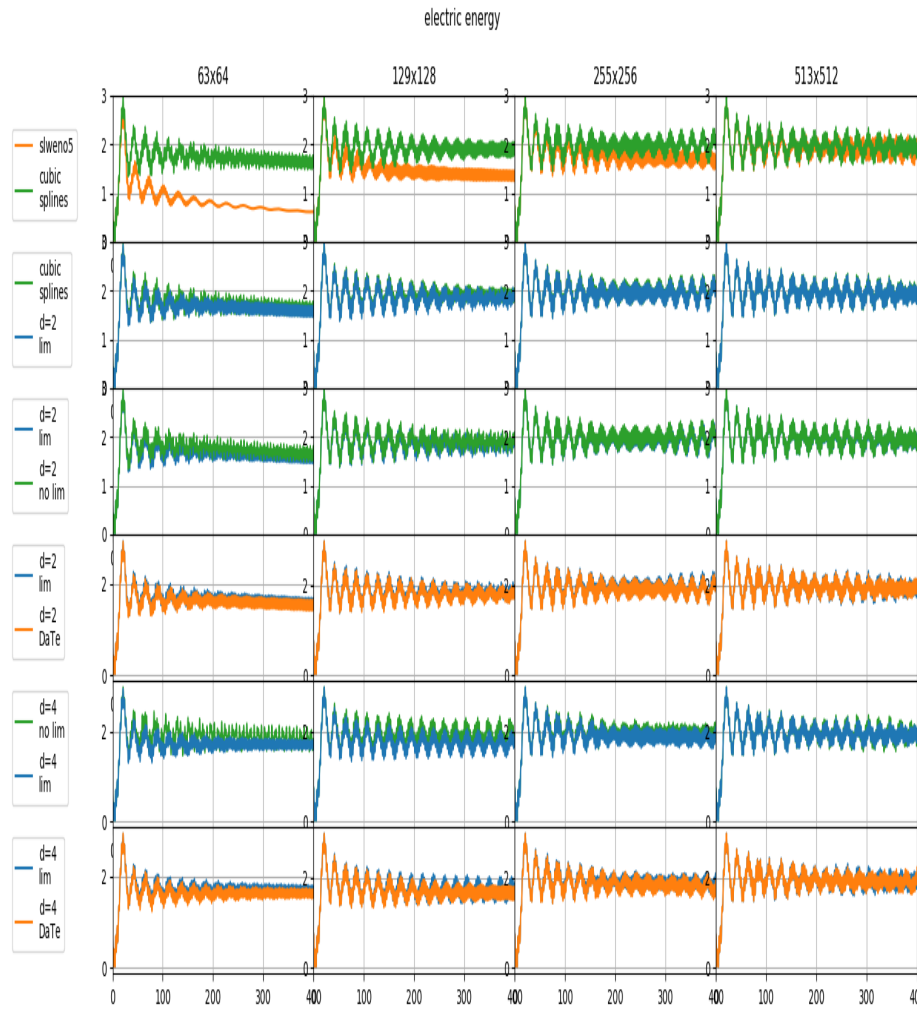


Figure 8: Time evolution of electric energy for bump on tail test case



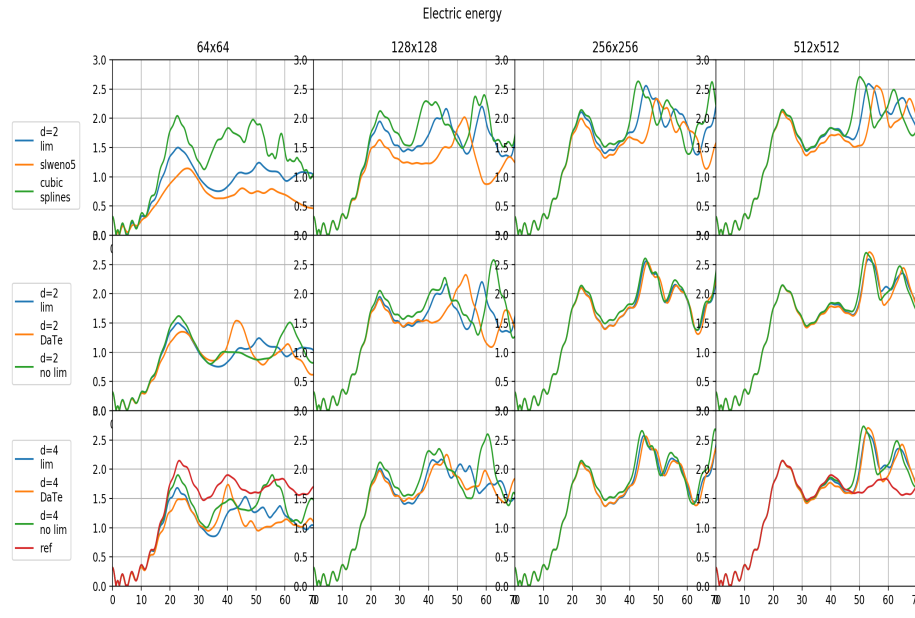


Figure 9: Time evolution of electric energy for short time ( $T = 70$ ) for two stream instability II test case

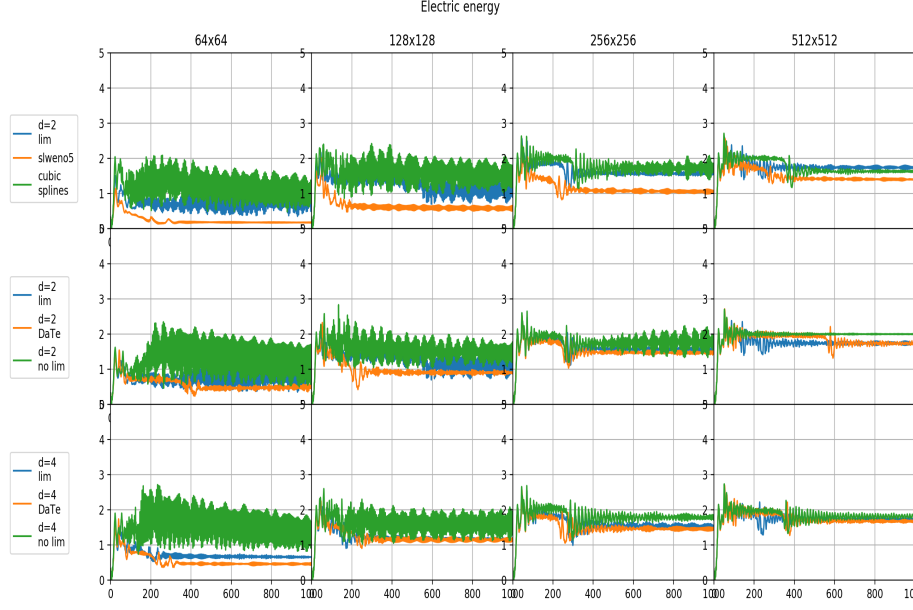


Figure 10: Time evolution of electric energy for two stream instability II test case

scheme can be recast as

$$\begin{aligned}
 f_i^{n+1} &= f_{j^L}^n - \nu_L \left( \sum_{\ell=-d}^d c_\ell^L(\nu_L) f_{j^L+\ell}^n - \sum_{\ell=-d}^d c_\ell^L(\nu_L) f_{j^L+\ell-1}^n \right) \\
 &= -\nu_L c_d^L(\nu_L) f_{j^L+d}^n + \nu_L c_{-d}^L(\nu_L) f_{j^L-d-1}^n - \nu_L \sum_{\ell=-d, \ell \neq 0}^{d-1} (c_\ell^L(\nu_L) - c_{\ell+1}^L(\nu_L)) f_{j^L+\ell}^n \\
 &\quad + (1 - \nu_L (c_0^L(\nu_L) - c_1^L(\nu_L))) f_{j^L}^n.
 \end{aligned}$$

Comparing the corresponding term in (A.1), we get

$$\begin{cases}
 \nu_L c_{-d}^L(\nu_L) &= \mathcal{L}_{d+1}(\nu_L), \\
 -\nu_L (c_\ell^L(\nu_L) - c_{\ell+1}^L(\nu_L)) &= \mathcal{L}_{-\ell}(\nu_L), \quad \ell = -d, \dots, -1, \\
 -\nu_L c_d^L(\nu_L) &= \mathcal{L}_{-d}(\nu_L), \\
 -\nu_L (c_\ell^L(\nu_L) - c_{\ell+1}^L(\nu_L)) &= \mathcal{L}_{-\ell}(\nu_L), \quad \ell = 1, \dots, d-1, \\
 1 - \nu_L (c_0^L(\nu_L) - c_1^L(\nu_L)) &= \mathcal{L}_0(\nu_L).
 \end{cases}$$

Notice that the above system holds for any  $\nu_L \in (0, 1)$ . Solving the linear system

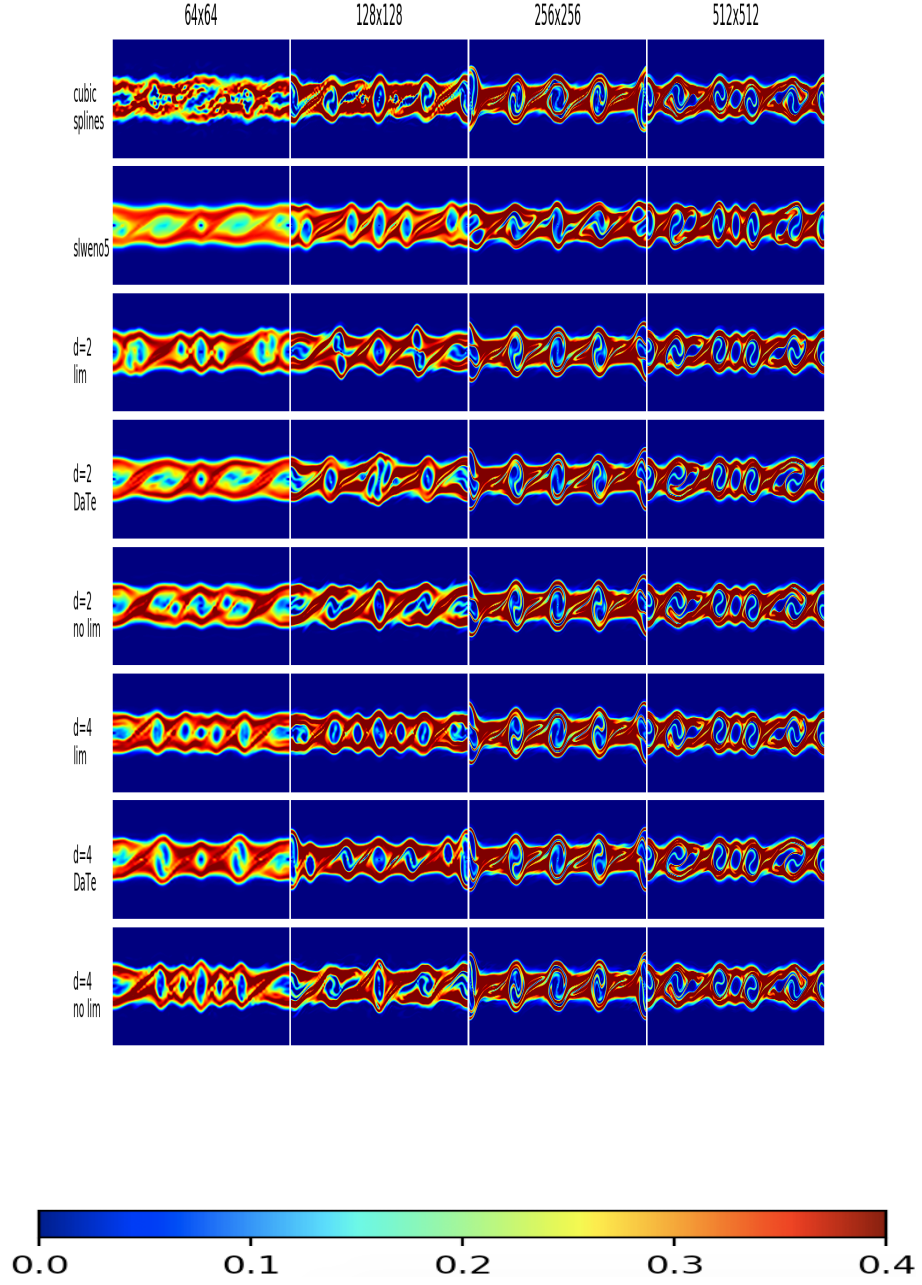


Figure 11: Two stream instability II at time  $T = 70$

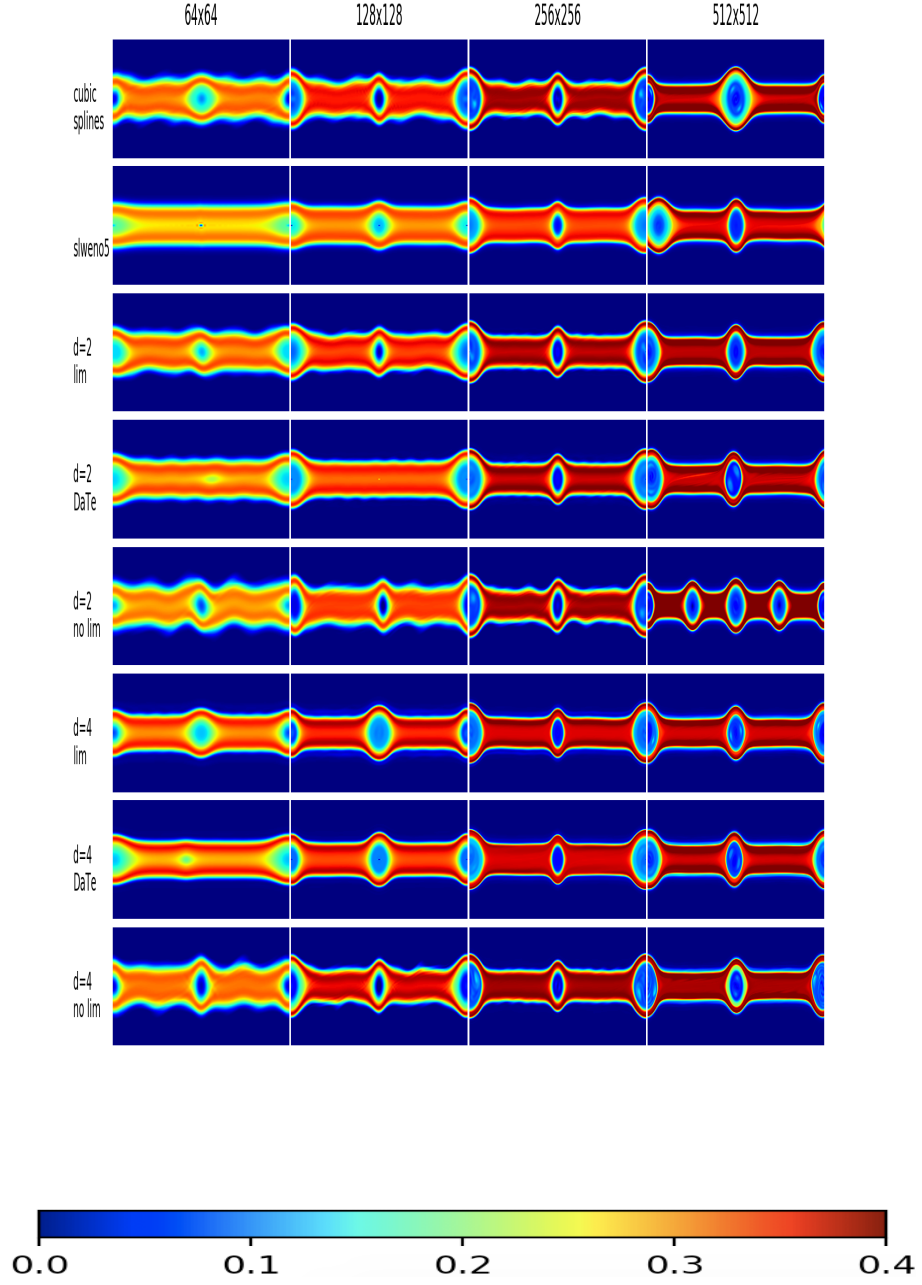


Figure 12: Two stream instability II at time  $T = 1000$

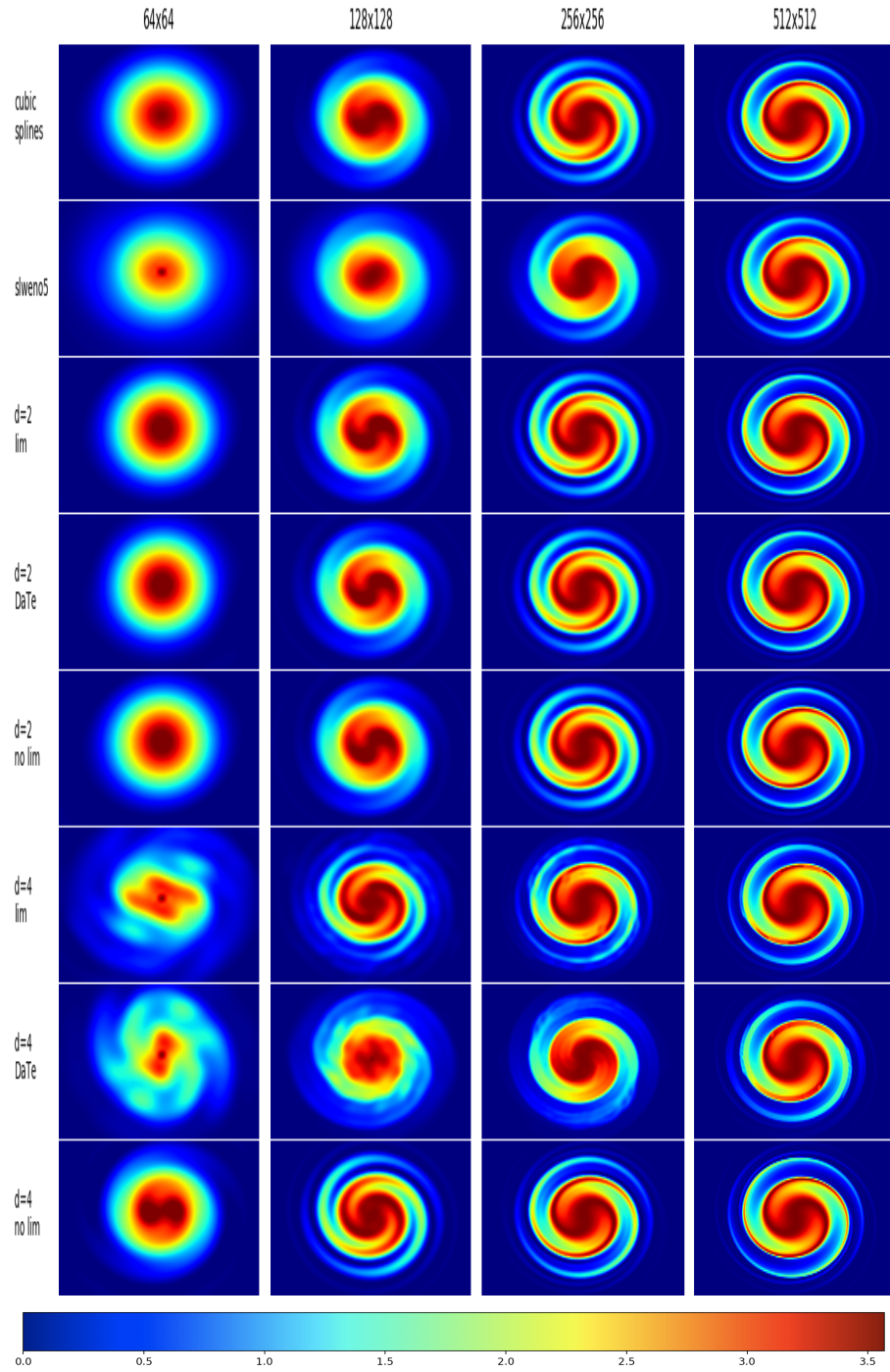


Figure 13: Beam test case at time  $T = 1000$

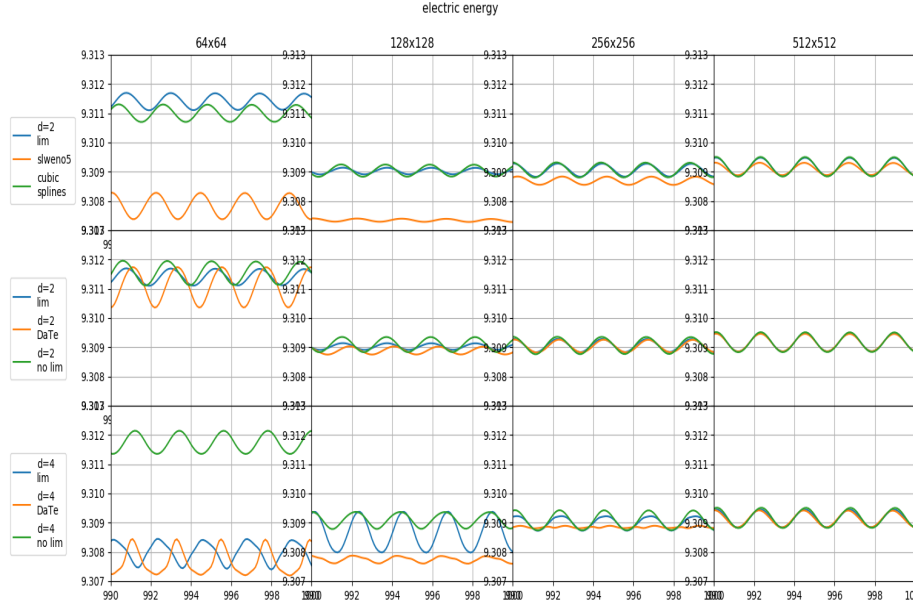


Figure 14: Time evolution of electric energy for beam test case

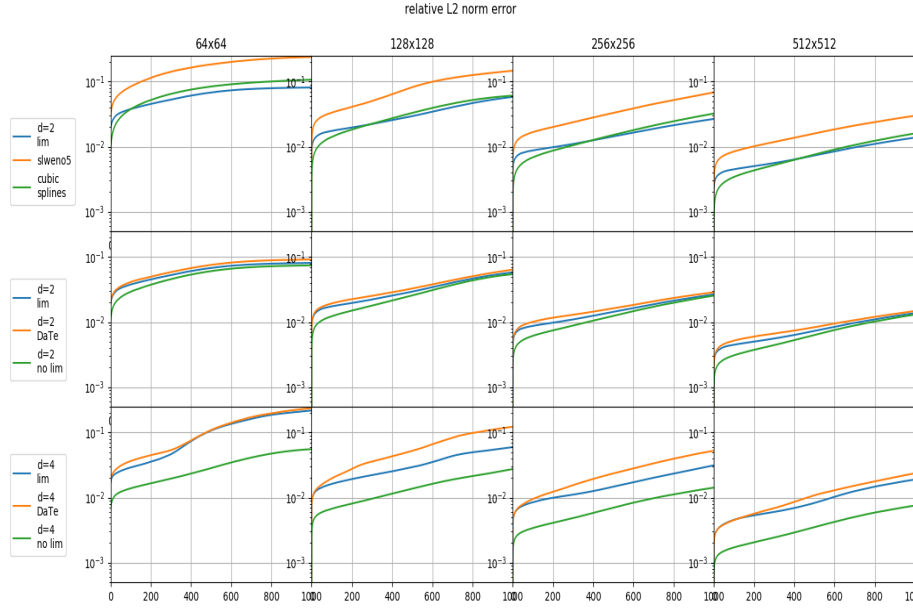


Figure 15: Time evolution of relative  $L^2$  norm error for beam test case

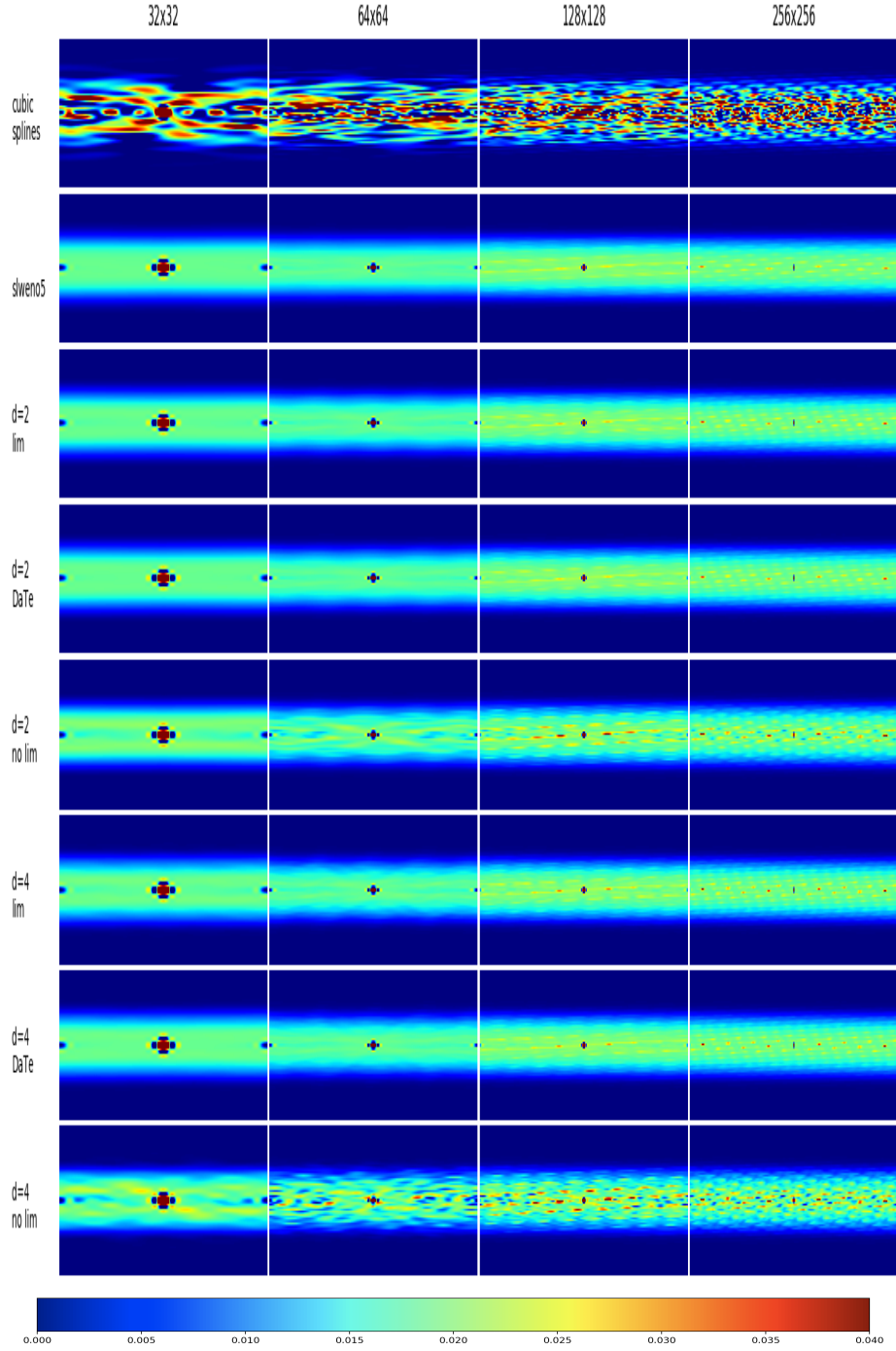


Figure 16: Plasma expansion at time  $T = 400$  on  $32 \times 32$  to  $256 \times 256$  grids



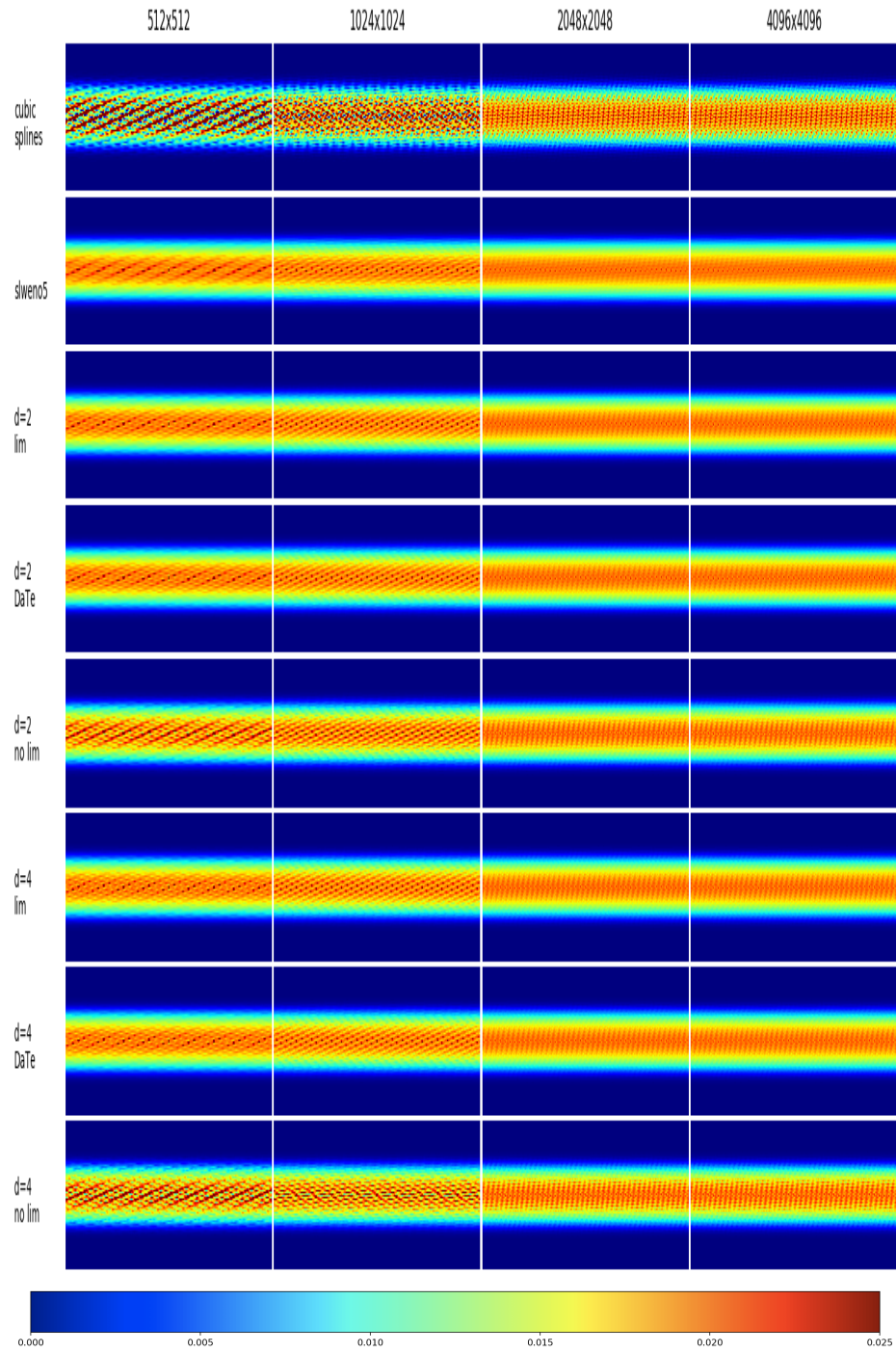


Figure 17: Plasma expansion at time  $T = 400$  on  $512 \times 512$  to  $4096 \times 4096$  grids



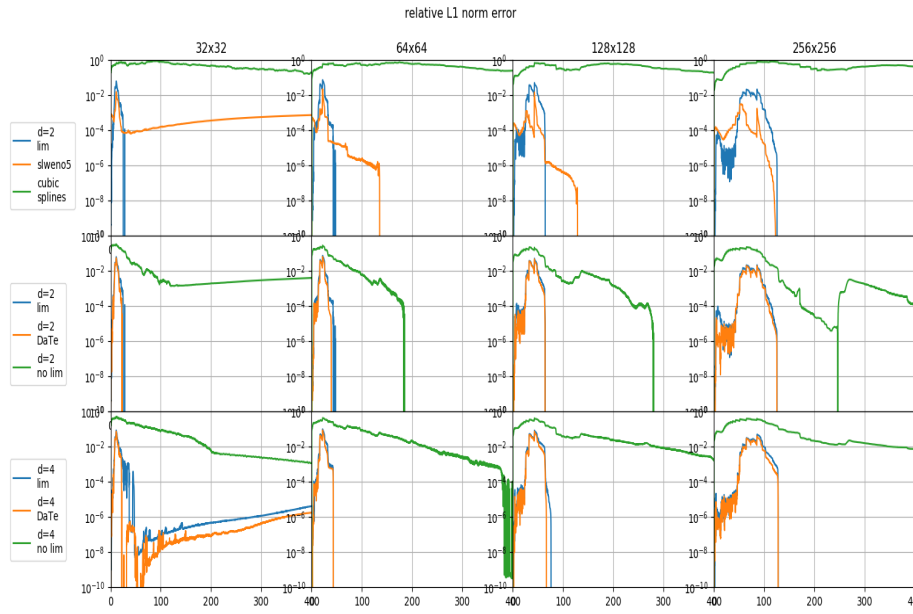


Figure 18: Time evolution of relative  $L^1$  norm error for plasma expansion test case on  $32 \times 32$  to  $256 \times 256$  grids. When we do not see the line of the plot, it means that the error is below  $10^{-10}$ .

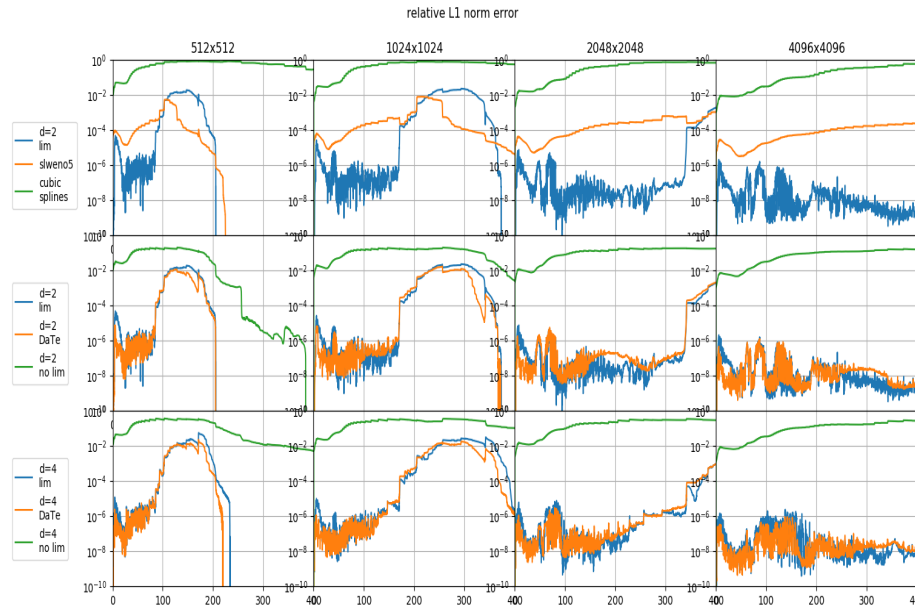


Figure 19: Time evolution of relative  $L^1$  norm error for plasma expansion test case on  $512 \times 512$  to  $4096 \times 4096$  grids

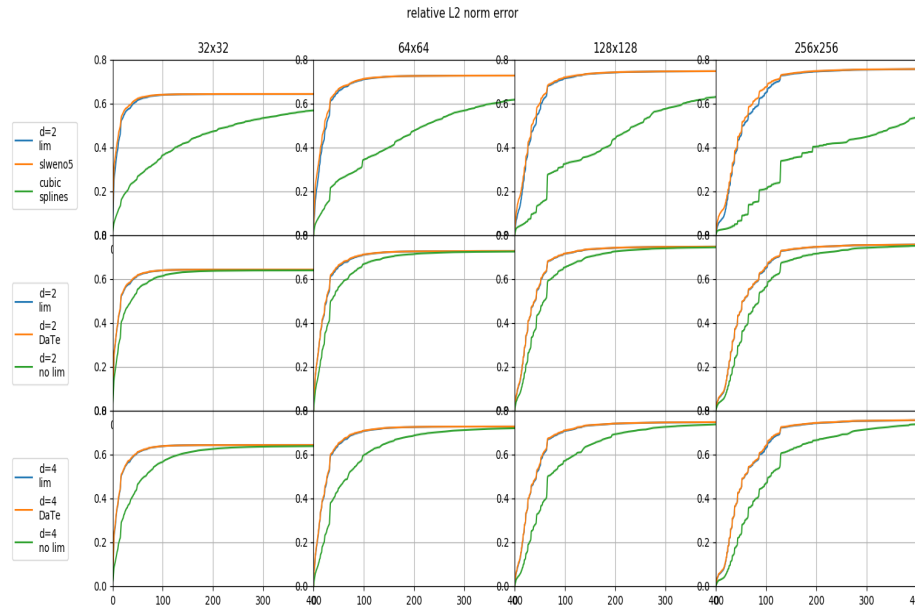


Figure 20: Time evolution of relative  $L^2$  norm error for plasma expansion test case on  $32 \times 32$  to  $256 \times 256$  grids

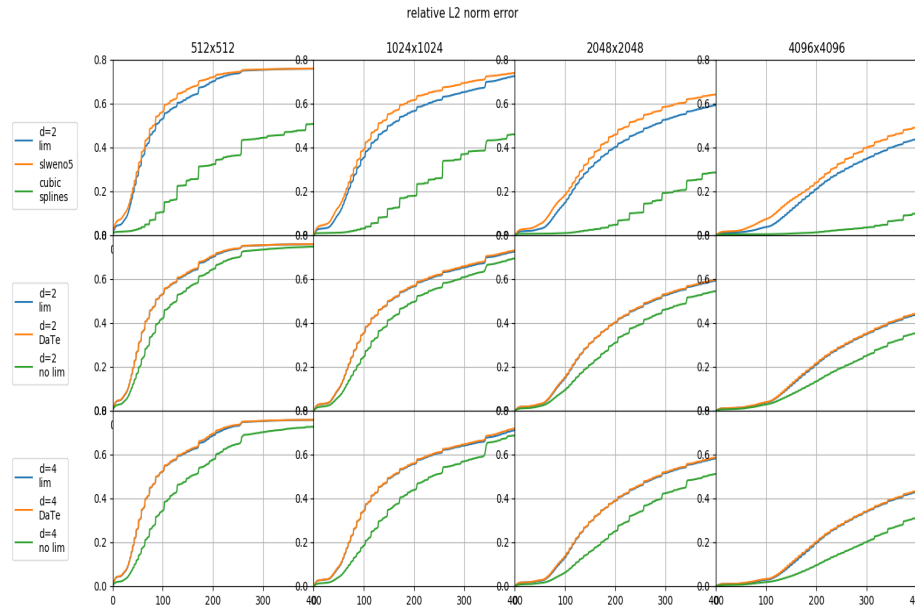


Figure 21: Time evolution of relative  $L^2$  norm error for plasma expansion test case on  $512 \times 512$  to  $4096 \times 4096$  grids

yields, for any  $\nu \neq 0$ , the functional coefficients

$$c_\ell^L(\nu) = \begin{cases} \frac{1}{\nu} \sum_{k=-d-1}^{\ell-1} \mathcal{L}_{-k}(\nu), & \ell = -d, \dots, 0, \\ -\frac{1}{\nu} \sum_{k=\ell}^d \mathcal{L}_{-k}(\nu), & \ell = 1, \dots, d, \end{cases} \quad (\text{A.2})$$

since  $\nu c_{-d}^L(\nu) + \sum_{\ell=-d}^{d-1} -\nu(c_\ell^L(\nu) - c_{\ell+1}^L(\nu)) - \nu c_d^L(\nu) + \sum_{\ell=1}^{d-1} -\nu(c_\ell^L(\nu) - c_{\ell+1}^L(\nu)) + 1 - \nu(c_0^L(\nu) - c_1^L(\nu)) = 1$ , and  $\mathcal{L}_{d+1}(\nu) + \sum_{\ell=-d}^{d-1} \mathcal{L}_{-\ell}(\nu) + \mathcal{L}_{-d}(\nu) + \sum_{\ell=1}^{d-1} \mathcal{L}_{-\ell}(\nu) + \mathcal{L}_0(\nu) = \sum_{\ell=-d}^{d+1} \mathcal{L}_\ell(\nu) = 1$ . For  $\nu = 0$ , the  $c_\ell^L(\nu)$  can be arbitrary from the system; we define them by taking the limit of (A.2), as  $\nu \rightarrow 0$ , which is well defined, since 0 is a root of  $\mathcal{L}_k$ , for  $k \neq 0$ . For implementing issues, we can use the following formulae which have no evaluation problem:

$$c_\ell^L(\nu) = \begin{cases} \sum_{k=-d-1}^{\ell-1} \tilde{\mathcal{L}}_{-k}(\nu), & \ell = -d, \dots, 0, \\ -\sum_{k=\ell}^d \tilde{\mathcal{L}}_{-k}(\nu), & \ell = 1, \dots, d, \end{cases} \quad (\text{A.3})$$

with

$$\tilde{\mathcal{L}}_\ell(\nu) = \frac{1}{\ell} \prod_{k=-d, k \notin \{\ell, 0\}}^{d+1} \frac{\nu - k}{\ell - k}, \quad (\text{A.4})$$

for  $\ell = -d, \dots, d+1$ ,  $\ell \neq 0$ .

#### Appendix A.2. Construction of right flux formula (symbolized by $R$ )

In order now to use the right flux formula, we suppose that the foot of characteristic curve locates between  $x_{jR}$  and  $x_{jR+1}$ , and we can write  $(x_i - v\Delta t) - x_{jR} = \nu_R h$ , thus the numerical solution by interpolation of degree one is now given by  $f_i^{n+1} = (1 - \nu_R)f_{jR}^n + \nu_R f_{jR+1}^n$ , and more generally of odd degree  $2d+1$ , with  $d \in \mathbb{N}$ , is given by

$$f_i^{n+1} = \sum_{\ell=-d}^{d+1} \mathcal{L}_\ell(\nu_R) f_{jR+\ell}^n. \quad (\text{A.5})$$

We get the same solution as before, but we now will make a link with the conservative form using the right flux formula instead of the left flux formula. We can write for  $d = 0$ ,  $f_i^{n+1} = f_{jR}^n + \nu_R(f_{jR+1}^n - f_{jR}^n)$ , and more generally write the Semi-Lagrangian scheme (A.5) with right flux formula, expressed as in (4)-(5).

Now, the Semi-Lagrangian scheme can be recast as

$$\begin{aligned}
f_i^{n+1} &= f_{j^R}^n + \nu_R \left( \sum_{\ell=-d}^d c_\ell^R(\nu_R) f_{j^R+\ell+1}^n - \sum_{\ell=-d}^d c_\ell^R(\nu_R) f_{j^R+\ell}^n \right) \\
&= \nu_R c_d^R(\nu_R) f_{j^R+d+1}^n - \nu_R c_{-d}^R(\nu_R) f_{j^R-d}^n \\
&\quad + \nu_R \sum_{\ell=-d+1, \ell \neq 0}^d (c_{\ell-1}^R(\nu_R) - c_\ell^R(\nu_R)) f_{j^R+\ell}^n + (1 + \nu_R(c_{-1}^R(\nu_R) - c_0^R(\nu_R))) f_{j^R}^n.
\end{aligned}$$

Comparing the corresponding term in (A.5), we get

$$\begin{cases} -\nu_R c_{-d}^R(\nu_R) &= \mathcal{L}_{-d}(\nu_R), \\ \nu_R(c_{\ell-1}^R(\nu_R) - c_\ell^R(\nu_R)) &= \mathcal{L}_\ell(\nu_R), \quad \ell = -d+1, \dots, -1, \\ \nu_R c_d^R(\nu_R) &= \mathcal{L}_{d+1}(\nu_R), \\ \nu_R(c_{\ell-1}^R(\nu_R) - c_\ell^R(\nu_R)) &= \mathcal{L}_\ell(\nu_R), \quad \ell = 1, \dots, d, \\ 1 + \nu_R(c_{-1}^R(\nu_R) - c_0^R(\nu_R)) &= \mathcal{L}_0(\nu_R). \end{cases}$$

We find similarly the functional coefficients as follows

$$c_\ell^R(\nu) = \begin{cases} -\sum_{k=-d}^{\ell} \tilde{\mathcal{L}}_k(\nu), & \ell = -d, \dots, -1, \\ \sum_{k=\ell+1}^{d+1} \tilde{\mathcal{L}}_k(\nu), & \ell = 0, \dots, d, \end{cases} \quad (\text{A.6})$$

where  $\tilde{\mathcal{L}}$  is the same as (A.4).

The functional coefficients  $c_\ell^R(\nu)$  is also related to  $c_\ell^L(\nu)$ . Indeed, on the one hand, we have, for  $\ell = -d, \dots, -1$ ,  $c_\ell^R(\nu) = -\sum_{k=-d}^{\ell} \tilde{\mathcal{L}}_k(\nu)$  and  $c_{-\ell}^L(\nu) = -\sum_{k=-\ell}^d \tilde{\mathcal{L}}_{-k}(\nu) = -\sum_{k=-d}^{\ell} \tilde{\mathcal{L}}_k(\nu) = c_\ell^R(\nu)$ . On the other hand, we also have for  $\ell = 0, \dots, d$ ,  $c_\ell^R(\nu) = \sum_{k=\ell+1}^{d+1} \tilde{\mathcal{L}}_k(\nu)$  and  $c_{-\ell}^L(\nu) = \sum_{k=-d-1}^{-\ell-1} \tilde{\mathcal{L}}_{-k}(\nu) = \sum_{k=\ell+1}^{d+1} \tilde{\mathcal{L}}_k(\nu) = c_\ell^R(\nu)$ . We conclude that  $c_\ell^R(\nu) = c_{-\ell}^L(\nu)$ , for  $\ell = -d, \dots, d$ .

## Appendix B. Explanations and proofs related to Section 3

### Appendix B.1. Relation between the limiters for left and right flux formulae

Let us show a relation between MP constraints for left flux formula and right flux formula in this part. As shown is Fig 1, we have  $j^L - \nu_L = j^R + \nu_R = j^L - 1 + 1 - \nu_L$ . So, we get  $j^R = j^L - 1$  and  $\nu_R = 1 - \nu_L$ . Thus for the right flux formula (4), we have

$$f_i^{n+1} = f_{j^L-1}^n + (1-\nu_L) \left( f_{j^R+1/2}^R - f_{j^R-1/2}^R \right) = f_{j^L}^n + f_{j^L-1}^n - f_{j^L}^n + (1-\nu_L) \left( f_{j^R+1/2}^R - f_{j^R-1/2}^R \right),$$

that is

$$f_i^{n+1} = f_{j^L}^n - (f_{j^L}^n - (1 - \nu_L)f_{j^R+1/2}^R) + f_{j^L-1}^n - (1 - \nu_L)f_{j^R-1/2}^R.$$

From the unicity of the formula, we have

$$\nu_L f_{j^L+1/2}^L = f_{j^L}^n - (1 - \nu_L)f_{j^R+1/2}^R = f_{j^L}^n - \nu_R f_{j^R+1/2}^R.$$

From (11), we have

$$\nu_L M_{j^L-1/2}^{(2,L)} + \left( f_{j^L}^n - M_{j^L-1/2}^{(2,L)} \right) \leq \nu_L f_{j^L+1/2}^L \leq \nu_L m_{j^L-1/2}^{(2,L)} + \left( f_{j^L}^n - m_{j^L-1/2}^{(2,L)} \right),$$

and so

$$(1 - \nu_L)m_{j^L-1/2}^{(2,L)} \leq f_{j^L}^n - \nu_L f_{j^L+1/2}^L \leq (1 - \nu_L)M_{j^L-1/2}^{(2,L)},$$

which leads to

$$m_{j^L-1/2}^{(2,L)} \leq f_{j^R+1/2}^R \leq M_{j^L-1/2}^{(2,L)}.$$

We have also from  $m_{j^L+1/2}^{(1,L)} \leq f_{j^L+1/2}^L \leq M_{j^L+1/2}^{(1,L)}$ , that

$$\nu_L m_{j^L+1/2}^{(1,L)} \leq \nu_L f_{j^L+1/2}^L \leq \nu_L M_{j^L+1/2}^{(1,L)},$$

leading to

$$f_{j^L}^n - \nu_L M_{j^L+1/2}^{(1,L)} \leq f_{j^L}^n - \nu_L f_{j^L+1/2}^L \leq f_{j^L}^n - \nu_L m_{j^L+1/2}^{(1,L)},$$

that is

$$f_{j^L}^n - (1 - \nu_R)M_{j^L+1/2}^{(1,L)} \leq \nu_R f_{j^R+1/2}^R \leq f_{j^L}^n - (1 - \nu_R)m_{j^L+1/2}^{(1,L)},$$

which leads to

$$M_{j^L+1/2}^{(1,L)} + \frac{1}{\nu_R}(f_{j^L}^n - M_{j^L+1/2}^{(1,L)}) \leq f_{j^R+1/2}^R \leq m_{j^L+1/2}^{(1,L)} + \frac{1}{\nu_R}(f_{j^L}^n - m_{j^L+1/2}^{(1,L)}).$$

We deduce that when we have  $m_{j^R+1/2}^{(1,R)} = m_{j^L-1/2}^{(2,L)}$ ,  $M_{j^R+1/2}^{(1,R)} = M_{j^L-1/2}^{(2,L)}$  and  $m_{j^R+1/2}^{(2,R)} = m_{j^L-1/2}^{(1,L)}$ ,  $M_{j^R+1/2}^{(2,R)} = M_{j^L-1/2}^{(1,L)}$ , then the limiter for right flux formula is equivalent to the one for left flux formula.

## Appendix B.2. Proofs related to Section 3.3

Now let us prove Proposition 3.3.

*Proof.* We only consider the case with increasing data, the proof for decreasing data follows the same arguments.

To prove the constraint (21) is MP, it is just necessary to verify

$$f_{\min,j^L} = f_{j^L}^n \text{ and } f_{\max,j^L} = \min \left( f_{j^L+1}^n, \Phi_{\nu_L,j^L}(f_{j^L-1}^n) \right).$$

So the constraint (21) becomes

$$f_{j^L}^n \leq f_{j^L+1/2}^L \leq \min \left( f_{j^L+1}^n, \Phi_{\nu_L, j^L}(f_{j^L-1}^n) \right). \quad (\text{B.1})$$

Moreover, since the constraint (21) is union of (18) and (20), we just need to prove both (18) and (20) are equivalent to (B.1).

Let us first look at (18). We obviously have

$$\begin{aligned} \max \left( m_{j^L+1/2}^{Um}, \Phi_{\nu_L, j^L}(\max(M_{j^L-1/2}^{Um}, f_{j^L-}^{LC})) \right) &= \max \left( f_{j^L}^L, \Phi_{\nu_L, j^L}(\max(f_{j^L}^L, f_{j^L-}^{LC})) \right) \\ &= \max \left( f_{j^L}^L, \min(f_{j^L}^L, \Phi_{\nu_L, j^L}(f_{j^L-}^{LC})) \right) = f_{j^L}^n. \end{aligned}$$

On the other hand,

$$\begin{aligned} \min \left( M_{j^L+1/2}^{Um}, \Phi_{\nu_L, j^L}(\min(m_{j^L-1/2}^{Um}, f_{j^L-}^{LC})) \right) &= \min \left( f_{j^L+1}^L, \Phi_{\nu_L, j^L}(\min(f_{j^L-1}^L, f_{j^L-}^{LC})) \right) \\ &= \min \left( f_{j^L+1}^n, \max(\Phi_{\nu_L, j^L}(f_{j^L-1}^n), \Phi_{\nu_L, j^L}(f_{j^L-}^{LC})) \right). \end{aligned}$$

We want to prove that, if the data are increasing, then  $\Phi_{\nu_L, j^L}(f_{j^L-1}^n) \geq \Phi_{\nu_L, j^L}(f_{j^L-}^{LC})$ , that is  $f_{j^L-1}^n \leq f_{j^L-}^{LC}$ , so that

$$d_{j^L-1/2} = \minmod(d_{j^L}, d_{j^L-1}) \leq f_{j^L}^n - f_{j^L-1}^n. \quad (\text{B.2})$$

If  $d_{j^L} d_{j^L-1} \leq 0$  or ( $d_{j^L} < 0$  and  $d_{j^L-1} < 0$ ), this is true, since  $f_{j^L}^n \geq f_{j^L-1}^n$ . Otherwise  $d_{j^L} > 0$  and  $d_{j^L-1} > 0$ , that is  $f_{j^L-1}^n - 2f_{j^L}^n + f_{j^L+1}^n > 0$  and  $f_{j^L-2}^n - 2f_{j^L-1}^n + f_{j^L}^n > 0$ . Thus (B.2) is true provided

$$f_{j^L-2}^n - 2f_{j^L-1}^n + f_{j^L}^n \leq f_{j^L}^n - f_{j^L-1}^n.$$

This is true thanks to

$$f_{j^L-2}^n - f_{j^L-1}^n \leq 0.$$

Then let us look at (20). We similarly have

$$\begin{aligned} \max \left( \min(m_{j^L+1/2}^{Um}, f_{j^L+}^{LC}), \Phi_{\nu_L, j^L}(M_{j^L-1/2}^{Um}) \right) &= \max \left( \min(f_{j^L}^n, f_{j^L+}^{LC}), \Phi_{\nu_L, j^L}(f_{j^L}^n) \right) \\ &= \max \left( \min(f_{j^L}^n, f_{j^L+}^{LC}), f_{j^L}^n \right) = f_{j^L}^n. \end{aligned}$$

On the other hand,

$$\min \left( \max(M_{j^L+1/2}^{Um}, f_{j^L+}^{LC}), \Phi_{\nu_L, j^L}(m_{j^L-1/2}^{Um}) \right) = \min \left( \max(f_{j^L+1}^n, f_{j^L+}^{LC}), \Phi_{\nu_L, j^L}(f_{j^L-1}^n) \right).$$

We want to prove that, if the data are increasing, then  $f_{j^L+1}^n \geq f_{j^L+}^{LC}$ , so that

$$-d_{j^L+1/2} = -\minmod(d_{j^L}, d_{j^L+1}) \leq f_{j^L+1}^n - f_{j^L}^n. \quad (\text{B.3})$$



If  $d_{j^L}d_{j^L+1} \geq 0$  or ( $d_{j^L} > 0$  and  $d_{j^L+1} > 0$ ), this is true, since  $f_{j^L+1}^n \geq f_{j^L}^n$ . Otherwise  $d_{j^L} < 0$  and  $d_{j^L+1} < 0$ , that is  $f_{j^L-1}^n - 2f_{j^L}^n + f_{j^L+1}^n < 0$  and  $f_{j^L}^n - 2f_{j^L+1}^n + f_{j^L+2}^n < 0$ . Thus (B.3) is true provided

$$-f_{j^L}^n + 2f_{j^L+1}^n - f_{j^L+2}^n \leq f_{j^L+1}^n - f_{j^L}^n.$$

This is true thanks to

$$f_{j^L+2}^n - f_{j^L+1}^n \geq 0.$$

Therefore, we have proven the results.  $\square$

## References

- [1] A. Back and E. Sonnendrucker. Finite element hodge for spline discrete differential forms. application to the Vlasov-Poisson system. Applied Numerical Mathematics, 79:124–136, 2014.
- [2] J. W. Banks, A. G. Odu, R. L. Berger, T. Chapman, W. Arrighi, and S. Brunner. High-order accurate conservative finite difference methods for Vlasov equations in 2d+2v. SIAM Journal on Scientific Computing, 41(5), 2019.
- [3] J. Bernier, F. Casas, and N. Crouseilles. Splitting methods for rotations: application to Vlasov equations. SIAM Journal on Scientific Computing, 42(2):A666–A697, 2020.
- [4] N. Besse and M. Mehrenberger. Convergence of classes of high-order semi-Lagrangian schemes for the Vlasov-Poisson system. Mathematics of Computation, 77(261):93–123, 2008.
- [5] X. Cai, W. Guo, and J. Qiu. A high order semi-lagrangian discontinuous Galerkin method for Vlasov-Poisson simulations without operator splitting. Journal of Computational Physics, 354:529–551, 2018.
- [6] X. Cai, W. Guo, and J.-M. Qiu. A high order semi-lagrangian discontinuous Galerkin method for the two-dimensional incompressible Euler equations and the guiding center Vlasov model without operator splitting. Journal of Scientific Computing, 79(2):1111–1134, 2019.
- [7] X. Cai, J.-M. Qiu, and Y. Yang. An Eulerian-Lagrangian discontinuous Galerkin method for transport problems and its application to nonlinear dynamics. arXiv preprint arXiv:2002.02930, 2020.
- [8] F. Casas, N. Crouseilles, E. Faou, and M. Mehrenberger. High-order hamiltonian splitting for the Vlasov–Poisson equations. Numerische Mathematik, 135(3):769–801, 2017.
- [9] F. Charles, B. Despres, and M. Mehrenberger. Enhanced convergence estimates for semi-Lagrangian schemes application to the Vlasov-Poisson equation. SIAM Journal on Numerical Analysis, 51(2):840–863, 2013.

- [10] C.-Z. Cheng and G. Knorr. The integration of the Vlasov equation in configuration space. Journal of Computational Physics, 22(3):330–351, 1976.
- [11] Y. Cheng, A. J. Christlieb, and X. Zhong. Energy-conserving discontinuous Galerkin methods for the Vlasov-Maxwell system. Journal of Computational Physics, 256(1):630–655, 2014.
- [12] Y. Cheng, I. M. Gamba, and P. J. Morrison. Study of conservation and recurrence of Runge-Kutta discontinuous Galerkin schemes for vlasov-poisson systems. Journal of Scientific Computing, 56(2):319–349, 2013.
- [13] S. Y. Cho, S. Boscarino, G. Russo, and S.-B. Yun. Conservative semi-lagrangian schemes for kinetic equations part ii: Applications. Journal of Computational Physics, 436:110281, 2021.
- [14] G. H. Cottet and P. A. Raviart. On particle-in-cell methods for the Vlasov-Poisson equations. Transport Theory and Statistical Physics, 15:1–31, 1986.
- [15] D. Coulette and G. Manfredi. An eulerian Vlasov code for plasma-wall interactions. In Journal of Physics: Conference Series, volume 561, page 012005. IOP Publishing, 2014.
- [16] N. Crouseilles, L. Einkemmer, and J. Massot. Exponential methods for solving hyperbolic problems with application to collisionless kinetic equations. Journal of Computational Physics, 420:109688, 2020.
- [17] N. Crouseilles, P. Glanc, S. A. Hirstoaga, E. Madaule, M. Mehrenberger, and J. Pétri. A new fully two-dimensional conservative semi-Lagrangian method: applications on polar grids, from diocotron instability to ITG turbulence. The European Physical Journal D, 68(9):252, 2014.
- [18] N. Crouseilles, P. Glanc, M. Mehrenberger, and C. Steiner. Finite volume schemes for Vlasov. CEMRACS 2011: Multiscale coupling of complex models in scientific computing. In ESAIM Proc., volume 38, pages 275–297, 2012.
- [19] N. Crouseilles, M. Lemou, F. Mehats, and X. Zhao. Uniformly accurate particle-in-cell method for the long time solution of the two-dimensional Vlasov-Poisson equation with uniform strong magnetic field. Journal of Computational Physics, 346:172–190, 2017.
- [20] N. Crouseilles, M. Mehrenberger, and E. Sonnendrücker. Conservative semi-lagrangian schemes for Vlasov equations. Journal of Computational Physics, 229(6):1927–1953, 2010.
- [21] V. Daru and C. Tenaud. High order one-step monotonicity-preserving schemes for unsteady compressible flow calculations. Journal of Computational Physics, 193(2):563–594, 2004.

- [22] B. A. De Dios, J. A. Carrillo, and C. Shu. Discontinuous Galerkin methods for the multi-dimensional Vlasov–Poisson problem. Mathematical Models and Methods in Applied Sciences, 22(12):1250042, 2012.
- [23] P. Degond, F. Deluzet, and D. Doyen. Asymptotic-preserving particle-in-cell methods for the Vlasov-Maxwell system in the quasi-neutral limit. Journal of Computational Physics, 330:467 – 492, 2017.
- [24] P. Degond, F. Deluzet, L. Navoret, A. Sun, and M. Vignal. Asymptotic-preserving particle-in-cell method for the Vlasov-Poisson system near quasineutrality. Journal of Computational Physics, 229(16):5630–5652, 2010.
- [25] B. Després. Finite volume transport schemes. Numerische Mathematik, 108(4):529–556, 2008.
- [26] B. Després. Uniform asymptotic stability of Strang’s explicit compact schemes for linear advection. SIAM journal on numerical analysis, 47(5):3956–3976, 2009.
- [27] B. Després. Polynomials with bounds and numerical approximation. Numerical Algorithms, 76(3):829–859, 2017.
- [28] B. Després and F. Lagoutière. Contact discontinuity capturing schemes for linear advection and compressible gas dynamics. Journal of Scientific Computing, 16(4):479–524, 2002.
- [29] L. Einkemmer. A study on conserving invariants of the vlasov equation in semi-lagrangian computer simulations. Journal of Plasma Physics, 83(2), 2017.
- [30] L. Einkemmer. A performance comparison of semi-lagrangian discontinuous Galerkin and spline based Vlasov solvers in four dimensions. Journal of Computational Physics, 376:937–951, 2019.
- [31] L. Fatone, D. Funaro, and G. Manzini. A semi-lagrangian spectral method for the Vlasov-Poisson system based on Fourier, Legendre and Hermite polynomials. arXiv: Numerical Analysis, 2018.
- [32] F. Filbet. Convergence d’un schéma de type volumes finis pour la résolution numérique du système de Vlasov–Poisson en dimension un. Comptes Rendus de l’Académie des Sciences - Series I - Mathematics, 330(11):979 – 984, 2000.
- [33] F. Filbet. Convergence of a finite volume scheme for the Vlasov-Poisson system. SIAM Journal on Numerical Analysis, 39(4):1146–1169, 2001.
- [34] F. Filbet and L. M. Rodrigues. Asymptotically stable particle-in-cell methods for the Vlasov-Poisson system with a strong external magnetic field. SIAM Journal on Numerical Analysis, 54(2):1120–1146, 2016.

- [35] F. Filbet, E. Sonnendrücker, and P. Bertrand. Conservative numerical schemes for the Vlasov equation. Journal of Computational Physics, 172(1):166–187, 2001.
- [36] A. Hamiaz, M. Mehrenberger, H. Sellama, and E. Sonnendrücker. The semi-lagrangian method on curvilinear grids. Communications in Applied and Industrial Mathematics, 7(3):99–137, 2016.
- [37] R. E. Heath, I. M. Gamba, P. J. Morrison, and C. Michler. A discontinuous Galerkin method for the Vlasov-Poisson system. Journal of Computational Physics, 231(4):1140–1174, 2012.
- [38] G.-S. Jiang and C.-W. Shu. Efficient implementation of weighted ENO schemes. Journal of computational physics, 126(1):202–228, 1996.
- [39] E. Madaule, M. Restelli, and E. Sonnendrücker. Energy conserving discontinuous Galerkin spectral element method for the Vlasov-Poisson system. Journal of Computational Physics, 279:261–288, 2014.
- [40] E. Malkov, S. Poleshkin, A. Shershnev, and A. Kudryavtsev. Numerical solution of the Vlasov-Poisson equations using a semi-lagrangian WENO scheme implemented on GPU. In Journal of Physics: Conference Series, volume 1404, page 012119. IOP Publishing, 2019.
- [41] E. A. Malkov and A. N. Kudryavtsev. Non-stationary antonov self-gravitating layer: analytics and numerics. Monthly Notices of the Royal Astronomical Society, 491(3):3952–3966, 2020.
- [42] G. Manzini, G. L. Delzanno, J. Vencels, and S. Markidis. A Legendre-Fourier spectral method with exact conservation laws for the Vlasov-Poisson system. Journal of Computational Physics, 317:82–107, 2016.
- [43] J. Qiu and C. Shu. Positivity preserving semi-lagrangian discontinuous Galerkin formulation: Theoretical analysis and application to the Vlasov-Poisson system. Journal of Computational Physics, 230(23):8386–8409, 2011.
- [44] J.-M. Qiu and C.-W. Shu. Conservative semi-lagrangian finite difference WENO formulations with applications to the Vlasov equation. Communications in Computational Physics, 10(4):979, 2011.
- [45] D. Sirajuddin and W. N. Hitchon. A truly forward semi-lagrangian WENO scheme for the Vlasov-Poisson system. Journal of Computational Physics, 392:619–665, 2019.
- [46] E. Sonnendrücker, J. Roche, P. Bertrand, and A. Ghizzo. The semi-Lagrangian method for the numerical resolution of the Vlasov equation. Journal of Computational Physics, 149(2):201–220, 1999.

- [47] C. Standar. On Finite Element Schemes for Vlasov-Maxwell System and Schrödinger Equation. PhD thesis, Chalmers Tekniska Hogskola, 2017.
- [48] A. Suresh and H. T. Huynh. Accurate monotonicity-preserving schemes with Runge-Kutta time stepping. Journal of Computational Physics, 136(1):83–99, 1997.
- [49] S. Tanaka, K. Yoshikawa, T. Minoshima, and N. Yoshida. Multidimensional Vlasov–Poisson simulations with high-order monotonicity- and positivity-preserving schemes. The Astrophysical Journal, 849(2):76, nov 2017.
- [50] Z. Tao, W. Guo, and Y. Cheng. Sparse grid discontinuous Galerkin methods for the Vlasov-Maxwell system. arXiv: Numerical Analysis, 2018.
- [51] T. Umeda. A conservative and non-oscillatory scheme for Vlasov code simulations. Earth, planets and space, 60(7):773–779, 2008.
- [52] T. Umeda, Y. Nariyuki, and D. Kariya. A non-oscillatory and conservative semi-Lagrangian scheme with fourth-degree polynomial interpolation for solving the Vlasov equation. Computer Physics Communications, 183(5):1094–1100, 2012.
- [53] H. D. Victory and E. J. Allen. The convergence theory of particle-in-cell methods for multidimensional Vlasov-Poisson systems. SIAM Journal on Numerical Analysis, 28(5):1207–1241, 1991.
- [54] G. Vogman, U. Shumlak, and P. Colella. Conservative fourth-order finite-volume Vlasov–Poisson solver for axisymmetric plasmas in cylindrical  $(r, vr, v\theta)$  phase space coordinates. Journal of Computational Physics, 373:877 – 899, 2018.
- [55] T. Xiong, J.-M. Qiu, Z. Xu, and A. Christlieb. High order maximum principle preserving semi-lagrangian finite difference WENO schemes for the Vlasov equation. Journal of Computational Physics, 273:618–639, 2014.
- [56] Z. Xu and C. Shu. Anti-diffusive flux corrections for high order finite difference WENO schemes. Journal of Computational Physics, 205(2):458–485, 2005.
- [57] C. Yang and F. Filbet. Conservative and non-conservative methods based on hermite weighted essentially non-oscillatory reconstruction for Vlasov equations. Journal of Computational Physics, 279:18–36, 2014.
- [58] C. Yang and L. M. Tine. A hybrid finite volume method for advection equations and its applications in population dynamics. Numerical Methods for Partial Differential Equations, 33(4):1114–1142, 2017.
- [59] H. Yang and F. Li. Discontinuous Galerkin methods for relativistic Vlasov-Maxwell system. Journal of Scientific Computing, 73(2):1216–1248, 2017.

- [60] S. Zaki, L. Gardner, and T. Boyd. A finite element code for the simulation of one-dimensional Vlasov plasmas. i. theory. Journal of Computational Physics, 79(1):184–199, 1988.