



Conservative Multi-dimensional Semi-Lagrangian Finite Difference Scheme: Stability and Applications to the Kinetic and Fluid Simulations

Tao Xiong¹ · Giovanni Russo² · Jing-Mei Qiu³ 

Received: 25 July 2018 / Revised: 13 October 2018 / Accepted: 10 December 2018 /

Published online: 19 December 2018

© Springer Science+Business Media, LLC, part of Springer Nature 2018

Abstract

In this paper, we propose a mass conservative semi-Lagrangian finite difference scheme for multi-dimensional problems *without dimensional splitting*. The semi-Lagrangian scheme, based on tracing characteristics backward in time from grid points, does not necessarily conserve the total mass. To ensure mass conservation, we propose a conservative correction procedure based on a flux difference form. Such procedure guarantees local mass conservation, while introducing time step constraints for stability. We theoretically investigate such stability constraints from an ODE point of view by assuming exact evaluation of spatial differential operators and from the Fourier analysis for linear PDEs. The scheme is tested by classical two dimensional linear passive-transport problems, such as linear advection, rotation and swirling deformation. The scheme is applied to solve the nonlinear Vlasov–Poisson system and guiding center Vlasov model using high order tracing schemes. The effectiveness of the proposed conservative semi-Lagrangian scheme is demonstrated numerically by extensive numerical tests.

Keywords Semi-Lagrangian · Conservative · High order · WENO · Linear stability analysis · Fourier analysis · Vlasov–Poisson system

1 Introduction

Semi-Lagrangian (SL) schemes have been used extensively in many areas of science and engineering, including weather forecasting [11,14,26], kinetic simulations [8,12] and fluid

T. Xiong: Research supported by NSF Grant of Fujian Province 2016J05022, NSFC Grant 11601455, NSAF Grant U1630247, the Fundamental Research Funds for the Central Universities No. 20720160009 and the Marie Skłodowska-Curie Individual Fellowships H2020-MSCA-IF-2014 of the European Commission, under the Project HNSKMAP 654175. G. Russo: The work has been partially supported by ITN-ETN Horizon 2020 Project *ModCompShock*, *Modeling and Computation on Shocks and Interfaces*, Project Reference 642768. J.-M. Qiu: Research supported by Air Force Office of Scientific Computing Grant FA9550-18-1-0257, NSF Grant NSF-DMS-1818924 and University of Delaware.

✉ Jing-Mei Qiu
jingqiu@udel.edu

Extended author information available on the last page of the article

simulations [16,29], interface tracing [7,27], etc. The schemes are designed to combine the advantages of Eulerian and Lagrangian approaches. In particular, the schemes are built upon a fixed computational mesh. Similar to the Eulerian approach, high spatial resolution can be realized by using high order interpolation/reconstruction procedures or by using piecewise polynomial solution spaces. On the other hand, in each time step evolution, the scheme is designed by propagating information along characteristics, relieving the CFL condition. Typically, the numerical time step size allowed for an SL scheme is larger than that of an Eulerian approach, leading to gain in computational efficiency.

Among high order SL schemes, depending on solution spaces, different classes of methods can be designed. For example, a finite difference scheme evolves point-wise values and realizes high spatial resolution by high order interpolation procedures [17,29], a finite volume scheme considers integrated cell-averages with high order reconstruction procedures [6,14], while a finite element method has piecewise continuous or discontinuous polynomial functions as its solution space [11,15,16,20,21]. Each class of the above mentioned SL methods has its own advantages. For example, the finite element method is more flexible with the geometry and handling boundary conditions, while the finite difference and finite volume schemes could perform better in resolving solution structures with sharp gradients, e.g. by using a weighted essentially non-oscillatory (WENO) procedure. To compare finite difference and finite volume schemes, finite volume scheme is often considered more physically relevant and the local mass conservation can be built up in a natural way; while the finite difference scheme is more computationally efficient for high-dimensional problems, if one considers schemes of third order or higher.

In this paper, we consider SL finite difference schemes with *local mass conservation property*. In fact, many existing SL finite difference schemes are built based on tracing characteristics backward in time together with a high order interpolation procedure [3]. Typically such schemes do not have local mass conservation property, which is fine for certain applications. However, for applications in weather forecasting or in kinetic simulations, ignoring local mass conservation could lead to significant loss of total mass, especially when the solution with sharp gradients becomes under-resolved by the computational mesh [13].

There have been many attempts to preserve the mass conservation of an SL finite difference scheme with large time stepping sizes, e.g. [17,19]. However, they are mostly designed for 1D problems by taking advantage of some special features in a 1D setting. Their generalization to high dimensional problems often relies on dimensional splitting which is subject to splitting errors. In this paper, we propose and investigate a truly multi-dimensional approach without dimensional splitting errors. To build an SL finite difference scheme with local mass conservation, one essential framework that we propose to work with is the flux-difference form. However, by working with the flux difference form, one often may observe time step constraint for numerical stability. Note that unlike the Eulerian approach, such time step constraint does not come from the CFL condition (i.e. numerical domain of dependence should include the physical domain of dependence), but from numerical stability one employed for temporal integration. As far as we are aware of, there is little work in quantifying the stability constraint and in optimizing the numerical strategies balancing stability, accuracy and computational efficiency. This paper aims to fill such gap, by understanding such time step constraints. In particular, we investigate the stability of time integration schemes based on a linear stability analysis around the imaginary axis in the complex plane, assuming the spatial differentiation is exact. We optimize quadrature rules for time integration by maximizing the stability interval along the imaginary axis. We further employ Fourier analysis to study the numerical stability of a fully-discretized scheme. The schemes are applied to 2D passive-transport problems, as well as to nonlinear Vlasov–Poisson (VP), guiding center

Vlasov model and incompressible Euler system in vorticity stream function formulation, by using high order characteristics tracing schemes proposed in [18,28]. Finally, we would like to mention a few of our previous work related to stability of SL finite difference schemes in a flux-difference form. In [19], a special treatment is introduced to relieve the time step constraint for 1D passive transport problems. However, such treatment is not possible for general high dimensional problems. In [5], the time step constraint is studied by Fourier analysis for an SL finite difference scheme coupled with integral deferred correction framework.

The paper is organized as follows. The SL finite difference scheme in flux-difference form is described in Sect. 2. The stability of time integration with quadrature rule is investigated in Sect. 3, assuming *exact* evaluation of spatial differentiation operators. We also optimize temporal integration rules. In Sect. 4, we study the numerical stability of a fully discretized scheme by Fourier analysis. In Sect. 5, numerical tests are performed for 2D linear passive-transport problems. In Sect. 6, we apply the scheme to the nonlinear VP system, the guiding center Vlasov equation and incompressible Euler system.

2 A Mass Conservative SL Finite Difference Scheme

In this section, we describe an SL finite difference scheme based on a flux-difference form to locally preserve mass. The scheme starts from a standard non-conservative procedure with backward characteristics tracing and high order spatial interpolation. Then a conservative correction is performed by a flux-difference formulation. We describe the scheme in a 1D linear setting, noting that its extension to nonlinear and high dimensional problems is straightforward, as long as characteristics can be properly traced backward in time, e.g. see our numerical examples in Sect. 5.

We consider a 1D linear advection equation,

$$\frac{\partial f}{\partial t} + \frac{\partial f}{\partial x} = 0, \quad f(x, 0) = f^0(x), \quad x \in [-\pi, \pi]. \tag{2.1}$$

For simplicity, we assume a periodic boundary condition. We assume a uniform discretization in space with $x_j = j\Delta x$, $j = 1, \dots, n_x$ and let f_j^n be an approximation of the solution at time t^n and position x_j . We describe below the conservative SL procedure to update $\{f_j^{n+1}\}_{j=1}^{n_x}$ from $\{f_j^n\}_{j=1}^{n_x}$.

In an Eulerian finite difference method, typically one would firstly approximate the spatial derivative by a flux difference form to ensure mass conservation, then the system of ODEs will be evolved in time by a high order numerical integrator such as the Runge-Kutta (RK) method via the method of lines. In the SL setting, however, we propose to perform the time integration based on quadrature rules first,

$$f_j^{n+1} = f_j^n - \frac{\partial}{\partial x} \left(\int_{t^n}^{t^{n+1}} f(x, t) dt \right) \Big|_{x=x_j} \approx f_j^n - \frac{\partial}{\partial x} (\mathcal{F}(x)) \Big|_{x_j}, \tag{2.2}$$

where we let

$$\mathcal{F}(x) \doteq \sum_{\ell=1}^s f(x, t^n + c_\ell \Delta t) b_\ell \Delta t \tag{2.3}$$

as a quadrature approximation of $\int_{t^n}^{t^{n+1}} f(x, t) dt$. Here (c_ℓ, b_ℓ) , $\ell = 1, \dots, s$ are the nodes and weights of an accurate quadrature formula and $f(x_j, t^n + c_\ell \Delta t)$, $\ell = 1, \dots, s$ (we call it stage values) can be approximated via a non-conservative SL scheme via backward

characteristics tracing and high order spatial interpolation. For the linear equation (2.2), $f(x_j, t^n + c_\ell \Delta t)$ can be traced back along characteristics to t^n at $f(x_j - c_\ell \Delta t, t^n)$, whose value can be obtained via interpolation from neighboring grid point values $\{f_j^n\}_{j=1}^{n_x}$, see our description for different interpolation procedures in Sect. 4. Then a conservative scheme, based on a flux-difference form, can be proposed in the spirit of the work by Shu and Osher [24]. In particular, the scheme can be formulated as

$$f_j^{n+1} = f_j^n - \frac{1}{\Delta x} \left(\hat{F}_{j+\frac{1}{2}} - \hat{F}_{j-\frac{1}{2}} \right), \tag{2.4}$$

where $\hat{F}_{j+\frac{1}{2}}$ comes from WENO reconstruction of fluxes from $\{\mathcal{F}_j\}_{j=1}^{n_x}$ with $\mathcal{F}_j \doteq \mathcal{F}(x_j)$. We refer to [23] for the basic principle and detailed procedures of WENO reconstruction. Also, Sect. 4 provides detailed discussions on different reconstruction procedures. It can be shown that the mass conservation is locally preserved due to the flux difference form (2.4).

Such conservative correction procedure can be directly generalized to problems with non-constant velocity fields in a multi-dimensional setting without any difficulty, e.g. rotation and swirling deformation. In addition to the procedures described above, a high order ODE integrator such as a RK method can be employed to locate the foot of a characteristic accurately. For example, we consider a 2D problem with a prescribed velocity field $a(x, y, t)$ and $b(x, y, t)$

$$f_t + (a(x, y, t)f)_x + (b(x, y, t)f)_y = 0.$$

Let the set of grid points

$$x_1 < \dots < x_i < \dots < x_{n_x}, \quad y_1 < \dots < y_j < \dots < y_{n_y} \tag{2.5}$$

be a uniform discretization of a 2D rectangular domain with $x_i = i \Delta x$ and $y_j = j \Delta y$. The foot of characteristic emanating from a 2D grid point, say (x_i, y_j) at $t^\ell \doteq t^n + c_\ell \Delta t$ can be located by solving the following final-value problem accurately with a high order RK method,

$$\frac{dx}{dt} = a(x, y, t), \quad \frac{dy}{dt} = b(x, y, t), \quad x(t^\ell) = x_i, \quad y(t^\ell) = y_j. \tag{2.6}$$

Once the foot of characteristic located, say at (x_i^*, y_j^*) , then $f(x_i, y_j, t^\ell)$ can be evaluated by approximating $f(x_i^*, y_j^*, t^n)$ via a high order 2D interpolation procedure [23]. A 2D conservative scheme based on a flux-difference form can be formulated as

$$f_{ij}^{n+1} = f_{ij}^n - \frac{1}{\Delta x} \left(\hat{F}_{i+\frac{1}{2},j} - \hat{F}_{i-\frac{1}{2},j} \right) - \frac{1}{\Delta y} \left(\hat{G}_{i,j+\frac{1}{2}} - \hat{G}_{i,j-\frac{1}{2}} \right), \tag{2.7}$$

where $\hat{F}_{i\pm\frac{1}{2},j}$ comes from WENO reconstruction of fluxes from $\{\mathcal{F}_{ij}\}_{i=1}^{n_x}$ for all j with

$$\mathcal{F}_{ij} \doteq \mathcal{F}(x_i, y_j) \approx \Delta t \sum_{\ell=1}^s f(x_i, y_j, t^n + c_\ell \Delta t) b_\ell.$$

The procedure for WENO reconstruction is the same as the 1D case for all j and we again refer to the review paper [23]. Similarly, $\hat{G}_{i,j\pm\frac{1}{2}}$ comes from WENO reconstruction of fluxes from $\{\mathcal{F}_{ij}\}_{j=1}^{n_y}$ for all i .

To generalize the conservative SL scheme to nonlinear systems, a problem-dependent high order characteristics tracing procedure needs to be designed for solving the final-value problem in the form of Eq. (2.6), but with the velocity field depending on the unknown function f . In many cases, a high order RK method could not be directly applied. In [18], a high order

multi-dimensional characteristics tracing scheme for the VP system is proposed, and in [28], it has been generalized for the guiding center Vlasov system and the incompressible Euler system in the vorticity-stream function formulation. Both can be applied in the above proposed conservative SL framework.

We close this section by making the following remark to motivate our discussions in the following two sections. There are two issues in the scheme formulation that contribute to the stability property of the above proposed SL scheme. One is the discretization by the quadrature rule (2.3). This part of stability is viewed as an ODE stability (assuming exact evaluation of spatial operators) and is investigated carefully in Sect. 3. The other issue concerns the spatial discretization, and can be explained by observing the following situation: if one changes the time stepping size slightly (could be arbitrary small), the root of characteristics $x_j - c_\ell \Delta t$ could come from a different grid cell, leading to a different interpolation stencil in the implementation. This aspect is investigated in Sect. 4.

3 Temporal Discretization and Stability

3.1 Linear stability functions and stability regions

We first investigate the linear stability of quadrature rules for temporal discretization (2.3) in an ODE setting, by assuming an *exact* evaluation of spatial derivative in Eq. (2.2). In particular, we look for the evolution of a Fourier mode, identified by a Fourier variable $\xi \in [-\pi, \pi]$, assuming exact evaluation of spatial interpolation and reconstruction procedure mentioned above. Such a discrete Fourier mode at time $t^n = n \Delta t$, will be denoted by,

$$f_\xi^n(x) = (Q(\xi))^n e^{i x \xi / \Delta x}, \quad \mathbf{i} = \sqrt{-1}, \tag{3.1}$$

where $Q(\xi)$ is the amplification factor associated with ξ . After plugging such ansatz into the expression for the flux, Eq. (2.3), and taking into account Eq. (3.1), we have:

$$\mathcal{F}(x) = \sum_{\ell=1}^s b_\ell f^n(x - c_\ell \Delta t) \Delta t = Q^n(\xi) \sum_{\ell=1}^s b_\ell \exp [i \xi (x - c_\ell \Delta t) / \Delta x] \Delta t$$

Using such expression in Eq. (2.2) gives:

$$f_j^{n+1} = f_j^n - Q^n(\xi) \exp(i \xi x / \Delta x) \frac{\Delta t}{\Delta x} i \xi \sum_{\ell=1}^s b_\ell \exp(-i \xi c_\ell \Delta t / \Delta x)$$

from which we deduce

$$Q(\xi) = 1 - i \xi a \sum_{\ell=1}^s b_\ell \exp(-i c_\ell \xi a), \tag{3.2}$$

where $a = \Delta t / \Delta x$ denotes the Courant number. The scheme is stable if

$$|Q(\xi)| \leq 1, \quad \forall \xi \in [-\pi, \pi].$$

Such stability property is closely related to the linear stability of the quadrature rule, which can be studied by the stability region for a scalar linear ODE.

3.2 Relation with A-Stability

Let us recall the definition of A-stability. Consider the scalar ODE

$$w'(t) = \lambda w(t), \quad w(0) = 1 \tag{3.3}$$

and $\lambda \in \mathbb{C}$. The solution after one time step is given by

$$w(\Delta t) = e^{\lambda \Delta t} = e^z,$$

where we set $z \equiv \lambda \Delta t$. Such solution is stable if and only if $\Re(z) \leq 0$, i.e. if and only if $\Re(\lambda) \leq 0$. Consider now the following identity

$$e^z = 1 + z \int_0^1 e^{\alpha z} d\alpha. \tag{3.4}$$

This is really an identity, because $\int_0^1 e^{\alpha z} d\alpha = (e^z - 1)/z =: \varphi(z)$. Now let us approximate the integral appearing in Eq. (3.4) by a quadrature formula, whose nodes and weights are, respectively, c_ℓ and b_ℓ , $\ell = 1, \dots, s$. Then one obtains the following approximation of the exact solution of Eq. (3.3) after one time step:

$$R(z) = 1 + z \sum_{\ell=1}^s b_\ell e^{c_\ell z}. \tag{3.5}$$

It is clear that in general it is not true that $|R(z)| \leq 1$ if $\Re(z) \leq 0$. Neglecting all other approximations, the core reason of the instability of the SL conservative method relies on the lack of unconditional stability in the approximation of e^z by $R(z)$ for z purely imaginary. This relation can be immediately applied to the study of the stability of the SL conservative scheme, since comparing Eq. (3.2) with Eq. (3.5), one has

$$Q(\xi) = R(-i\xi a).$$

Because $\xi \in [-\pi, \pi]$, the stability problem can be stated as follows: look for the largest interval $I^* \equiv [-y^*, y^*]$ of the imaginary axis such that $|R(iy)| \leq 1, \forall y \in I^*$. Then the maximum CFL number for the SL scheme that guarantees stability will be

$$a^* = y^*/\pi \tag{3.6}$$

Below, we report the stability regions for the following commonly used quadrature rules in the left panel of Fig. 1.

1. midpoint: $q_1 = 1/2, w_1 = 1$.
2. trapezoidal: $q_1 = 0, q_2 = 1, w_1 = w_2 = 1/2$.
3. Simpson: $q_1 = 0, q_2 = 1/2, q_3 = 1, w_1 = w_3 = 1/6, w_2 = 2/3$.
4. two-point Gauss–Legendre formulas (GL2): $q_1 = \frac{1}{2} - \frac{1}{2\sqrt{3}}, q_2 = \frac{1}{2} + \frac{1}{2\sqrt{3}}, w_1 = w_2 = 1/2$.

As it is apparent from the plot, midpoint and Simpson’s rule do not include a portion of the imaginary axis, while the trapezoidal rule and the two-point Gauss–Legendre rule do. The boundary of the stability region of the trapezoidal rule intersects the imaginary axis at π , and therefore the maximum CFL number that guarantees linear stability for the conservative scheme is $\pi/\pi = 1$. The two point Gauss–Legendre quadrature formula provides a wider stability interval, since in this case $y^* \approx 5.43$, giving a maximum CFL number of approximately $5.43/\pi = 1.72$. Higher order Gauss–Legendre quadrature formulas, hereafter

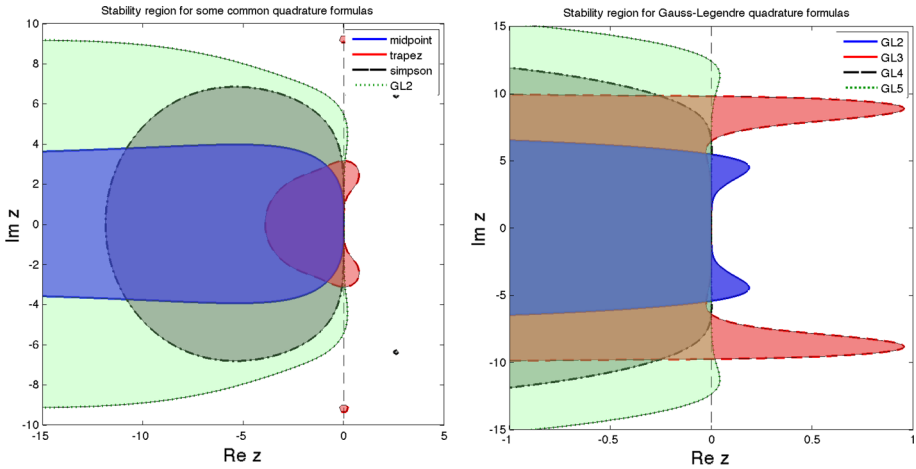


Fig. 1 Stability region for midpoint, trapezoidal, Simpson’s and two-point Gauss–Legendre rules (left) and GLs with $s = 2, 3, 4, 5$ (right)

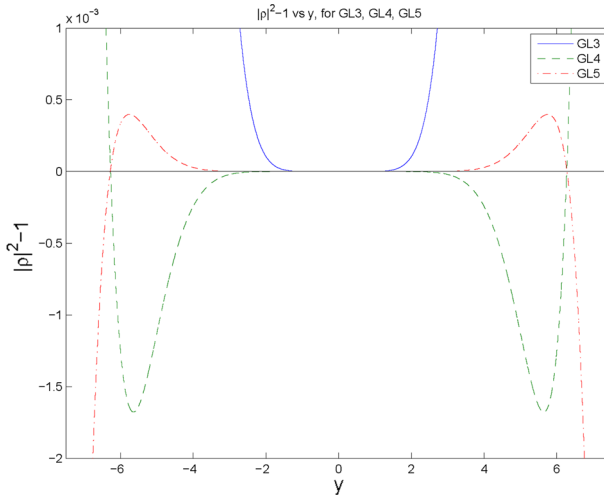


Fig. 2 $R(iy)^2 - 1$ versus y for GL3, GL4, GL5 formulas. GL3 and GL5 is observed to be unstable

denoted by GLs, where s indicates the number of nodes, may provide wider stability interval, as is illustrated in the right panel of Fig. 1. To better appreciate the stability region, we plot in Fig. 2 $|R(iy)|^2 - 1$ as a function of y . GL4 is observed to have better stability property, as $|R(iy)|^2 - 1 \leq 0$ for an interval with boundary $y^* \approx 6.2765$ leading to a maximum CFL number of approximately $6.2765/\pi = 1.99$. Gauss–Legendre rule formulas with odd number of points, such as GL3 and GL5, are unstable near the origin, see the right panel of Fig. 1 as well as Fig. 2.

3.3 Maximize the Stability Interval on Imaginary Axis

In order to analyze the stability of quadrature formulas, let us consider the expression $R(\mathbf{i}y)$ from Eq. (3.5), and write it in the form

$$R(\mathbf{i}y) = 1 + \mathbf{i}y(C_s(y) + iS_s(y)) = 1 - yS_s(y) + \mathbf{i}yC_s(y) \tag{3.7}$$

where

$$C_s(y) \equiv \sum_{\ell=1}^s b_\ell \cos(c_\ell y), \quad S_s(y) \equiv \sum_{\ell=1}^s b_\ell \sin(c_\ell y). \tag{3.8}$$

The stability condition therefore becomes

$$|R(\mathbf{i}y)|^2 = 1 - 2yS_s(y) + y^2(C_s^2(y) + S_s^2(y)) \leq 1.$$

Such condition can be written in the form

$$yF_s(y) \geq 0, \tag{3.9}$$

where

$$F_s(y) \equiv S_s(y) - \frac{1}{2}y(C_s^2(y) + S_s^2(y)). \tag{3.10}$$

The problem of finding quadrature formulas with the widest stability region can be stated as: determine the coefficients $\mathbf{b} = (b_1, \dots, b_s)$ and $\mathbf{c} = (c_1, \dots, c_s)$ so that the interval in which (3.9) is satisfied is the widest.

Rather than directly solving this optimization problem, we consider a particular case of quadrature formulas, i.e. those for which the nodes are symmetrically located with respect to point 1/2 in the interval [0, 1]. The choice is motivated by the observation that several quadrature formulas have a symmetric distribution of nodes and weights around the center of the integration interval. Furthermore, such an assumption greatly simplifies the derivation and analysis of the methods. Among such formulas we restrict to the case in which s even, as the schemes are observed to be unstable for odd s , see Fig. 2.

Let us denote by $\tilde{c}_\ell = 1 - 2c_\ell, \ell = 1, \dots, s$. Then $c_\ell = (1 - \tilde{c}_\ell)/2$. Since the nodes are symmetric and the quadrature formula is interpolatory, we have

$$\tilde{c}_\ell = -\tilde{c}_{s-\ell+1}, \quad b_\ell = b_{s-\ell+1}. \tag{3.11}$$

The absolute stability function $R(\mathbf{i}y)$ can then be written, after simple manipulations, as

$$R(\mathbf{i}y) = 1 - 2y \sin(y/2)\tilde{C}_s(y) + 2\mathbf{i}y \cos(y/2)\tilde{C}_s(y),$$

where

$$\tilde{C}_s(y) \equiv \sum_{\ell=1}^{s/2} b_\ell \cos(\tilde{c}_\ell y/2),$$

leading to

$$|R(\mathbf{i}y)|^2 = 1 - 4y \sin(y/2)\tilde{C}_s(y) + 4y^2\tilde{C}_s^2(y). \tag{3.12}$$

The function $F_s(y)$ can then be written, after simple manipulations

$$F_s(y) = 2\tilde{C}_s(y) \left(\sin(y/2) - y\tilde{C}_s(y) \right). \tag{3.13}$$

Then the stability condition (3.9) becomes

$$\tilde{C}_s(y) \left(\sin(y/2) - y\tilde{C}_s(y) \right) \geq 0.$$

Because function F_s contains the product of two factors, the condition to ensure that the function does not change sign at roots is that the two factors vanish simultaneously at simple roots, therefore $\tilde{C}_s(y)$ has to vanish also at the same points $y_k > 0$ at which $\sin(y/2) - y\tilde{C}_s(y) = 0$. There is no need to impose that \tilde{C}_s vanishes at the origin, since, because of symmetry, $yF_s(y)$ does not change sign at the origin.

In order to determine the coefficients that define the quadrature formula for maximizing the stability interval on imaginary axis, we proceed as follows. Because of the symmetry constraints (3.11), we have to find s coefficients, i.e. $b_1, \dots, b_{s/2}$ and $\tilde{c}_1, \dots, \tilde{c}_{s/2}$ by imposing a total of s conditions. Such conditions will be a balance between accuracy and stability. If we want that the quadrature formulas have degree of precision $s - 1$, i.e. if we want that they are exact for polynomials of degree less or equal to $s - 1$, we have to impose

$$\frac{1}{2} \int_{-1}^1 \xi^{2k} d\xi = \sum_{\ell=1}^s b_\ell (\tilde{c}_\ell)^{2k}, \quad k = 0, \dots, s/2 - 1. \tag{3.14}$$

We only impose the condition for even polynomials, since for odd polynomials the conditions are automatically satisfied because of the symmetry. The condition that $\tilde{C}_s(y)$ vanishes when $\sin(y/2) - y\tilde{C}_s(y)$ vanishes becomes

$$\tilde{C}_s(2\pi k) = 0, \quad k = 1, \dots, s/2. \tag{3.15}$$

For $k = 0$ the stability condition (3.15) is marginally satisfied since $\tilde{C}_s(0) = \sum_{\ell=1}^{s/2} b_\ell = 1/2$. Eqs. (3.14) and (3.15) constitute a nonlinear set of equations for the s coefficients $b_1, \dots, b_{s/2}$ and $\tilde{c}_1, \dots, \tilde{c}_{s/2}$. Because the equations are nonlinear, we have to resort to Newton’s method for their solution. In practice, for large values of s , it is hard to find an initial guess which lies in the convergence basin of Newton’s method. We had to resort to a relaxed version of Newton’s method, coupled with continuation techniques, in order to solve the system.

We numerically compute nodes and weights for $s = 2, 4, 6, 8, 10, 12$ and check *a posteriori* whether the stability condition is actually satisfied. The following phenomena are observed:

- $s = 2$: the quadrature nodes and weights are consistent with those of the two-point Gauss–Legendre formula.
- $s = 4, 8, 12$: In Fig. 3 we plot the functions $|R_s(iy)|^2 - 1$ (left panel) and the corresponding stability regions in the complex plane (right panel) for $s = 4, 8, 12$. A wide interval with stability on the imaginary axis is shown. We report the coefficients in Table 1. Only $s/2$ coefficients are reported, since the other satisfy the symmetry relation (3.11). We also report the maximum CFL number $a^* = y^*/\pi$ [see Eq. (3.6)].
- $s = 6$ and $s = 10$: In Fig. 4, we plot the functions $|R_s(iy)|^2 - 1$ for $s = 6$ and $s = 10$. It is observed that $|R_s(iy)|^2 \geq 1$ for any interval containing the origin, i.e. these two quadrature formulas are not stable.

The stability regions for the quadrature formulas obtained for $s = 4, 8, 12$ reported in Table 1 are computed under the assumption that one considers the exact space dependence of the Fourier mode, so that the only error is in time integration. In reality there are several other causes of errors, that may affect the stability region of the quadrature. In the next section, we take spatial discretization into account and quantify the corresponding stability interval.

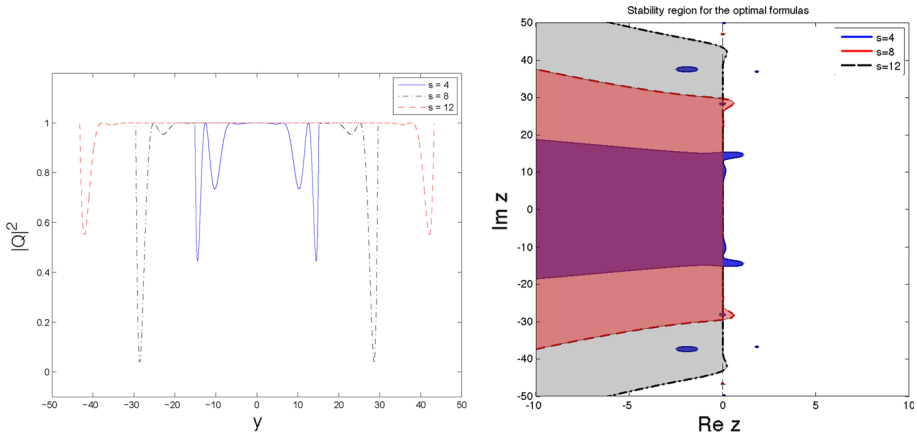


Fig. 3 Left: plot of $|R(iy)|^2$ for the symmetric quadrature formulas with $s = 4, 8, 12$. Right: plot of stability regions for the corresponding quadrature formulas. These formulas show a wide stability interval on the imaginary axis, thus allowing, in principle, large CFL numbers

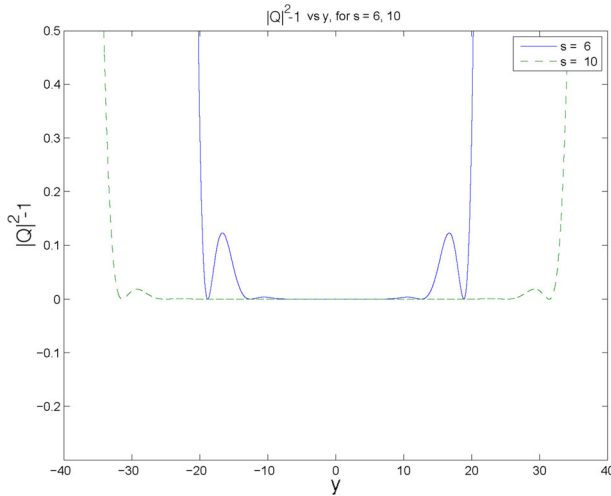


Fig. 4 Plot of $|R(iy)|^2 - 1$ for the symmetric quadrature formulas with $s = 6, 10$. Such formulas are not stable because the stability region do not contain a portion of the imaginary axis

4 Spatial Discretization

There are two spatial discretization processes in the scheme. One is the interpolation in approximating $f(x_j, t^n + c_\ell \Delta t) = f(x_j - c_\ell \Delta t, t^n)$ from neighboring grid point values $\{f_j^n\}_{j=1}^{n_x}$. The other is the reconstruction in obtaining numerical fluxes $\hat{F}_{j+\frac{1}{2}}$ in (2.4) from $\{\mathcal{F}_j\}_{j=1}^{n_x}$. In this paper, we consider the following two classes of spatial discretizations.

- **Odd order approximations.** For the linear equation (2.1), we use a right-biased stencil to approximate $f(x_j - c_\ell \Delta t, t^n)$ and use a left-biased stencil for reconstructing the flux $\hat{F}_{j+\frac{1}{2}}$. For example, for a first order scheme with $\Delta t / \Delta x < 1$, $f(x_j - c_\ell \Delta t, t^n)$

Table 1 Weights and nodes of accurate and stable quadrature formulas

$s = 4, a^* = 4.8125674352016$		
1	0.199889211759008	0.083205952308564
2	0.300110788240992	0.347904700949451
$s = 8, a^* = 9.4130380474585$		
1	0.058702317190867	0.023248965963790
2	0.119923212650690	0.114686793929813
3	0.154113350301760	0.253867587586135
4	0.167261119856682	0.415892817555109
$s = 12, a^* = 13.7671988660496$		
1	0.027182888487959	0.010668025829619
2	0.059633412276882	0.054560771376909
3	0.084799522112170	0.127471263371368
4	0.101625491473440	0.221353922812027
5	0.111259037829236	0.328318059665840
6	0.115499647820313	0.442082833046309

Each formula is exact for polynomials of degree not greater than $s - 1$. The maximum CFL number a^* that guarantees stability in the theoretical Fourier analysis is reported

is approximated from the *interpolation* stencil $\{f_j\}$ and the numerical flux $\hat{F}_{j+\frac{1}{2}}$ is approximated from the *reconstruction* stencil $\{\mathcal{F}_j\}$. With such stencil arrangement, the SL scheme is reduced to a first order upwind scheme when $\Delta t/\Delta x < 1$,

$$f_j^{n+1} = f_j^n - \Delta t/\Delta x(f_j^n - f_{j-1}^n).$$

Third, fifth, seventh and ninth order schemes can be constructed by including one, two, three, four more points symmetrically from left and from right, respectively, in the interpolation and reconstruction stencils. We list them as follows.

- Third order : $\{f_{j-1}, f_j, f_{j+1}\}, \{\mathcal{F}_{j-1}, \mathcal{F}_j, \mathcal{F}_{j+1}\}.$
- Fifth order : $\{f_{j-2}, f_{j-1}, f_j, f_{j+1}, f_{j+2}\}, \{\mathcal{F}_{j-2}, \mathcal{F}_{j-1}, \mathcal{F}_j, \mathcal{F}_{j+1}, \mathcal{F}_{j+2}\}.$
- Seventh order : $\{f_{j-3}, f_{j-2}, f_{j-1}, f_j, f_{j+1}, f_{j+2}, f_{j+3}\},$
 $\{\mathcal{F}_{j-3}, \mathcal{F}_{j-2}, \mathcal{F}_{j-1}, \mathcal{F}_j, \mathcal{F}_{j+1}, \mathcal{F}_{j+2}, \mathcal{F}_{j+3}\}.$
- Ninth order : $\{f_{j-4}, f_{j-3}, f_{j-2}, f_{j-1}, f_j, f_{j+1}, f_{j+2}, f_{j+3}, f_{j+4}\},$
 $\{\mathcal{F}_{j-4}, \mathcal{F}_{j-3}, \mathcal{F}_{j-2}, \mathcal{F}_{j-1}, \mathcal{F}_j, \mathcal{F}_{j+1}, \mathcal{F}_{j+2}, \mathcal{F}_{j+3}, \mathcal{F}_{j+4}\}.$

- Even order approximations. For the linear equation (2.1), we use symmetric stencils to approximate $f(x_j - c_\ell \Delta t, t^n)$ by interpolation and to approximate $\hat{F}_{j+\frac{1}{2}}$ by reconstruction. For example, for a second order scheme with $\Delta t/\Delta x < 1$, $f(x_j - c_\ell \Delta t, t^n)$ is approximated from the *interpolation* stencil $\{f_{j-1}, f_j\}$ and the numerical flux $\hat{F}_{j+\frac{1}{2}}$ is approximated from the *reconstruction* stencil $\{\mathcal{F}_j, \mathcal{F}_{j+1}\}$. Fourth, sixth and eighth order schemes can be constructed by including one, two, three more points symmetrically from left and from right, respectively, in the interpolation and reconstruction stencils. We list

them as follows.

- Fourth order : $\{f_{j-2}, f_{j-1}, f_j, f_{j+1}\}, \{\mathcal{F}_{j-1}, \mathcal{F}_j, \mathcal{F}_{j+1}, \mathcal{F}_{j+2}\}.$
- Sixth order : $\{f_{j-3}, f_{j-2}, f_{j-1}, f_j, f_{j+1}, f_{j+2}\},$
 $\{\mathcal{F}_{j-2}, \mathcal{F}_{j-1}, \mathcal{F}_j, \mathcal{F}_{j+1}, \mathcal{F}_{j+2}, \mathcal{F}_{j+3}\}.$
- Eighth order : $\{f_{j-4}, f_{j-3}, f_{j-2}, f_{j-1}, f_j, f_{j+1}, f_{j+2}, f_{j+3}\},$
 $\{\mathcal{F}_{j-3}, \mathcal{F}_{j-2}, \mathcal{F}_{j-1}, \mathcal{F}_j, \mathcal{F}_{j+1}, \mathcal{F}_{j+2}, \mathcal{F}_{j+3}, \mathcal{F}_{j+4}\}.$

Remark 4.1 We follow the same principle in the interpolation and reconstruction procedures in more general settings, for example the situation when the time stepping size is greater than the CFL restriction, i.e $\Delta t/\Delta x \geq 1$ for Eq. (2.1). For example, if $\Delta t/\Delta x$ is between 1 and 2 then all the stencils for the computation of $f(x_j - \Delta t)$ are shifted to the left by one, and so on. For general high dimensional problems, e.g. the Vlasov equation, similar procedures can be applied in a truly multi-dimensional fashion.

To access the stability property of the conservative method, we perform Fourier analysis via the linear equation (2.1) with $x \in [0, 2\pi]$ and periodic boundary condition. In particular, we make the ansatz $f_j^n = \hat{f}^n e^{i j \xi}$ with $\mathbf{i} = \sqrt{-1}$ and $\xi \in [0, 2\pi]$. Plugging the ansatz into the SL conservative scheme as described in Sect. 2, but with both linear interpolation and linear reconstruction (i.e. the WENO coefficients are frozen to constant values that guarantee maximum accuracy), we obtain $\hat{f}^{n+1}(\xi) = Q_\lambda(\xi) \hat{f}^n(\xi)$ with $Q_\lambda(\xi)$ being the amplification factor for the Fourier mode associated with ξ and $\lambda = \frac{\Delta t}{\Delta x}$. To ensure linear stability, it is sufficient to have

$$|Q_\lambda(\xi)| \leq 1, \quad \forall \xi \in [0, 2\pi], \quad \forall \lambda \in [0, \lambda^*], \quad \text{for some } \lambda^*. \tag{4.1}$$

We seek for λ^* by numerically checking the inequality (4.1) for 100 discretized grid points on $\xi \in [0, 2\pi]$, and by gradually increasing λ with a step size of 0.01 starting from $\lambda = 0$. Taking the machine precision into account in our implementation, we check the inequality $|Q_\lambda(\xi)| \leq 1+10^{-11}$ instead. We tabulate such λ^* in Table 2 for different quadrature formulas as discussed in Sect. 3 and with different choices of spatial interpolation and reconstruction stencils with odd and even order respectively. One can observe that the second order trapezoidal rule and the fourth order GL2 perform much better than the mid-point rule in terms of stability, especially for high order spatial approximations. The time stepping sizes allowed for stability of fully discretized schemes with $s = 4, 8, 12$ are observed to be much less than the one provided by ODE stability analysis in the previous section. The reason for this behavior is not clear, and a deeper analysis is needed to study the stability of fully discrete schemes, on order to obtain more effective quadrature formulas. Such analysis is however beyond the scope of the present paper and will be subject of future investigation.

In the following, we take the linear advection equation $u_t + u_x = 0$ with a smooth initial function $\sin(2\pi x)$ on the domain $[0, 1]$, to test the CFL bounds in Table 2. Here for better illustration, only linear interpolation and linear reconstruction are used (i.e. the WENO weights are frozen to obtain maximum order of accuracy). We consider schemes of coupling GL2 for temporal integration with third or fourth order spatial approximations. Errors and orders of convergence at a final integration time $T = 100.1$ for the 3rd order and $T = 1001.1$ for the 4th order are recorded in Table 3. Clear third order and fourth order spatial accuracies are observed at the corresponding upper bounds for CFL (1.22 for 3rd order and 1.84 for 4th order as in Table 2.) The code will blow up with the CFL increased by 0.01 at the corresponding time, which confirms the validity of the CFL bounds in the table. We have similar observations for schemes of higher order, but omitting them to save space. Although even order schemes comparatively have larger CFL bounds than odd order ones, for solutions

Table 2 Upper bounds of CFL for FD SL scheme with odd and even order interpolation and reconstruction

Temp/spatial	1st	3rd	5th	7th	9th	Exact
Mid-point	1.00	1.00	0.14	0.04	0.02	0.00
Trapezoid	1.99	1.68	1.52	1.44	1.38	1.00
Simpson	1.33	1.50	1.35	0.71	0.37	0.00
GL2	1.00	1.22	1.19	1.16	1.15	1.72
s = 4	1.00	1.37	1.27	1.22	1.19	4.81
s = 8	1.00	1.35	1.26	1.21	1.18	9.41
s = 12	1.00	1.37	1.25	1.21	1.18	13.76
Temp/spatial	2nd	4th	6th	8th	10th	Exact
Mid-point	2.00	0.04	0.01	0.00	0.00	0.00
Trapezoid	1.29	1.26	1.24	1.22	1.20	1.00
Simpson	3.00	2.91	0.83	0.34	0.20	0.00
GL2	1.85	1.84	1.84	1.83	1.83	1.72
s = 4	1.96	1.97	1.98	1.98	1.98	4.81
s = 8	1.99	1.99	1.99	1.99	1.99	9.41
s = 12	1.99	1.99	1.99	1.99	2.00	13.76

The amplification factor is bounded by $1 + 10^{-11}$. $N = 100$

Table 3 Accuracy test of the linear advection equation $u_t + u_x = 0$ with the initial function $\sin(2\pi x)$ for the 3rd order scheme with CFL = 1.22 at $T = 100.1$ and 4th order with CFL = 1.84 at $T = 1001.1$

Scheme	N	L^1 error	Order	L^∞ error	Order
3rd order	240	4.12E-04	–	6.47E-04	–
	480	5.15E-05	3.00	8.09E-05	3.00
	960	6.44E-06	3.00	1.01E-05	3.00
	1920	8.04E-07	3.00	1.26E-06	3.00
4th order	120	1.76E-03	–	2.77E-03	–
	240	1.10E-04	4.00	1.73E-04	4.00
	480	6.89E-06	4.00	1.08E-05	4.00
	960	4.76E-07	3.85	7.48E-07	3.85

with discontinuities, we can observe that odd order schemes with upwind mechanism can resolve the discontinuities better, see Fig. 5 for the numerical solutions of our schemes with linear weights for advecting a step function. Due to the above considerations, we will only consider the scheme with 5th order spatial approximation and GL2 for temporal integration in the following numerical sections. We would mention that the numerical oscillations for discontinuous solutions due to a linear scheme can be well controlled by only using WENO in the reconstruction step, not in the interpolation, see Fig. 5 for 5th order linear interpolation with 5th order WENO reconstruction, and also the tests in the numerical section.

5 Numerical Tests on 2D Linear Passive-Transport Problems

In this section, the conservative truly multi-dimensional SL scheme will be tested for passive transport equations, such as linear advection, rotation and swirling deformation. Since the

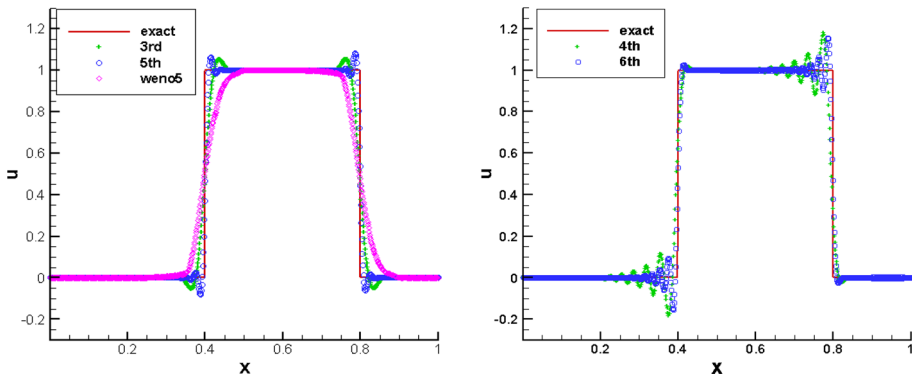


Fig. 5 Numerical solution for the linear advection equation $u_t + u_x = 0$ with an initial step function at $T = 100.1$. Left: 3rd order with CFL = 1.22 and 5th order with CFL = 1.19; Right: 4th order and 6th order with CFL = 1.84. $N = 800$. Here “weno5” means 5th order linear interpolation with 5th order WENO reconstruction, while others all using linear interpolation with linear reconstruction in corresponding orders

Table 4 Errors and orders for the linear equation in space

$N_x \times N_y$	20×20	40×40	60×60	80×80	100×100
L^1 error	2.76E-4	8.38E-6	1.11E-6	2.64E-7	8.68E-8
Order	–	5.04	4.99	4.98	4.99

$T = 1.2$. CFL = 1.15

velocity of the field is given a priori, characteristics can be traced by a high order RK ODE integrator.

In this and next sections, we use 5th order spatial approximations with linear interpolation and WENO reconstruction for evaluating flux functions in (2.4). We use GL2 for temporal integration, while characteristics are traced back in time by RK to locate feet of characteristics. For a general two dimensional problem $u_t + f(u)_x + g(u)_y = 0$, the time step is taken as

$$\Delta t = \text{CFL}/(a/\Delta x + b/\Delta y),$$

where $a = \max |f'(u)|$ and $b = \max |g'(u)|$. From Table 2, the maximum CFL number allowed by stability is 1.22 for a 3rd order discretization and 1.19 for the 5th order. In the following, we take CFL = 1.15 without specification.

Example 5.1 We first test our problem for the linear equation $u_t + u_x + u_y = 0$ with initial condition $u(x, y, 0) = \sin(x) \sin(y)$. The exact solution is $u(x, y, t) = \sin(x - t) \sin(y - t)$. For this example, the roots of characteristics are located exactly. Tables 4 and 5 present spatial and temporal orders of convergence of the proposed scheme. Both 5th order spatial accuracy and 4th order temporal accuracy from GL2 can be observed.

Example 5.2 Now we consider two problems defined on the domain $[-\pi, \pi]^2$. One is the rigid body rotating problem

$$u_t - yu_x + xu_y = 0,$$

the other is the swirling deformation flow problem

$$u_t - \left(\cos^2\left(\frac{x}{2}\right) \sin(y)g(t)u \right)_x + \left(\sin(x) \cos^2\left(\frac{y}{2}\right) g(t)u \right)_y = 0,$$

Table 5 Errors and orders for the linear equation in time

CFL	1.1	1.0	0.9	0.8	0.7
L^1 error	3.34E-9	2.27E-9	1.49E-9	9.29E-10	5.46E-10
Order	–	4.07	3.97	4.02	3.98

$N_x = N_y = 200. T = 1$

with $g(t) = \cos(\pi t/T)\pi$. Both have the initial condition, which includes a slotted disk, a cone as well as a smooth hump, see Figs. 6 (top left) and 7 (top left). To trace characteristics backward in time in the SL framework, we use a fourth order Runge-Kutta method in solving final value problems.

For the rigid body rotating problem, its period is 2π . In left panels of Fig. 6, we have shown the results at a half period and one period. As we can see, the shape of the bodies are well preserved. For the swirling deformation flow problem, after a half period the bodies are deformed; and they regain its initial shape after one period, see left panels in Fig. 7. On right panels, 1D cuts of numerical solutions are shown, benchmarked with the exact reference solutions.

6 Numerical Tests of Nonlinear Systems

In this section, we test the conservative SL scheme, which is denoted as “Cons-SL”, on the nonlinear VP system, the guiding center Vlasov system and the incompressible Euler system in vorticity-stream function formulation. Despite different application backgrounds, the latter two systems have indeed almost the same mathematical formulation, only with different signs in the Poisson’s equation.

We will compare our conservative scheme to the nonconservative one [17], which is denoted as “NonCons-SL”. For the nonconservative scheme “NonCons-SL”, we take 6th order spatial reconstruction, which is 5th order accurate in space globally, and CFL = 1.15 the same as the conservative scheme.

6.1 Vlasov–Poisson System

Arising from collisionless plasma applications, the rescaled VP system [4,8,25] writes

$$\frac{\partial f}{\partial t} + \mathbf{v} \cdot \nabla_{\mathbf{x}} f + \mathbf{E}(\mathbf{x}, t) \cdot \nabla_{\mathbf{v}} f = 0, \tag{6.1}$$

where

$$\mathbf{E}(\mathbf{x}, t) = -\nabla_{\mathbf{x}}\phi(\mathbf{x}, t). \quad -\Delta_{\mathbf{x}}\phi(\mathbf{x}, t) = \rho(\mathbf{x}, t) - 1. \tag{6.2}$$

The system describes the temporal evolution of the particle distribution function in six dimensional phase space. $f(\mathbf{x}, \mathbf{v}, t)$ is the probability distribution function which describes the phase space density of charged particles with velocity \mathbf{v} at position \mathbf{x} at time t , \mathbf{E} is the electric field, and ϕ is the self-consistent electrostatic potential. The probability distribution function couples to the long range fields via the charge density, $\rho(t, \mathbf{x}) = \int_{\mathbb{R}^{d_v}} f(\mathbf{x}, \mathbf{v}, t) d\mathbf{v}$, where d_v is the dimension in the velocity field, and we take the limit of uniformly distributed infinitely massive ions in the background, whose charge density is assumed to be 1 (with mines sign)

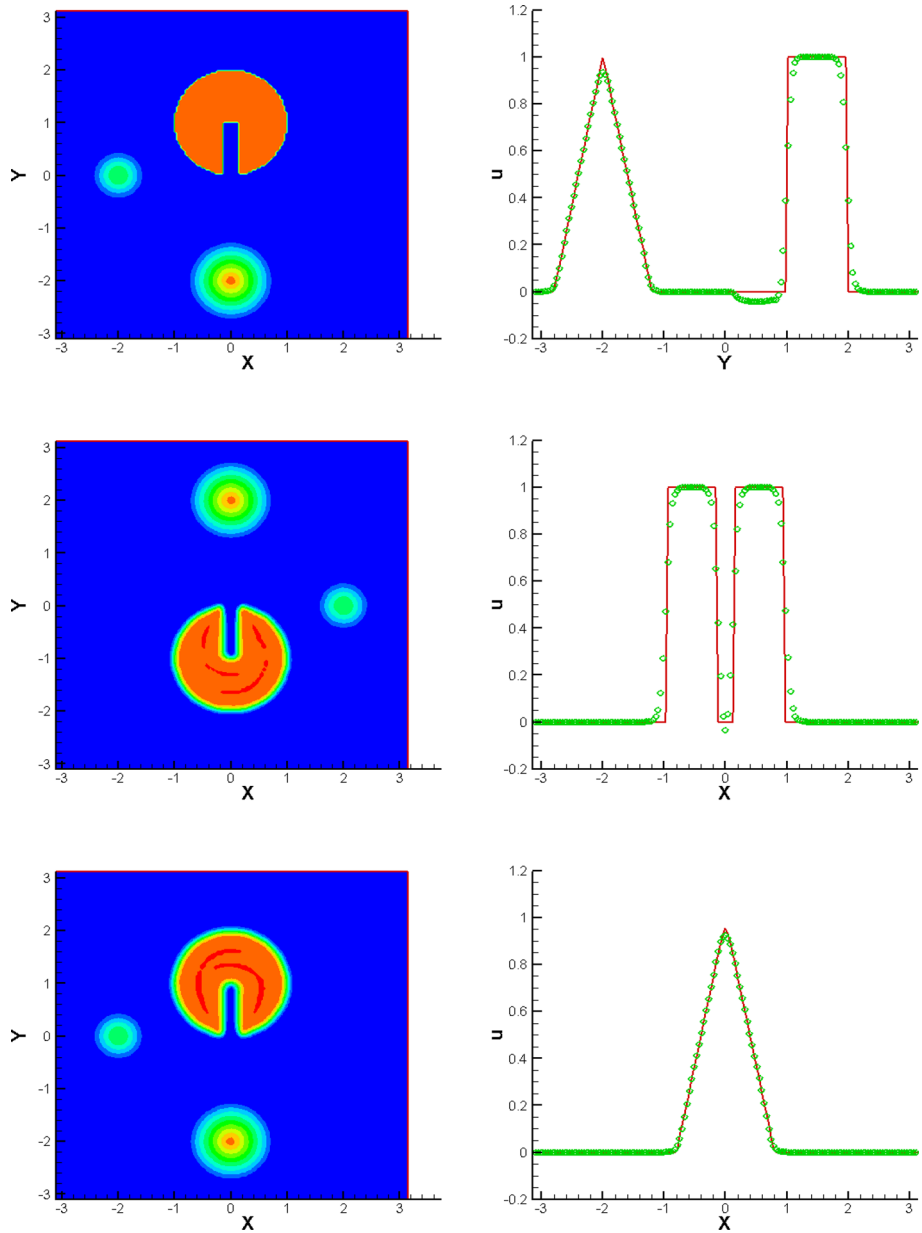


Fig. 6 Rigid body rotating problem. Mesh size: 128×128 . Left: from top to bottom, time at $T = 0, \pi, 2\pi$ respectively. Right: from top to bottom, cuts at $T = 2\pi$ along $x = 0, y = \pi/4$ and $y = -5\pi/8$

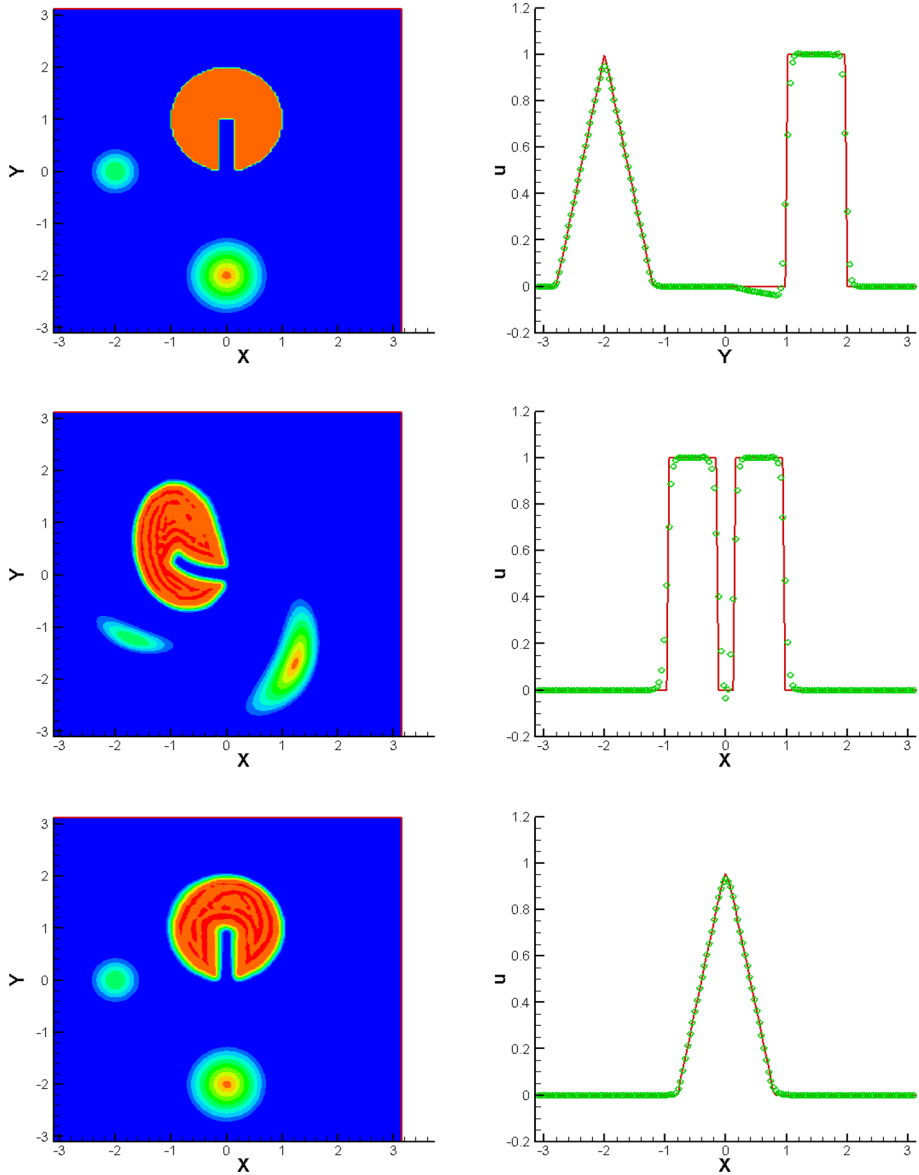


Fig. 7 Swirling deformation flow problem. Mesh size: 128×128 . Left: from top to bottom, time at $T = 0, 0.75, 1.5$ respectively. Right: from top to bottom, cuts at $T = 1.5$ along $x = 0, y = \pi/4$ and $y = -5\pi/8$

appearing in the Poisson equation (6.2). In this paper, we consider the VP system with 1-D in \mathbf{x} and 1-D in \mathbf{v} and simply denote as (x, v) the coordinates in the phase space. Periodic boundary condition is imposed in the x -direction, while zero boundary condition is imposed in v -direction. The equations for tracking characteristics are

$$\frac{dx}{dt} = v, \quad \frac{dv}{dt} = E, \tag{6.3}$$

where the electric field E is now a scalar function and nonlinearly depends on f via the Poisson system (6.2). To locate the foot of characteristics accurately, we apply the high order procedure proposed in [18].

Next we recall several norms in the VP system below, which should remain constant in time [10].

1. Mass:

$$\text{Mass} = \int_v \int_x f(x, v, t) dx dv.$$

2. L^p norm $1 \leq p < \infty$:

$$\|f\|_p = \left(\int_v \int_x |f(x, v, t)|^p dx dv \right)^{\frac{1}{p}}. \tag{6.4}$$

3. Energy:

$$\text{Energy} = \frac{1}{2} \int_v \int_x f(x, v, t) v^2 dx dv + \int_x E^2(x, t) dx, \tag{6.5}$$

where $E(x, t)$ is the electric field.

4. Entropy:

$$\text{Entropy} = \int_v \int_x f(x, v, t) \log(f(x, v, t)) dx dv. \tag{6.6}$$

Tracking relative deviations of these quantities numerically will be a good measure of the quality of numerical schemes. The relative deviation is defined to be the deviation away from the corresponding initial value divided by the magnitude of the initial value. We also check the mass conservation over time $\int_v \int_x f(x, v, t) dx dv$, which is the same as the L^1 norm if f is positive. However, since our scheme is not positivity preserving, the time evolution of the mass could be different from that of the L^1 norm due to the negative values appearing in numerical solutions.

In our numerical tests, we let the time step size

$$\Delta t = \text{CFL} \min(\Delta x/v_{max}, \Delta v/\max(E)),$$

where CFL is specified as 1.15, and let $v_{max} = 2\pi$ to minimize the error from truncating the domain in v -direction. The spatial domain will be $[0, L]$, where $L = \frac{2\pi}{k}$ with k being the wave number appeared in the initial condition specified in the examples. In all examples, we compare performance of conservative and non-conservative schemes, using the same set of computational mesh, by comparing the evolution of relative deviations of norms mentioned above.

Example 6.1 (Weak Landau damping) For the VP system, we first consider the weak Landau damping with the initial condition:

$$f(t = 0, x, v) = \frac{1}{\sqrt{2\pi}} (1 + \alpha \cos(kx)) \exp\left(-\frac{v^2}{2}\right), \tag{6.7}$$

where $\alpha = 0.01$ and $k = 0.5$. The length of the domain in the x -direction is $L = \frac{2\pi}{k} = 4\pi$.

We take mesh size 128×128 . In Fig. 8 (top left), we plot the time evolution of the electric field E in L^2 norm and L^∞ norm. The results are comparable to those in [17]. The relative derivation of mass, the discrete L^1 norm and L^2 norm, kinetic energy and entropy are plotted in Fig. 9, and are compared to the results from the nonconservative scheme

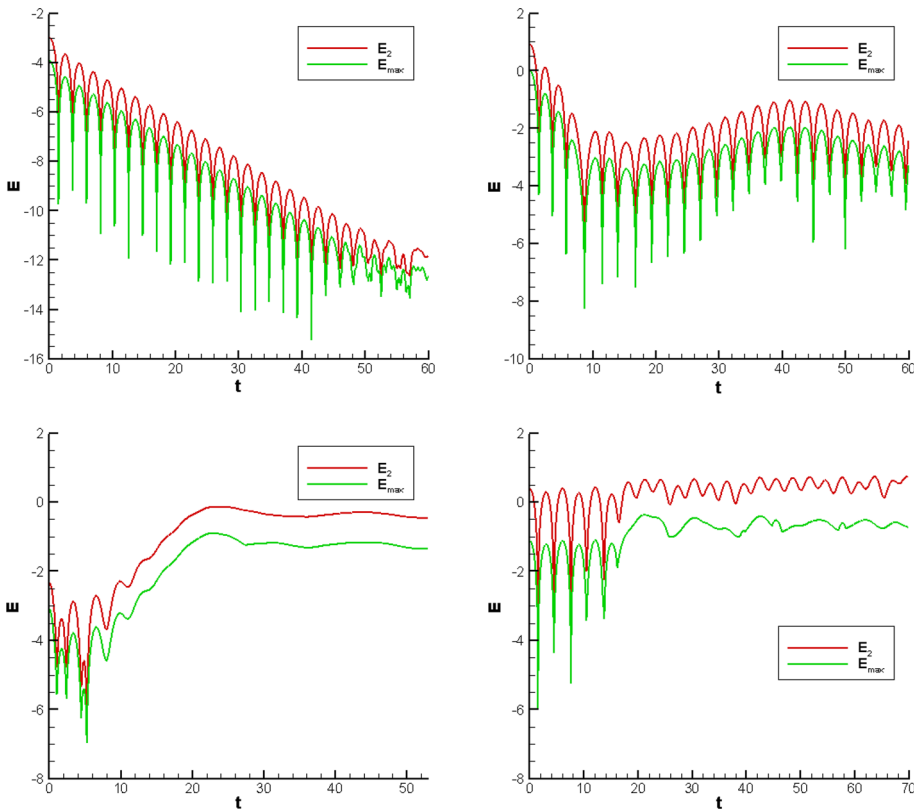


Fig. 8 Time evolution of the electric field in L^2 norm and L^∞ norm. Mesh: 128×128 . Top left: weak Landau damping; top right: strong Landau damping; bottom left: two stream instability; bottom right: symmetric two stream instability

[17]. The mass and L^1 norm will be the same when the distribution function $f(t, x, v)$ is always positive, otherwise they are different. We can observe that the conservative scheme “Cons-SL” preserves the mass exactly and also the L^1 norm, which is much better than the nonconservative scheme “NonCons-SL”. For the L^2 norm, energy and entropy, “Cons-SL” also preserves them better than “NonCons-SL”.

Example 6.2 (Strong Landau damping) The initial condition of strong Landau damping is (6.7), with $\alpha = 0.5$ and $k = 0.5$. Similarly we take mesh 128×128 and in Fig. 8 (top right), we plot the time evolution of the electric field in L^2 norm and L^∞ norm for “Cons-SL”. In Fig. 10, we compare the relative derivation of mass, the discrete L^1 norm and L^2 norm, kinetic energy and entropy of “Cons-SL” and “NonCons-SL” schemes. Except entropy, “Cons-SL” preserves those quantities better than “NonCons-SL”.

We also use this example to show the spatial and temporal accuracies of the conservative scheme proposed in this paper. In Tables 6 and 7, we can observe 5th order spatial accuracy and almost 4th order temporal accuracy respectively. Here the errors are computed by comparing to a reference solution on the mesh size 256×256 with $CFL = 0.01$. In order to measure the temporal accuracy, 9th order linear interpolation and 9th order linear reconstructions are used. As we can see that the temporal accuracy is almost one order higher than expected,

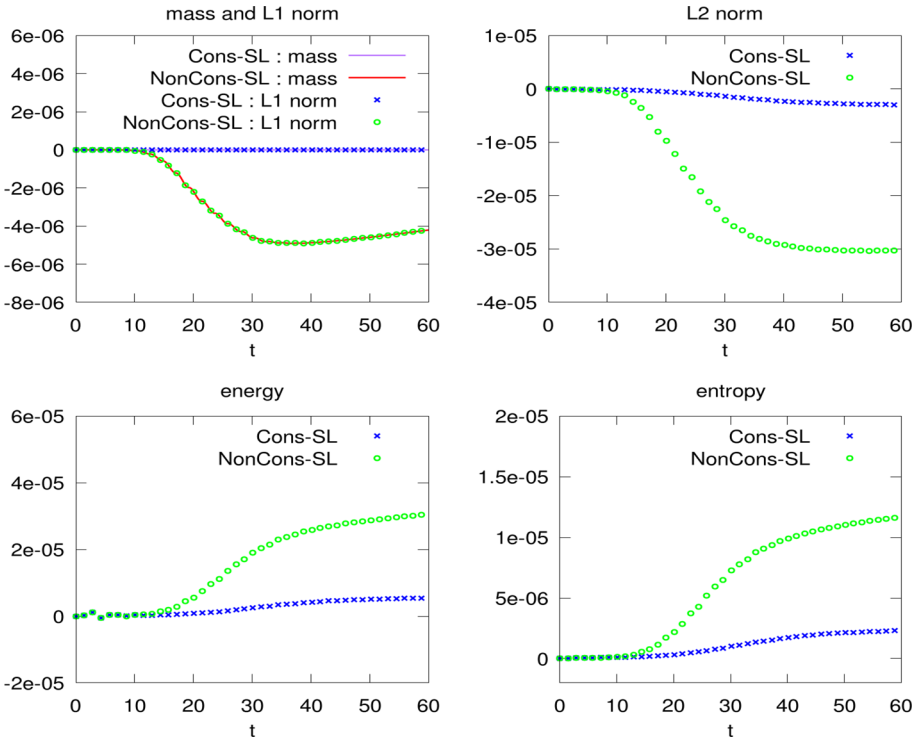


Fig. 9 Preservation of conserved quantities for weak Landau damping. Mesh: 128 × 128

that is we use a third order scheme in tracking characteristics proposed in [18] and observe higher than third order convergence in the table. Such phenomenon does not appear in the nonconservative case [18], which might be due to some special cancellation in the flux difference form.

Example 6.3 (Two stream instability) Now we consider the two stream instability problem, with an unstable initial distribution function given by:

$$\begin{aligned}
 f(t = 0, x, v) = & \frac{2}{7\sqrt{2\pi}}(1 + 5v^2)(1 + \alpha((\cos(2kx) + \cos(3kx))/1.2 \\
 & + \cos(kx)) \exp\left(-\frac{v^2}{2}\right)
 \end{aligned}
 \tag{6.8}$$

where $\alpha = 0.01$ and $k = 0.5$. We show the time evolution of electric field in L^2 norm and L^∞ norm in Fig. 8 (bottom left) with a mesh of 128×128 . We plot the numerical solution at $T = 53$ in Fig. 12 (left) with a mesh of 256×256 . The results are comparable to those in [17]. The relative derivation of mass, the discrete L^1 norm and L^2 norm, kinetic energy and entropy, as compared to the nonconservative scheme, are shown in Fig. 11. “Cons-SL” preserves those quantities better than “NonCons-SL”. Better performances in preserving norms are observed for “Cons-SL”.

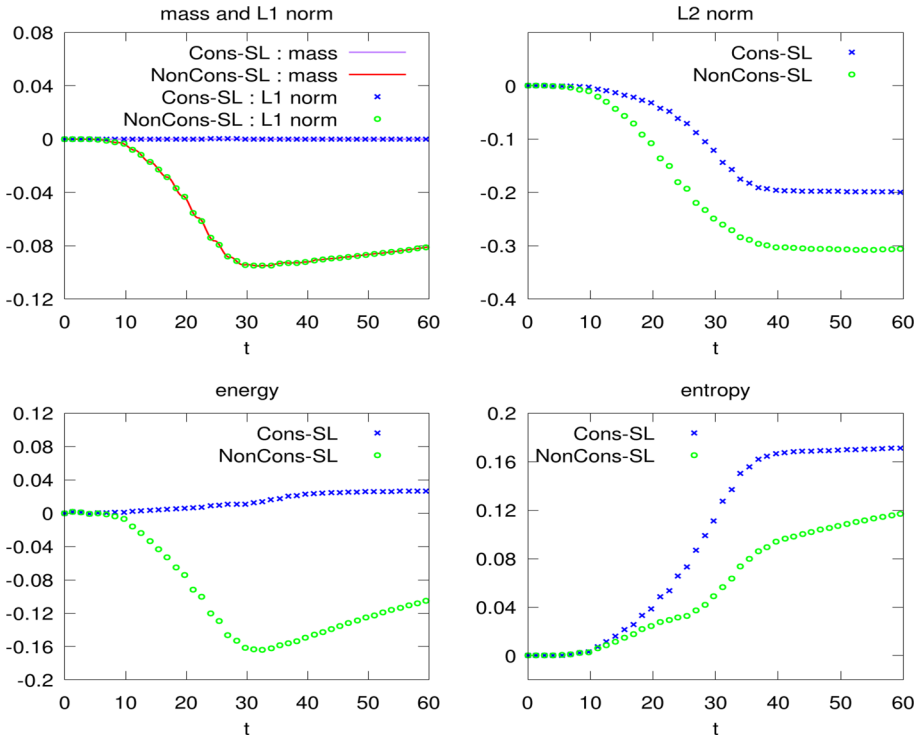


Fig. 10 Preservation of conserved quantities for strong Landau damping. Mesh: 128×128

Table 6 Errors and orders for strong Landau damping in Example 6.2

$N_x \times N_y$	16×16	32×32	64×64	128×128
L^1 error	1.90E-4	8.45E-6	2.91E-7	9.04E-9
Order	–	4.49	4.86	5.01

$T = 0.1$

Table 7 Errors and orders for strong Landau damping in Example 6.2

CFL	1.15	1.05	0.95	0.85
L^1 error	1.34E-11	9.74E-12	6.53E-12	4.26E-12
Order	–	3.49	3.99	3.84

$N_x = N_y = 256. T = 0.1$

Example 6.4 (Symmetric two stream instability) We consider the symmetric two stream instability with the initial condition:

$$f(t = 0, x, v) = \frac{1}{2v_{th}\sqrt{2\pi}} \left[\exp\left(-\frac{(v-u)^2}{2v_{th}^2}\right) + \exp\left(-\frac{(v+u)^2}{2v_{th}^2}\right) \right] (1 + \alpha \cos(kx)) \tag{6.9}$$

with $\alpha = 0.05, u = 0.99, v_{th} = 0.3, k = \frac{2}{13}$ and $L = 13\pi$. The time evolution of the electric field in L^2 norm and L^∞ norm are shown in Fig. 8 (bottom right). We plot the numerical solution with mesh size 256×256 at $T = 70$ in Fig. 12 (right). These results

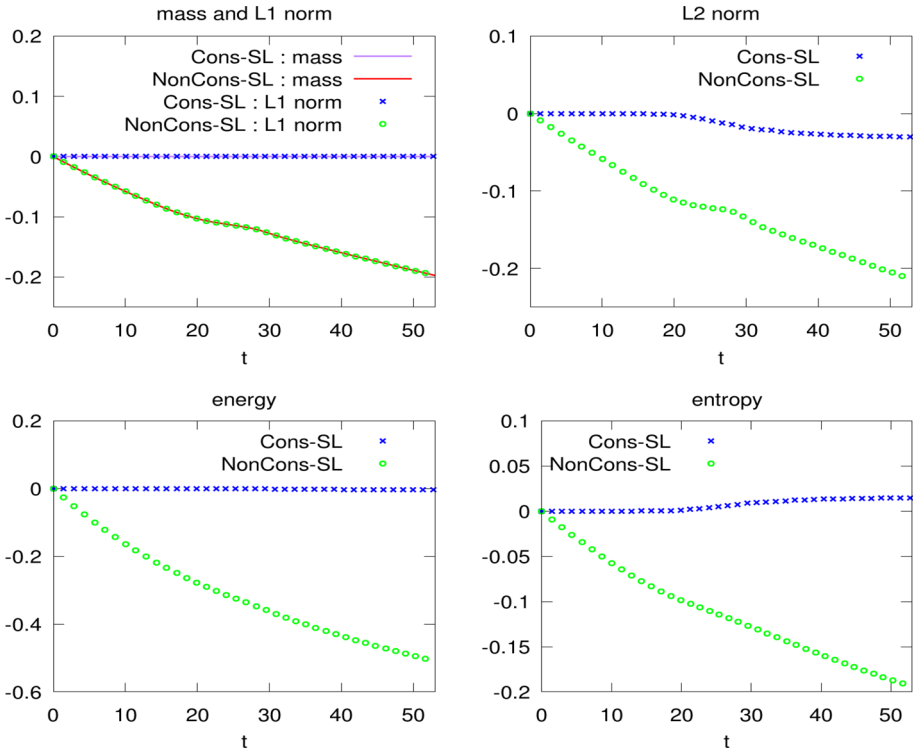


Fig. 11 Preservation of conserved quantities for two stream instability with the initial data (6.8). Mesh: 128×128

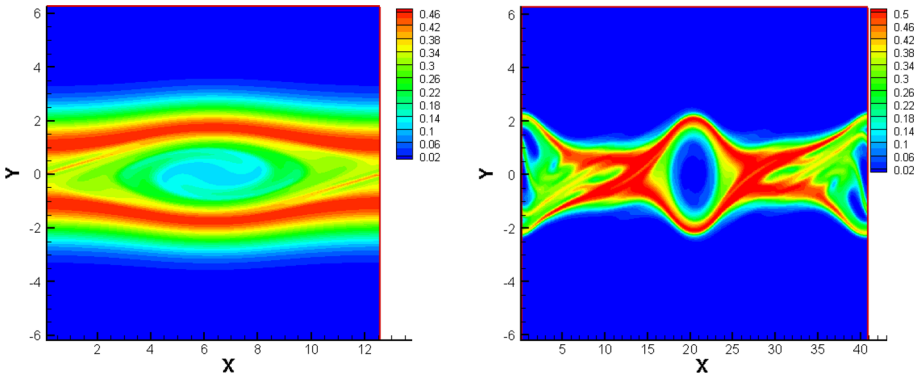


Fig. 12 Left: two stream instability at $T = 53$, mesh: 256×256 . Right: symmetric two stream instability at $T = 70$, mesh: 256×256

are similar to those in [17]. The relative derivation of mass, the discrete L^1 norm and L^2 norm, kinetic energy and entropy, comparing ‘Cons-SL’ and ‘NonCons-SL’, are reported in Fig. 13. As observed in previous examples, ‘Cons-SL’ preserves these quantities better than ‘NonCons-SL’.

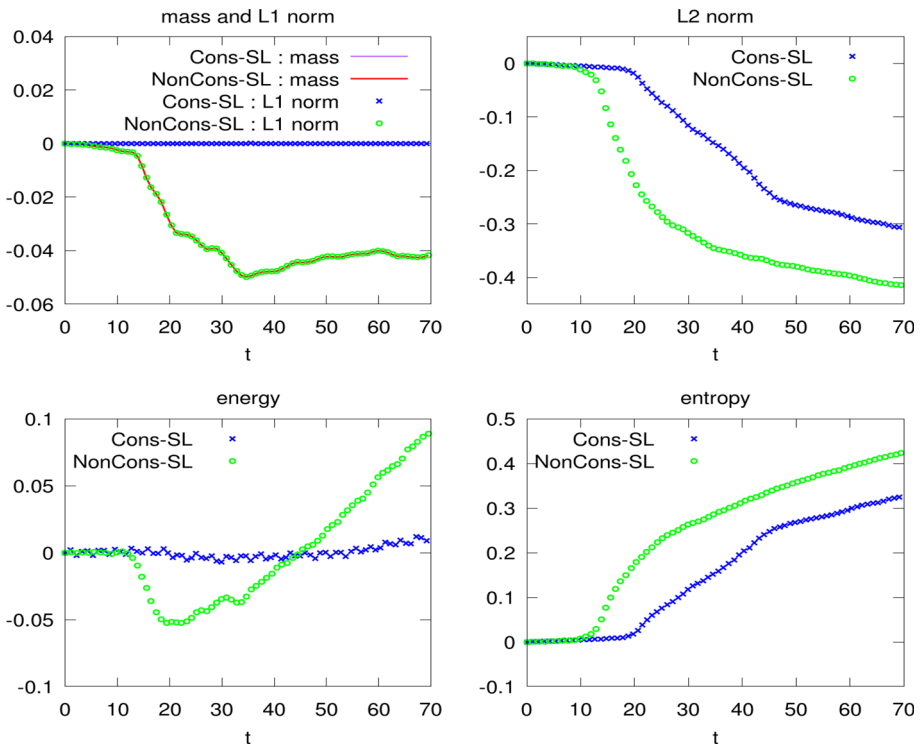


Fig. 13 Preservation of conserved quantities for symmetric two stream instability. Mesh: 128×128

6.2 The Guiding Center Vlasov Model

The guiding center model is an approximation of the 2D Vlasov model [9,30] under a strong constant external magnetic field, which can be written as

$$\frac{\partial \rho}{\partial t} + E_2 \frac{\partial \rho}{\partial x} - E_1 \frac{\partial \rho}{\partial y} = 0, \tag{6.10}$$

or equivalently in a conservative form as

$$\frac{\partial}{\partial t} \rho + \frac{\partial}{\partial x} (\rho E_2) + \frac{\partial}{\partial y} (-\rho E_1) = 0, \tag{6.11}$$

where $\mathbf{E} = (E_1, E_2) = -\nabla \Phi$ with Φ determined from the Poisson’s equation

$$\Delta \Phi = -\rho.$$

It is well-known that the mass, energy and entropy are preserved over time in the PDE level.

1. Mass:

$$\text{Mass} = \int_y \int_x \rho(x, y, t) dx dy. \tag{6.12}$$

2. Energy:

$$\|\mathbf{E}\|_2^2 = \int_y \int_x \|\mathbf{E}\|^2 dx dy. \tag{6.13}$$

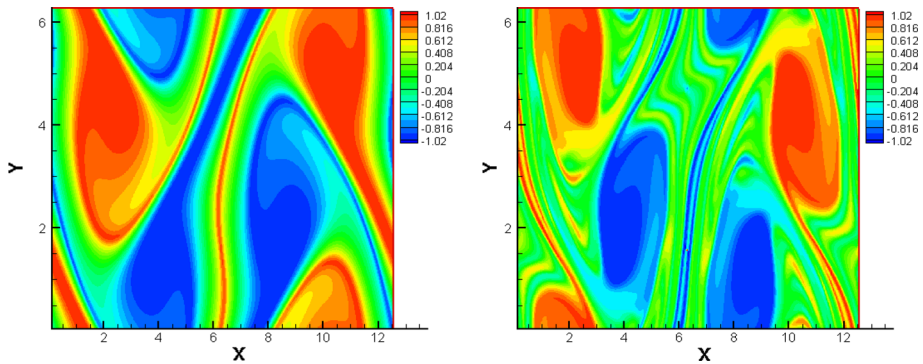


Fig. 14 Kelvin–Helmholtz instability problem. Mesh: 256×256 . $T = 20$ (left), $T = 40$ (right)

3. Entropy:

$$\|\rho\|_2^2 = \int_y \int_x |\rho(x, y, t)|^2 dx dy. \tag{6.14}$$

Tracking relative deviations of these quantities numerically will be a good measure of the quality of numerical schemes. In a SL framework, we adopt a third order high order characteristic tracing scheme as proposed in our recent work [28]. For this problem, the time step size is taken as

$$\Delta t = \text{CFL} \min(\Delta x / \max(E_2), \Delta y / \max(E_1)),$$

with CFL specified as 1.15.

Example 6.5 (Kelvin–Helmholtz instability problem) This example is the 2-D guiding center model problem with the initial condition

$$\rho_0(x, y) = \sin(y) + 0.015 \cos(kx) \tag{6.15}$$

and periodic boundary conditions on the domain $[0, 4\pi] \times [0, 2\pi]$. We let $k = 0.5$, which creates a Kelvin–Helmholtz instability [22].

For this example, we take mesh size 256×256 and show the surface of the solution at $T = 20$ and 40 in Fig. 14; the solution profiles are comparable to those previously presented in other papers [9,28]. Then we take mesh size 128×128 and show the relative derivation of mass, energy $\|\mathbf{E}\|_2^2$ and entropy $\|\rho\|_2^2$ in Fig. 15. We compare them to the nonconservative scheme [28] with 6th order space reconstruction and $\text{CFL} = 1.15$. For this example, “Cons-SL” has exact mass conservation, while “NonCons-SL” does not. For the energy and entropy, “NonCons-SL” seems to perform a little better.

6.3 Incompressible Euler Equation

The incompressible Euler equations in the vorticity-stream function formulation in two space dimensions reads

$$\omega_t + \nabla \cdot (\mathbf{u} \omega) = 0, \quad \Delta \psi = \omega, \quad \mathbf{u} = \left(-\frac{\partial \psi}{\partial y}, \frac{\partial \psi}{\partial x} \right), \tag{6.16}$$

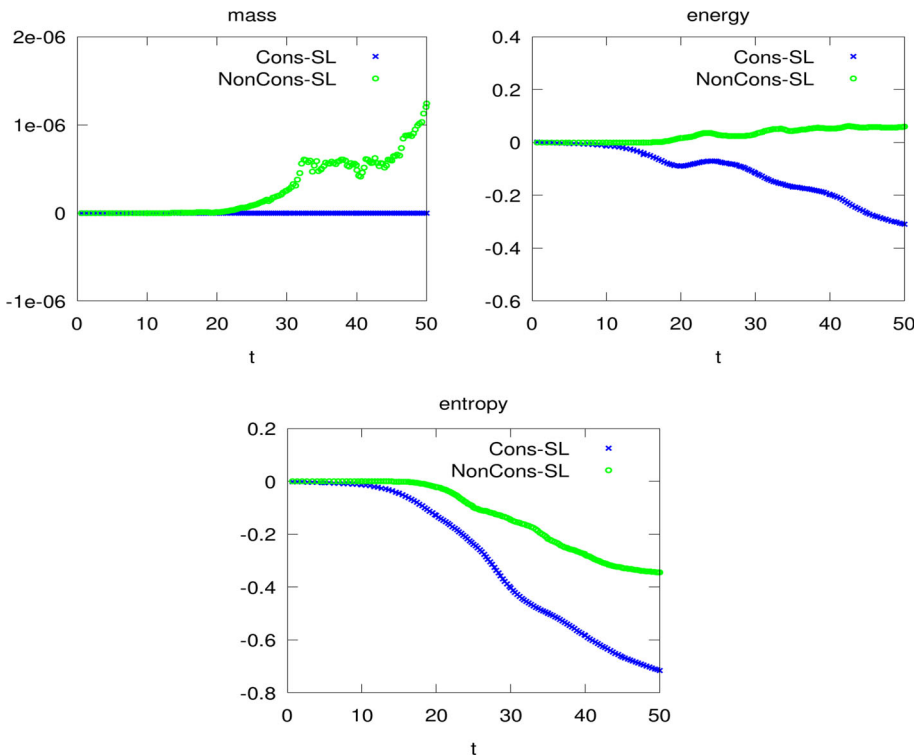


Fig. 15 Preservation of conserved quantities for Kelvin–Helmholtz instability problem. Mesh: 128×128

where $\mathbf{u} = (u_1, u_2)$ is the velocity vector, $\omega = \frac{\partial u_2}{\partial x} - \frac{\partial u_1}{\partial y}$ is the vorticity and ψ is a stream function. The form is almost equivalent as the guiding center Vlasov model, up to a sign difference for the field equation [28]. The conservation of mass, energy and entropy shown in Eqs. (6.12)–(6.14) is the same as the guiding center model. We use the third order characteristics tracing scheme, similar to that in the guiding center model, as presented in our previous work [28]. Similar to the guiding center problem, the time step size is taken as

$$\Delta t = \text{CFL} \min(\Delta x / \max(u_1), \Delta y / \max(u_2)),$$

with CFL to be 1.15.

Example 6.6 We first consider the incompressible Euler system on the domain $[0, 2\pi] \times [0, 2\pi]$ with an initial condition $\omega_0(x, y) = -2 \sin(x) \sin(y)$. The exact solution will stay stationary with $\omega(x, y, t) = -2 \sin(x) \sin(y)$. Similarly as in Tables 6 and 7, the 5th order spatial accuracy and 4th order temporal accuracy are clearly observed in Tables 8 and 9 respectively. Here in Table 9, in order to measure the temporal convergence, we use a 9th order linear interpolation and 9th order linear reconstruction in space to minimize errors from spatial approximations.

Table 8 Errors and orders for the incompressible Euler equation in Example 6.6

$N_x \times N_y$	32×32	64×64	128×128	256×256
L^1 error	6.14E-4	2.10E-5	6.61E-7	2.04E-8
Order	–	4.87	4.99	5.02

$T = 1.2$

Table 9 Errors and orders for the incompressible Euler equation in Example 6.6

CFL	1.15	1.05	0.95	0.85
L^1 error	1.06E-12	7.39E-13	4.97E-13	3.21E-13
Order	–	3.96	3.95	3.94

$N_x = N_y = 256. T = 1$

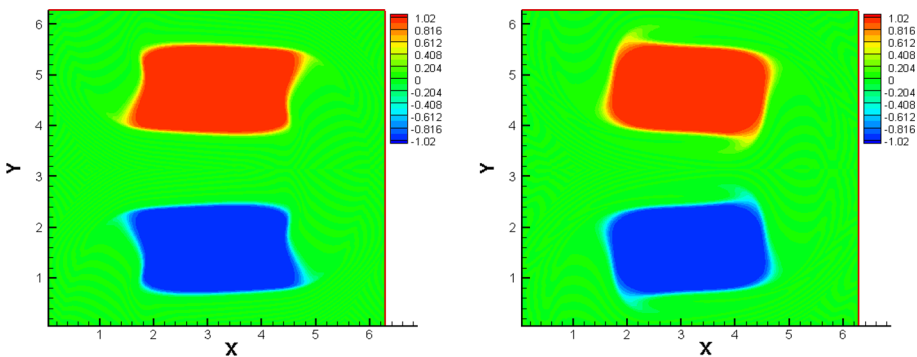


Fig. 16 Vortex patch problem. Mesh: 256×256 . $T = 5$ (left) and $T = 10$ (right)

Example 6.7 (The vortex patch problem) In this example, we consider the incompressible Euler equations with the initial condition given by

$$\omega_0(x, y) = \begin{cases} -1, & \frac{\pi}{2} \leq x \leq \frac{3\pi}{2}, \frac{\pi}{4} \leq y \leq \frac{3\pi}{4}; \\ 1, & \frac{\pi}{2} \leq x \leq \frac{3\pi}{2}, \frac{5\pi}{4} \leq y \leq \frac{7\pi}{4}; \\ 0, & \text{otherwise.} \end{cases} \tag{6.17}$$

We show the surface of ω at $T = 5$ and 10 in Fig. 16. The mesh size is 256×256 . The results are similar to those in [19]. Then we take mesh size 128×128 and show the relative derivation of mass, energy $\|\mathbf{E}\|_2^2$ and entropy $\|\rho\|_2^2$ in Fig. 15 (left). We compare them to the nonconservative scheme [28] with 6th space reconstruction and $CFL = 1.15$. ‘‘Cons-SL’’ has exact mass conservation, while ‘‘NonCons-SL’’ does not. For the energy and entropy, ‘‘Cons-SL’’ performs better than ‘‘NonCons-SL’’.

Example 6.8 (Shear flow problem) This example is the same as above but with following initial conditions (Fig. 17)

$$\omega_0(x, y) = \begin{cases} \delta \cos(x) - \frac{1}{\rho} \operatorname{sech}^2((y - \pi/2)/\rho)^2, & y \leq \pi; \\ \delta \cos(x) + \frac{1}{\rho} \operatorname{sech}^2((3\pi/2 - y)/\rho)^2, & y > \pi. \end{cases} \tag{6.18}$$

where $\delta = 0.05$ and $\rho = \frac{\pi}{15}$. We show the surface of ω at $T = 6$ and 8 in Fig. 18. The mesh size is 256×256 . The results are also similar to those in [19], which can also show our

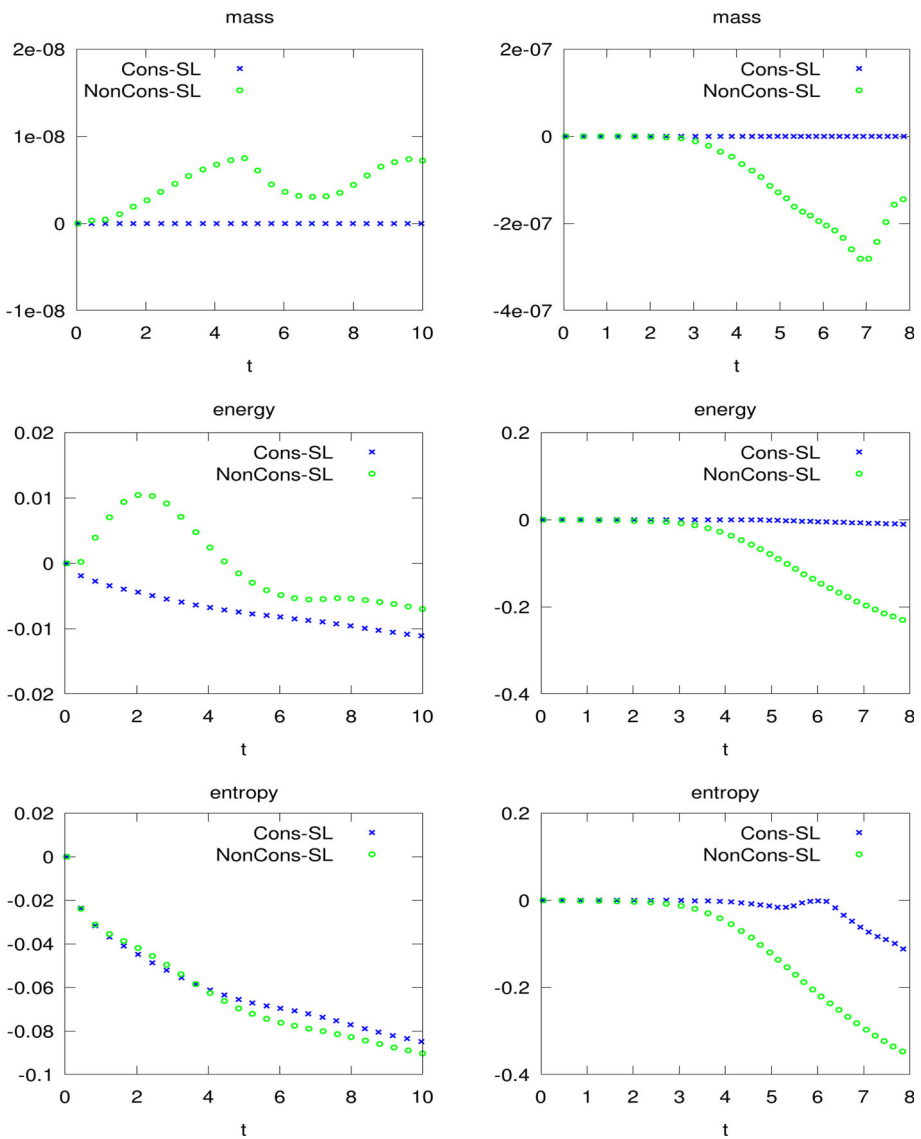


Fig. 17 Preservation of conserved quantities for the incompressible flow system. Mesh: 128×128 . Left: vortex patch; right: shear flow

multi-dimensional non-splitting conservative scheme can perform well. Then we take mesh size 128×128 and show the relative deviation of mass, energy $\|\mathbf{E}\|_2^2$ and entropy $\|\rho\|_2^2$ in Fig. 15 (right). We compare them to the nonconservative scheme. Similar to the vortex patch problem, “Cons-SL” has exact mass conservation, while “NonCons-SL” does not. For the energy and entropy conservation, “Cons-SL” performs better than “NonCons-SL”.

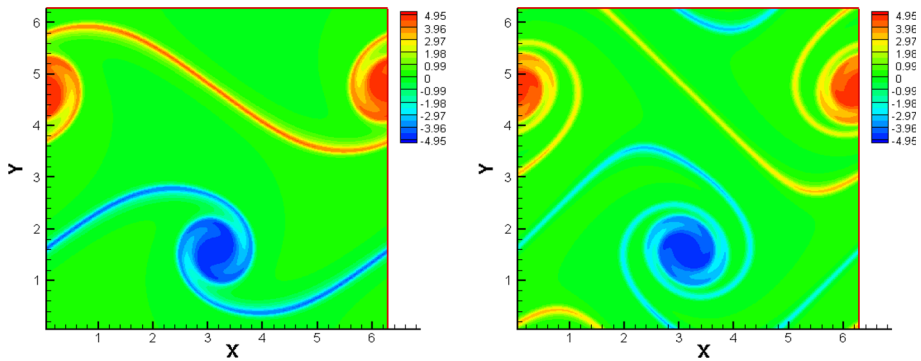


Fig. 18 Shear flow problem. Mesh: 256×256 . $T = 6$ (left) and $T = 8$ (right)

7 Conclusion

In this paper, we propose a *mass conservative* semi-Lagrangian finite difference WENO scheme based on a flux difference formulation. We investigate its numerical stability from the linear ODE and PDE point of view via Fourier analysis. The upper bound of time step constraints have been found in the linear setting and have been numerically verified. These upper bounds are only slightly greater than those from the Eulerian approach, unfortunately. The schemes are applied to passive transport problems as well as nonlinear Vlasov systems and the incompressible Euler system to showcase its effectiveness. The properties of the schemes are verified numerically. In several circumstances the conservative scheme provides more accurate solutions than the corresponding non conservative counterpart with the same discretization parameters. The development of conservative, large time step, multi-dimensional semi-Lagrangian scheme is currently under investigation. If one is interested in using semi-Lagrangian schemes with mass conservation, and with stability for large CFL numbers, the discontinuous Galerkin method would be a good framework to use [1,2].


Acknowledgements The final draft of this paper is completed, while the authors were in residence at the Institute for Computational and Experimental Research in Mathematics in Providence, RI, during the Collaboration@ICERM program, supported by the National Science Foundation under Grant No. DMS-1439786.

References

1. Cai, X., Guo, W., Qiu, J.-M.: A high order conservative semi-Lagrangian discontinuous Galerkin method for two-dimensional transport simulations. *J. Sci. Comput.* **73**, 514–542 (2017)
2. Cai, X., Guo, W., Qiu, J.-M.: A high order semi-Lagrangian discontinuous Galerkin method for Vlasov–Poisson simulations without operator splitting. *J. Comput. Phys.* **354**, 529–551 (2018)
3. Carrillo, J.A., Vecil, F.: Nonoscillatory interpolation methods applied to Vlasov-based models. *SIAM J. Sci. Comput.* **29**, 1179–1206 (2007)
4. Cheng, C.-Z., Knorr, G.: The integration of the Vlasov equation in configuration space. *J. Comput. Phys.* **22**, 330–351 (1976)
5. Christlieb, A., Guo, W., Morton, M., Qiu, J.-M.: A high order time splitting method based on integral deferred correction for semi-Lagrangian Vlasov simulations. *J. Comput. Phys.* **267**, 7–27 (2014)
6. Crouseilles, N., Mehrenberger, M., Sonnendrücker, E.: Conservative semi-Lagrangian schemes for Vlasov equations. *J. Comput. Phys.* **229**, 1927–1953 (2010)
7. Enright, D., Losasso, F., Fedkiw, R.: A fast and accurate semi-Lagrangian particle level set method. *Comput. Struct.* **83**, 479–490 (2005)

8. Filbet, F., Sonnendrücker, E., Bertrand, P.: Conservative numerical schemes for the Vlasov equation. *J. Comput. Phys.* **172**, 166–187 (2001)
9. Frenod, E., Hirstoaga, S.A., Lutz, M., Sonnendrücker, E.: Long time behaviour of an exponential integrator for a Vlasov–Poisson system with strong magnetic field. *Commun. Comput. Phys.* **18**, 263–296 (2015)
10. Glassey, R.T.: *The Cauchy Problem in Kinetic Theory*, vol. 52. SIAM, Philadelphia (1996)
11. Guo, W., Nair, R.D., Qiu, J.-M.: A conservative semi-Lagrangian discontinuous Galerkin scheme on the cubed sphere. *Mon. Weather Rev.* **142**, 457–475 (2014)
12. Guo, W., Qiu, J.-M.: Hybrid semi-Lagrangian finite element-finite difference methods for the Vlasov equation. *J. Comput. Phys.* **234**, 108–132 (2013)
13. Huot, F., Ghizzo, A., Bertrand, P., Sonnendrücker, E., Coulaud, O.: Instability of the time splitting scheme for the one-dimensional and relativistic Vlasov–Maxwell system. *J. Comput. Phys.* **185**, 512–531 (2003)
14. Lin, S.-J., Rood, R.B.: Multidimensional flux-form semi-Lagrangian transport schemes. *Mon. Weather Rev.* **124**, 2046–2070 (1996)
15. Morton, K., Priestley, A., Suli, E.: Stability of the Lagrange–Galerkin method with non-exact integration. *RAIRO-Modélisation mathématique et analyse numérique* **22**, 625–653 (1988)
16. Pironneau, O.: On the transport-diffusion algorithm and its applications to the Navier–Stokes equations. *Numerische Mathematik* **38**, 309–332 (1982)
17. Qiu, J.-M., Christlieb, A.: A conservative high order semi-Lagrangian WENO method for the Vlasov equation. *J. Comput. Phys.* **229**, 1130–1149 (2010)
18. Qiu, J.-M., Russo, G.: A high order multi-dimensional characteristic tracing strategy for the Vlasov–Poisson system. *J. Sci. Comput.* **71**, 414–434 (2017)
19. Qiu, J.-M., Shu, C.-W.: Conservative high order semi-Lagrangian finite difference WENO methods for advection in incompressible flow. *J. Comput. Phys.* **230**, 863–889 (2011)
20. Qiu, J.-M., Shu, C.-W.: Positivity preserving semi-Lagrangian discontinuous Galerkin formulation: theoretical analysis and application to the Vlasov–Poisson system. *J. Comput. Phys.* **230**, 8386–8409 (2011)
21. Rossmanith, J.A., Seal, D.C.: A positivity-preserving high-order semi-Lagrangian discontinuous Galerkin scheme for the Vlasov–Poisson equations. *J. Comput. Phys.* **230**, 6203–6232 (2011)
22. Shoucri, M.M.: A two-level implicit scheme for the numerical solution of the linearized vorticity equation. *Int. J. Numer. Methods Eng.* **17**, 1525–1538 (1981)
23. Shu, C.-W.: High order weighted essentially nonoscillatory schemes for convection dominated problems. *SIAM Rev.* **51**, 82–126 (2009)
24. Shu, C.-W., Osher, S.: Efficient implementation of essentially non-oscillatory shock-capturing schemes. *J. Comput. Phys.* **77**(2), 439–471 (1988)
25. Sonnendrücker, E., Roche, J., Bertrand, P., Ghizzo, A.: The semi-Lagrangian method for the numerical resolution of the Vlasov equation. *J. Comput. Phys.* **149**, 201–220 (1999)
26. Staniforth, A., Côté, J.: Semi-Lagrangian integration schemes for atmospheric models—a review. *Mon. Weather Rev.* **119**, 2206–2223 (1991)
27. Strain, J.: Semi-Lagrangian methods for level set equations. *J. Comput. Phys.* **151**, 498–533 (1999)
28. Xiong, T., Qiu, J.-M., Russo, G.: High order multi-dimensional characteristic tracing for the incompressible Euler equation and the guiding-center Vlasov equation. *J. Sci. Comput.* (2018). <https://doi.org/10.1007/s10915-018-0705-y>
29. Xiu, D., Karniadakis, G.E.: A semi-Lagrangian high-order method for Navier–Stokes equations. *J. Comput. Phys.* **172**, 658–684 (2001)
30. Yang, C., Filbet, F.: Conservative and non-conservative methods based on hermite weighted essentially non-oscillatory reconstruction for Vlasov equations. *J. Comput. Phys.* **279**, 18–36 (2014)

Affiliations

Tao Xiong¹ · Giovanni Russo² · Jing-Mei Qiu³ 

Tao Xiong
txiong@xmu.edu.cn

Giovanni Russo
russo@dmi.unict.it

¹ School of Mathematical Sciences, Fujian Provincial Key Laboratory of Mathematical Modeling and High-Performance Scientific Computing, Xiamen University, Xiamen 361005, Fujian, People's Republic of China

² Department of Mathematics and Computer Science, University of Catania, 95125 Catania, Italy

³ Department of Mathematical Science, University of Delaware, Newark, DE 19716, USA