Observable Error in Mass, Momentum, and Energy Conservative Dynamical Low-rank Integration Schemes for the Vlasov-Poisson Equation

April Herwig

February 19, 2025

Abstract

Vlasov-Poisson has (infinitely) many conserved quantities. Low-rank tensor approximation [5]. Does this conserve more observables than we expect?

Contents

1 Introduction

Efficient numerical solutions to kinetic equations is the basis for many fields in physics and has applications such as fusion energy generation in magnetically confined chambers, or dosage calculation in radiation therapy. Kinetic models are dependent on both physical space coordinates $x \in \Omega_x \subset \mathbb{R}^d$, $d = 1, 2, 3$ and velocity coordinates $v \in \Omega_v \subset \mathbb{R}^d$. A grid-based discretization with n grid points in each dimension therefore incurs a computational cost of order $O(n^{2d})$. Hence, direct solutions for the (kinetic) density function are prohibitively expensive.

Recently, dynamical low-rank solvers for kinetic equations such as the Vlasov-Poisson equation have been proposed [4]. These methods are based on [7], in which matrix differential equations are approximated by low-rank matrices using a Galerkin principle. These solvers have been shown to be successful [1], but without additional care the projection to low-rank tensor products destroys important physical structure in the system. This can be alleviated by altering the approximation space such that some moments of the velocity coordinate always lie in the space spanned by the low-rank matrices [5]. This is sufficient to guarantee that continuity equations for mass, momentum, and energy are conserved up to an error that is linear in the time step size.

In the following report we examine whether the conservative dynamical low-rank integrator also preserves further observables and if so, to what order. The report is structured as follows: in section 2 the Vlasov-Poisson equation is introduced and conservation properties are discussed; in section 3 the conservative dynamical low-rank integrator is shown; in section 4 the integrator is demonstrated on an analytically understood example system.

if some observable works, mention here

2 The Vlasov-Poisson Equation and Conservation Properties

The Vlasov-Poisson equation models the evolution of the electron density $f = f(x, v, t)$ in plasma. It is assumed that (1) slow movement of heavy positive ions, (2) pair collisions between electrons, (3) variations in magnetic field, and (4) relativistic effects are all negligible. The resultant (non-dimensionalized) evolution equation reads

$$\partial_t f(x, v, t) + v \cdot \nabla_x f(x, v, t) + E(x, t) \cdot \nabla_v f(x, v, t) = 0 \tag{2.1}$$

where $E = E(f)$ is the self-consistent electric field. The mass

$$\mathcal{M}_0(t) = \int_{\Omega_x} \int_{\Omega_v} f(x, v, t) \; dv \, dx \tag{2.2}$$

is wlog. assumed to be initially equal to 1. Moreover E is written as a potential $E = -\nabla\phi$ where

$$-\Delta\phi\,(x) = 1 - \int_{\Omega_v} f(x, v, t) \; dv. \tag{2.3}$$

Equation 2.1 contains infinitely many invariants. Indeed, all L^p norms $\|f\|_p$ as well as entropy

$$\mathcal{S}(t) = \iint_{\Omega_x \times \Omega_v} f(x, v, t) \;\; \log f(x, v, t) \; dv \, dx \tag{2.4}$$

are conserved [**<empty citation>**]. Additionally it can be shown [**<empty citation>**] that all moments of the velocity distrbution

cite entropy conservation

cite moment equations

$$M_k(x, t) = \int_{\Omega_v} vv \dots v f(x, v, t) \; dv \tag{2.5}$$

(wich are k-tensors) satisfy continuity equations and correspondingly their space-integrals $\mathcal{M}_k = \int_{\Omega_x} M_k \, dx$ are also conserved. Notable are the first few moments and their continuity equations

$$\partial_t M_0 + \nabla_x \cdot M_1 = 0, \tag{2.6}$$
$$\partial_t M_1 + \nabla_x \cdot M_2 = -E \, M_0, \tag{2.7}$$
$$\partial_t M_2 + \nabla_x \cdot M_3 = -2E \otimes M_1 \tag{2.8}$$

Verify this

which imply conservation of mass $\mathcal{M}_0$, momentum $\mathcal{M}_1$, and (kinetic) energy $\mathrm{tr}\mathcal{M}_2$ after integrating over Ω_x. The continuity equation for (total) energy density is then

$$\partial_t e + \frac{1}{2}\nabla_x \cdot (\mathrm{tr}M_3) = (\partial_t E - M_1) \cdot E \tag{2.9}$$

where

$$e(x,t) = \frac{1}{2}\mathrm{tr}M_2(x,t) + \frac{1}{2}E(x,t)^2 = \frac{1}{2}\int_{\Omega_v} v^2 f(x,v,t)\ dv + \frac{1}{2}E(x,t)^2 \tag{2.10}$$

and

$$\mathrm{tr}M_3(x,t) = \int_{\Omega_v} v^2\ v f(x,v,t)\ dv. \tag{2.11}$$

this is trace of a 3-tensor ie. tensor contraction - should maybe simplify?

The dynamical low-rank integrator presented in the following section is designed to satisfy equations 2.6, 2.7, and 2.9.

3 A Low-rank Tensor Approximation Scheme

The density function $f(x,v,t) \in L^2(\Omega_x \times \Omega_v)$ is approximated by a tensor product of functions in x and in v:

$$f(x,v,t) = f_0(v) \sum_{i,j=1}^{r} X_i(x,t)\ S_{ij}(t)\ V_j(v,t) =: f_0(v)\ X(x,t)^T\, S(t)\, V(v,t) \tag{3.1}$$

where r is the approximation rank and $f_0(v) = \exp(-v^2)$ is a Gaussian weight.

Let

$$\langle g,h\rangle_x = \int_{\Omega_x} f\ g\ dx, \quad (g,h)_v = \int_{\Omega_v} g\ h\ dv \quad \langle g,h\rangle_v = \int_{\Omega_v} g\ h\ f_0\ dv. \tag{3.2}$$

We require that X and V satisfy the orthonormality conditions

$$\langle X_i,\ X_j\rangle_x = \delta_{ij}, \quad \langle V_i,\ V_j\rangle_v = \delta_{ij}, \quad 1 \le i,j \le r \tag{3.3}$$

where δ_{ij} refers the the Kronecker delta, as well as the gauge conditions

$$\langle \partial_t X_i,\ X_j\rangle_x = 0, \quad \langle \partial_t V_i,\ V_j\rangle_v = 0, \quad 1 \le i,j \le r. \tag{3.4}$$

Let further $\bar{X} = \mathrm{span}\,\{X_1,\dots,X_r\}$, $\bar{V} = \mathrm{span}\,\{V_1,\dots,V_r\}$. The Galerkin condition yields the equation [4]:

$$\partial_t f = \Pi\,(\,RHS(f)\,) \quad \text{where} \quad \Pi\, g = \Pi_{\bar{V}}\, g - \Pi_{\bar{X}}\Pi_{\bar{V}}\, g + \Pi_{\bar{X}}\, g \tag{3.5}$$

and $\Pi_{\bar{X}}$, $\Pi_{\bar{V}}$ are the orthogonal projections onto $\bar{X}$ and $\bar{V}$, respectively. A first-order Lie-Trotter splitting based on the three terms in equation 3.5 yields equations of motion for the components of X, S, and V.

The key insight of [3] is that if the functions $v \mapsto 1$, $v \mapsto v_1, \ldots, v \mapsto v_d$, and $v \mapsto v^2$ lie in $\bar{V}$, then discrete versions of equations 2.6, 2.7, and 2.9 hold. Hence the integration scheme is altered to guarantee this condition. We split $V(v,t) \in \mathbb{R}^r$ into two blocks $U(v,t) \in \mathbb{R}^m$ and $W \in \mathbb{R}^{r-m}$:

$$V = \begin{bmatrix} U \\ W \end{bmatrix} \tag{3.6}$$

where U is fixed throughout the integration and contains the desired functions

$$U(v,t) = U(v) = \begin{bmatrix} 1 \\ v_1 \\ \vdots \\ v_d \\ v^2 \end{bmatrix}. \tag{3.7}$$

Performing the analogous calculations as in [4] for the altered basis functions yields [3]

$$\sum_i \partial_t X_i \, S_{ik} = (\, V_k,\, RHS(f)\,)_v - \sum_i X_i \, \partial_t S_{ik}, \qquad 1 \le k \le r \tag{3.8}$$

$$\sum_{ip} S_{i\,q+m} \, S_{ip} \, \partial_t W_p = \frac{1}{f_0} \sum_i S_{i\,q+m} \, \langle\, X_i,\, RHS(f)\,\rangle_x - \sum_{il} S_{i\,q+m} \, \partial_t S_{il} \, V_l, \qquad 1 \le q \le r-m \tag{3.9}$$

$$\partial_t S_{kl} = (\, X_k,\, (V_l,\, RHS(f))_v\,)_x \qquad 1 \le k,l \le r. \tag{3.10}$$

Concrete equations for computing the above inner products are given in [5]. It should be reemphasized that the orthonormality and guage conditions 3.3, 3.4 must still hold, so the components of U must be appropriately scaled.

To solve equations 3.8, 3.9 directly, the matrix S must be inverted. However, as the approximation rank r increases, S has progressively smaller singular values. Hence the scheme becomes increasingly ill-conditioned as the accuracy increases. We therefore need to alter the low-rank scheme again to address this issue.

Notice that the low-rank approximation $f = X^T S V$ can be written as $f = K^T V$ for some $K = K(x,t) \in \mathbb{R}^r$. X^T and S are then (up to a unitary basis transformation) the result of a (semidiscrete[1]) QR factorization of K^T. Analogously, f can be written as $f = X^T L$ and the elements W can also be reconstructed by a QR factorization. We may therefore rewrite equations 3.8 and 3.9 as

[1]The standard Gram-Schmidt process to construct a QR decomposition can just as easily be viewed in a semidiscrete setting: each "column" of K^T is a function of x evaluated across Ω_x, instead of each column being a discrete vector. The algorithm does not need to be changed at all.

$$\partial_t K_k = (\, V_k,\, RHS(f)\,)_v, \qquad 1 \le k \le r \tag{3.11}$$

$$\partial_t L_q = \frac{1}{f_0} \sum_i S_{i\,q+m}\, \langle\, X_i,\, RHS(f)\,\rangle_x - \sum_{il} S_{i\,q+m}\, \partial_t S_{il}\, V_l, \qquad 1 \le q \le r-m \tag{3.12}$$

mMn sollte die Ableitung im letzten Term eigentlich aud das erste S sein, aber so haben die es im Einkemmer Paper gemacht und es functioniert

with

$$K_k = \sum_i X_i\, S_{ik}, \qquad L_q = \sum_{il} S_{i\,q+m}\, S_{il} W_l. \tag{3.13}$$

Using a stepping scheme

$$K(t+\tau) = K(t) + \tau \partial_t K(t), \qquad L(t+\tau) = L(t) + \tau \partial_t L(t) \tag{3.14}$$

we may obtain $X(t+\tau)$, $V(t+\tau)$ via QR factorization

$$K_k(t+\tau) = \sum_i X_i(t+\tau)\, R^1_{ik}, \qquad L_q(t+\tau) = \sum_i W_i(t+\tau)\, R^2_{iq}. \tag{3.15}$$

Finally, we compute $S(t+\tau)$ using equation 3.10 and a best-approximation of f:

$$f \approx \widetilde{f} := f_0\, X(t+\tau)^T\, (M^T S N)\, V(t+\tau), \tag{3.16}$$

$$S_{kl}(t+\tau) = \sum_{ij} M^T_{ki}\, S_{ij}\, N_{jl} + \tau \left(X_k,\, \left(V_l,\, RHS(\widetilde{f}) \right)_v \right)_x \tag{3.17}$$

where

$$M_{ki} = \langle X_k(t),\, X_i(t+\tau) \rangle_x\,, \qquad N_{jl} = \langle V_j(t),\, V_l(t+\tau) \rangle_v\,. \tag{3.18}$$

Rename M and N so that it doesn't conflict with moments

Crucially, the approximation $f \approx \widetilde{f}$ doe *not* conserve any of the invariants since the projections in 3.18 are not conservative. Therefore we need to expand the basis onto which we project. Specifically, let $(\widetilde{X}_j)_j$ be an orthonormal basis of span $\{X_i(t), \nabla X_i(t), K_i(t+\tau)\}_i$ and $(\widetilde{V}_j)_j$ an orthonormal basis of span $\{V_i(t), L_q(t+\tau)\}_{iq}$. Then, the projections

$$M_{ki} = \left\langle \widetilde{X}_k,\, X_i(t+\tau) \right\rangle_x\,, \qquad N_{jl} = \left\langle \widetilde{V}_j,\, V_l(t+\tau) \right\rangle_v \tag{3.19}$$

are mass, momentum, and energy conservative [5]. However, this has increased the rank of the approximation f. Thus, we need to truncate the approximation in a way which ensures that the fixed basis functions of U remain unchanged. For convenience write

$$\widetilde{S} = M^T S(t+\tau) N \tag{3.20}$$

where M and N are as in equation 3.19. Letting $\widetilde{K}^T = \widetilde{X}^T \widetilde{S}$ and using the structure of $\widetilde{V}$,

$$f(t+\tau) \approx \widetilde{K}^T \widetilde{V} = \begin{bmatrix} (\widetilde{K}^{cons})^T & (\widetilde{K}^{rem})^T \end{bmatrix} \begin{bmatrix} U \\ \widetilde{W} \end{bmatrix} \tag{3.21}$$

where $\widetilde{K}^{cons}$ is the first m components of $\widetilde{K}$, and $\widetilde{K}^{rem}$, $\widetilde{W}$ are the last components of $\widetilde{K}$, $\widetilde{V}$ respectively.

Hence, by truncating $\widetilde{K}^{rem}$ and $\widetilde{W}$, the desired components of U remain unaffected. We perform the truncation as follows: QR factorizations of $\widetilde{K}^{cons}$, $\widetilde{K}^{rem}$ yield

$$\widetilde{K}^{cons}_k = \sum_i X^{cons}_i \, S^{cons}_{ik}, \qquad \widetilde{K}^{rem}_q = \sum_j \widetilde{X}^{rem}_j \, \widetilde{S}^{rem}_{jq}. \tag{3.22}$$

By a truncated singular value decomposition of $\widetilde{S}^{rem}$, keeping only the largest $r-m$ singular values, we have

$$\widetilde{S}^{rem} \approx \widehat{U} \widehat{S} \widehat{W}. \tag{3.23}$$

Now set $S^{rem} = \widehat{S}$ and

$$X^{rem}_q = \sum_i \widetilde{X}^{rem}_i \, \widehat{U}_{iq}, \qquad \widetilde{W}_q = \sum_j \widetilde{W}_j \, \widehat{W}_{jq}, \qquad 1 \le q \le r-m. \tag{3.24}$$

Combining $\widehat{X} = \begin{bmatrix} X^{cons} \\ X^{rem} \end{bmatrix}$ and performing a final QR factorization

$$\widehat{X}_k = \sum_i \check{X}_i \, R_{ik} \tag{3.25}$$

finishes the truncation, as we set

$$X(t+\tau) = \check{X}, \qquad S(t+\tau) = R \begin{bmatrix} S^{cons} & \\ & S^{rem} \end{bmatrix}, \qquad V(t+\tau) = \begin{bmatrix} U \\ \widetilde{W} \end{bmatrix}. \tag{3.26}$$

While in this case we have performed the time-stepping in equations 3.21, 3.17 via a simple explicit Euler scheme, the extension to time steps of higher order is immediate. Indeed, we refer to [2] for an extension of the robust integrator using the midpoint rule. Pseudocode for the presented algorithm can be found in [5] and implementation of the algorithm (as well as the midpoint-rule extension) in [6].

4 Numerical Example: Landau Damping

1. Basic takeaway of Landau damping: dissipation can occur without loss of energy; different than in fluid equations

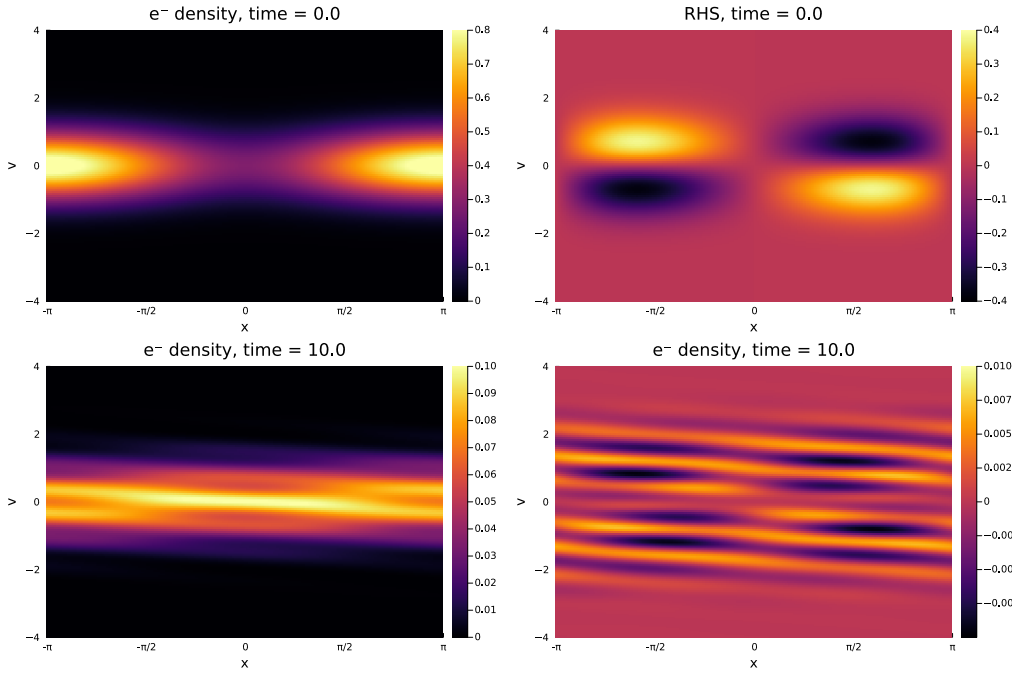

Figure 4.1: Landau damping. Density 4.1 and RHS for various times. !! Note the differing color bar axes!

2. Specific initial conditions that we consider:

$$f(x, v, 0) = (1 + \alpha \cos(x)) \exp(-v^2) \tag{4.1}$$

for the parameter value $\alpha = 1/2$ computed over the domain $\Omega_x = [-\pi, \pi)_{per}$, $\Omega_v = [-5, 5]$. f is normalized to have initial mass 1.

3. This is too large of a parameter value to use linear approx from Landau's original argument, but the derivation can be done to higher order and dissipation is still analytically predicted. Indeed, we see this numerically as well, see fig. 4.1.

4. Odd power moments seem to have very predictable behavior. The error follows a power/exponential law very well (see fig. 4.2). This must have something to do with some form of symmetry wrt. Ω_v. Are the kinetic moment density functions

$$\underbrace{vv\ldots v}_{k \text{ times}} f(x, v, t) \tag{4.2}$$

symmetric wrt. v? All odd moments $\mathcal{M}_k$ are essentially 0. Even power moments seem more erratic but are lower long-term. This could be a power law, though it is difficult to tell. Another, longer integration may be needed.

5. Observables seem to be near-constant whenever the density function is positive, see fig. **??**. As soon as f becomes negative, the observables go haywire. Should a positivity-preserving method be considered? How would this affect the conservation?

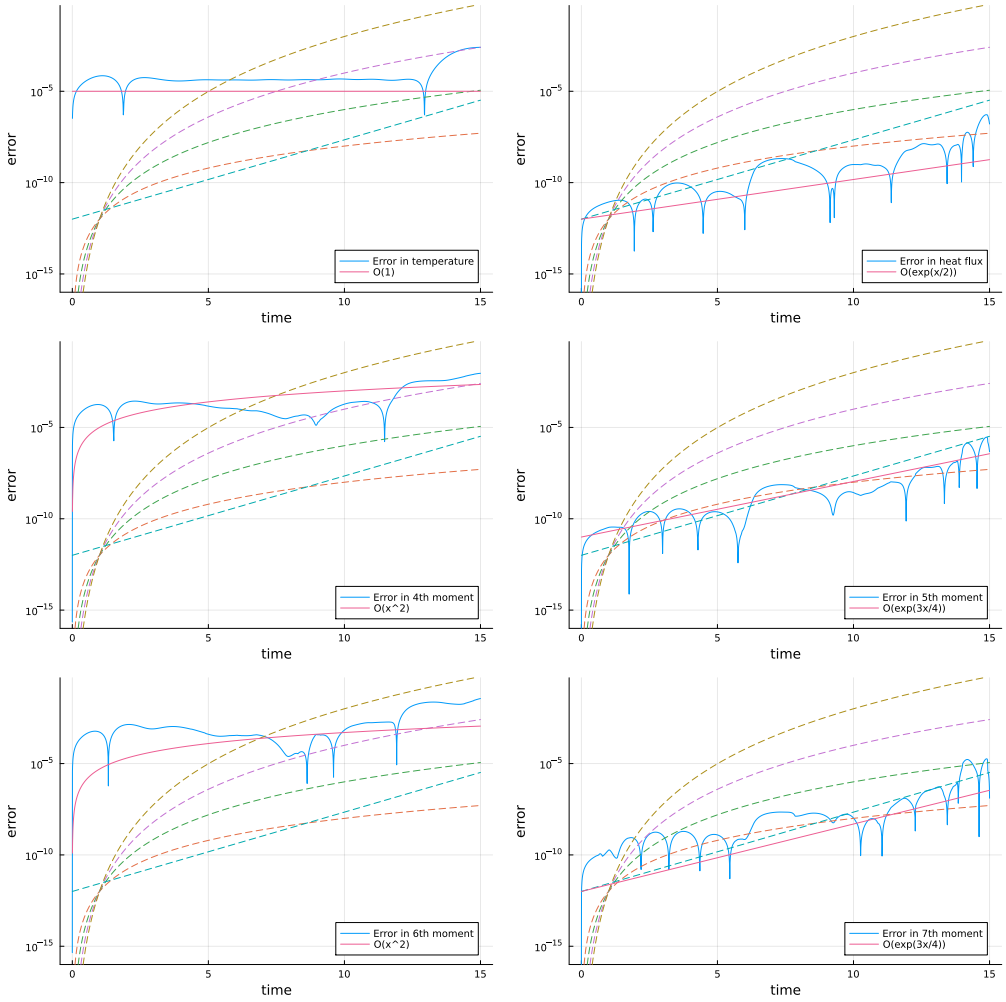

Figure 4.2: Even- and odd-power moment errors based on one very exact numerical integration. Dashed overlay are various power laws: orange: $O(\tau^4)$, green: $O(\tau^6)$, pink: $O(\tau^4)$, yellow: $O(\tau^1 0)$, light blue: $O(exp(\tau))$.

5 Conclusion

References

[1] Fabio Cassini and Lukas Einkemmer. "Efficient 6D Vlasov simulation using the dynamical low-rank framework Ensign". In: *Computer Physics Communications* 280 (2022), p. 108489. ISSN: 0010-4655. DOI: 10.1016/j.cpc.2022.108489. URL: http://dx.doi.org/10.1016/j.cpc.2022.108489.

[2] Gianluca Ceruti et al. *A robust second-order low-rank BUG integrator based on the midpoint rule*. 2024. arXiv: 2402.08607 [math.NA]. URL: https://arxiv.org/abs/2402.08607.

[3] Lukas Einkemmer and Ilon Joseph. "A mass, momentum, and energy conservative dynamical low-rank scheme for the Vlasov equation". In: *Journal of Computational Physics* 443 (2021), p. 110495. ISSN: 0021-9991. DOI: 10.1016/j.jcp.2021.110495. URL: http://dx.doi.org/10.1016/j.jcp.2021.110495.

[4] Lukas Einkemmer and Christian Lubich. *A low-rank projector-splitting integrator for the Vlasov–Poisson equation*. 2018. eprint: 1801.01103. URL: https://arxiv.org/abs/1801.01103.

[5] Lukas Einkemmer, Alexander Ostermann, and Carmela Scalone. "A robust and conservative dynamical low-rank algorithm". In: *Journal of Computational Physics* 484 (2023), p. 112060. ISSN: 0021-9991. DOI: https://doi.org/10.1016/j.jcp.2023.112060.

[6] April Herwig. *DynamicalLowRankIntegrators.jl*. https://github.com/April-Hannah-Lena/DynamicalLowRankIntegrators.jl. 2025.

[7] Othmar Koch and Christian Lubich. "Dynamical Low-Rank Approximation". In: *SIAM Journal on Matrix Analysis and Applications* 29.2 (2007), pp. 434–454. DOI: 10.1137/050639703. URL: https://doi.org/10.1137/050639703.