SUPERCONVERGENCE AND ACCURACY ENHANCEMENT OF DISCONTINUOUS GALERKIN SOLUTIONS FOR VLASOV-MAXWELL EQUATIONS AND NUMERICAL ANALYSIS OF A HYBRID METHOD FOR RADIATION TRANSPORT

By

Andrés Felipe Galindo Olarte

A DISSERTATION

Submitted to
Michigan State University
in partial fulfillment of the requirements
for the degree of

Applied Mathematics—Doctor of Philosophy

2023

ABSTRACT

In this thesis we will analyze and enhance two schemes for kinetic equations. Namely the discontinous Galerkin (DG) methods for solving the the Vlasov-Maxwell (VM) system and a hybrid method for solving the time-dependent radiation transport equation (RTE).

In Chapter 2 we will consider the DG methods for solving the VM system, a fundamental model for collisionless magnetized plasma. The DG methods provide accurate numerical description with conservation and stability properties. However, to resolve the high dimensional probability distribution function, the computational cost is the main bottleneck even for modern-day supercomputers. The first part of this thesis studies the applicability of a post-processing technique to the DG solution to enhance its accuracy and resolution for the VM system. This postprocessor is applied at the final time of the simulation, and its cost is negligible, it succeeds by producing a high-resolution solution with the same cost of computing a low-resolution one, thus saving computational time in the process. In particular, we prove the superconvergence of order $(2k + \frac{1}{2})$ in the negative order norm for the probability distribution function and the electromagnetic fields when piecewise polynomial degree k is used. Numerical tests including Landau damping, two-stream instability and streaming Weibel instabilities are considered showing the performance of the post-processor. This is based on joint work with Yingda Cheng, Juntao Huang and Jennyfer Ryan [1].

In Chapter 3, we prove rigorous error estimates for a hybrid method introduced in [2] for solving the time-dependent RTE. The method relies on a splitting of the kinetic distribution function for the radiation into uncollided and collided components. A high-resolution method (in angle) is used to approximate the uncollided components and a low-resolution method is used to approximate the the collided component. After each time step, the kinetic distribution is reinitialized to be entirely uncollided. For this analysis, we consider a mono-energetic problem on a periodic domains, with constant material cross-sections of arbitrary size. We assume the uncollided equation is solved ex-

actly and the collided part is approximated in angle via a spherical harmonic expansion (P_N method). Using a non-standard set of semi-norms, we obtain estimates of the form $C(\varepsilon, \sigma, \Delta t)N^{-s}$ where $s \geq 1$ denotes the regularity of the solution in angle, ε and σ are scattering parameters, Δt is the time-step before reinitialization, and C is a complicated function of ε, σ, and Δt. These estimates involve analysis of the multiscale RTE that includes, but necessarily goes beyond, usual spectral analysis. We also compute error estimates for the monolithic P_N method with the same resolution as the collided part in the hybrid. Our results highlight the benefits of the hybrid approach over the monolithic discretization in both highly scattering and streaming regimes. This is based in a joint work with Cory D. Hauck and Victor Decaria [3].

"Las cosas solo son puras si uno las mira desde lejos, es muy importante conocer nuestras raíces , saber de donde venimos, conocer nuestra historia , pero al mismo tiempo tan importante como saber de donde somos, es entender que todos en el fondo somos de ningún lado del todo y de todos lados un poco."

Jorge Drexler

ACKNOWLEDGEMENTS

The last five years at Michigan State University have been the most fulfilling and professionally rewarding of my life thus far. This journey would not have been possible without Professor Yingda Cheng's guidance. I thank her for taking me as her student, for her patience, valuable feedback, and for helping me navigate the job market. I will always be grateful for the opportunity to work with her.

I also want to thank Irene M Gamba and the Oden institute at University of Texas at Austin for giving me a job opportunity for the next two years.

At MSU, there are many people to thank. First, I would like to acknowledge my committee members, Professor Daniel Appelö, Professor Andrew Christlieb, and Professor Zhengfang Zhou, for their guidance and for taking the time to read and evaluate my work. I would also like to thank Professor Peter Bates for continuously believing in me and recruiting me to this special place called Michigan State University (MSU). Thank you for always giving me advice and helping me prepare my job application material and interviews. I would also like to thank my fellow group members, Amit Rotem, Kai Huang, Zhichao Peng, Yann-Meing Law, and Spencer Lee, for their support. I would also like to thank my friends from the math department who have made my stay more pleasant: Luis Suarez, Danika Van Niel, Cullen Haselby, Shih-Fang Yeh, Yuta Hozumi, Chloe Lewis, Gokul Bhusal, Bowen Su, Azzam Alfarraj, Joe Melby, Leonardo Abbrescia, Rodrigo Becerra Matos, Ben Jones, Zhixin Wang, Chen Zhang, Max Throm, Shikha Bhutani, Rithwik Vidyarthi, Owen Ekblad, Rolando Ramos, Aldo Garcia, Valerie Jean Pierre, Christopher Potvin, and Albert Chua. I want to thank the math department staff, particularly Laura Willoughby, Estrella Starn, Taylor Alvarado, and Sabrina Walton. Thanks for always helping me!

I want to acknowledge my collaborators in all the projects in this dissertation; it would have been impossible to carry out my research without their help, Jennifer Ryan, Juntao Huang, Cory D Hauck, and Victor Decaria. Thank you also for guiding me in my profes-

sional career.

My bachelor's degree was fundamental to my formation. I want to thank the Universidad Distrital Francisco José de Caldas for giving this kid from the south of Bogotá the opportunity to afford high-level education and, thus, the first step to becoming a mathematician. The first person I would like to acknowledge is Professor Arturo Sanjuán. He was my first mentor, and I am grateful to call him my friend. Thanks for all the advice on life and mathematics throughout the years and for trusting me to become his first student. He showed me that I could be a person and a mathematician. I also want to thank professors Alejandro Masmela, Deccy Trejos, Luis Fernando Villarraga, Herbert Sarmiento, Oriol Mora, and Carlos Ochoa.

My time at the Universidad de los Andes was terrific. Those three years and a half were fundamental for me. I learned more about math and the profession. On top of that, I had the best friends/colleagues to support me at every step. Thank you Jerson Caro, Daniel Avila, Rodolfo Quintero, Sebastián Osorio, Nicolás Escobar, Santiago Pinzon, Hernán Garcia, Edison Leguizamon, Gustavo Chaparro, David Moreno Paris, Andrés Patiño, Mariana Vicaria, Duvan Cardona, Weimar Astaiza, Juanita Duque and Laura Gamboa. Your friendship means a lot to me! I also want to thank my advisor, Jean Cortissoz, for helping me through my thesis process and navigating my doctoral program search. I also want to thank professors Monika Winklemeir, Florent Schaffhauser, Alexander Getmanenko, and Andrei Giniatoulline. Thank you for all your lessons and mentoring.

I would also like to recognize all my friends from the Comunidad Latinoamericana at MSU. I appreciate their vote of confidence in me as president of this organization for three years. Thank you for helping me create a space where international students from Latin America and the Caribbean could find a place where they feel at home, correctly represented, respected, and understood. I would like to especially mention Francisco Flores, Marisol Masso, Angelica Herreño, Guillermo Huanes, Vanessa Maldonado, Diego

Sierra, Diego Granados, Akash Saxena, Jorge Nevares, Erik Amezquita, Skyy Pineda, Jennifer Mojica Santana, Viridiana Garcia, Giovanny Salazar, Henry Gonzalez, Gerardo Melgar, Martina Borges, Luisa Parrado, Marcela Tabares, Paulo Izquierdo, Viviana Ortiz, Eloy Moreno, Astrid Olave, Ian Fisher, Tayna Carrasquillo, Joon Chung, Joelis Lama, Santiago Hernandez, Santiago Rodriguez, Lina Gomez, Jose Quintero, Francisco Santos, Maria Buitrago, Angelica Bernabe, Julian Bello, William Torres, Juan Sebastian Hernandez, Dennise Celis, Carolina Vargas, Catalina Barraza, Daniel Hoffman, Juan Sandoval, Angie Vega, Marco Lopez, Mayra Florez, Pepe Montero, Andrea Bernardes, Christian Gonzalez, Daniel Maldonado, Rafael Castro, Laura Castro, María Alejandra Garcia, Nicolas Scamardi, Martina Gimenez, Omar Posos, Francisco Campos, Pedro Chu, Gianna Mendez, Andres Lanzas, Julian Quiroga, Valeria Obando, and Jean-Paul Ortiz. Thanks to the Center for Latin American and Caribbean Studies (CLACS) for supporting us. Thank you, Laura Medina, for serving as our faculty advisor for three years. I cannot finish this paragraph without thanking the Office for International Students and Scholars (OISS) for helping us with resources and support at different events—thanks to Leidy Egan, Clara Graucob, and Krista McCallum.

I cannot express my gratitude to my parents, Martha Olarte and José Galindo. Thank you for always supporting me and my brother through difficult times. You always gave me the utmost care and filled my life with love and support. Thanks for constantly believing in me.

Last but not least, I want to give all my love and thanks to my partner Melina Jimenez. Thank you for the gift of your love and company. You make my life better. This Ph.D. journey would not have been the same without you. TE AMO.

TABLE OF CONTENTS

CHAPTER 1

INTRODUCTION

The evolution of a large particle system such as gases or plasmas at the microscopic level is described by systems of ODE's, however in general solving these systems numerically is extremelly costly and brings little to no insight into the macroscopic behaviour of the system. Thus we seek a reduced model of particle dynamics, that bridges the gap between the microscopic and macroscopic description of the physical phenomena and that is also accurate. One of such models are the kinetic equations [4],[5]. Kinetic equations intend to describe these particles systems by means of a distribution functions f in phase space. This phase space includes macroscopic variables i.e the position in physical space, but also microscopic variables which describe the state of the particle, in this thesis the only microscopic variables that will be considered are the velocity components, other examples of microscopic variables are, internal energy and spin variables, etc. This object f represents a particle density in phase space i.e. $f\,dx\,dv$ is the number of particles in a small volume [4]. Kinetic equations have applications in different fields such as: gas dynamics, plasma physics, biology, socieconomics, nuclear engineering, etc.

In the classical kinetic theory of rarified gases, the variation of a non-negative function, $f = f(x, v, t)$, characterizing the particle densities having velocity $v \in \mathbb{R}^3$ in position $x \in \mathbb{R}^3$ at time t, is obtained via the equation

$$\frac{\partial f}{\partial t} + v \cdot \nabla_x f = Q(f)$$

The operator $Q(f)$, on the right-hand side of equation (1), describes the effects of internal forces due to particle interactions, and its form depends on the details of the microscopic dynamic [5]. The most well-known examples are Boltzmann's equation and the Vlasov mean-field equation.

There are numerous challenges for numerical solvers for kinetic equations. The first one coming from the intrinsic high-dimensionality of the problem, in general $(x, v, t) \in \mathbb{R}^7$.

There are additional difficulties and requirements specific to kinetic equations:

- *Conservation properties* We would like to preserve physicial laws, such as conservation. [5].
- *Computational cost.* Aside from the computational cost that comes from high-dimensions. Computations using the operator $Q(f)$ involve the computation of high-dimensional integrals in velocity space at each point in physical space [5].
- *Presence of multiple scales*. Usually one have to deal with multiple space-time scales, which span different regimes. [5].

In this work we will consider deterministic numerical methods for kinetic equations. For a survey on these methods see [5].

In this thesis we intend to analyze two different schemes for kinetic equations. In the rest of this chapter we would like introduce the reader to two schemes: Discontinuous Galerkin method for the Vlasov-Maxwell equations and a hybrid method for radiation transport.

1.1 Discontinuous Galerkin method for the Vlasov-Maxwell equations

In the first part of this thesis, we consider numerical solutions of the Vlasov-Maxwell (VM) system, a fundamental model for collisionless magnetized plasma. The dimensionless form of the equations that describes the evolution of a single species of non-relativistic electrons under the self-consistent electromagnetic field while the ions are treated as uniform fixed background is given by

$$\partial_t f + \mathbf{v}\cdot\nabla_{\mathbf{x}} f + (\mathbf{E} + \mathbf{v}\times\mathbf{B})\cdot\nabla_{\mathbf{v}} f = 0, \tag{1.1a}$$

$$\frac{\partial \mathbf{E}}{\partial t} = \nabla_{\mathbf{x}}\times\mathbf{B} - \mathbf{J}, \quad \frac{\partial \mathbf{B}}{\partial t} = -\nabla_{\mathbf{x}}\times\mathbf{E}, \tag{1.1b}$$

$$\nabla_{\mathbf{x}}\cdot\mathbf{E} = \rho - \rho_i, \quad \nabla_{\mathbf{x}}\cdot\mathbf{B} = 0, \tag{1.1c}$$

with

$$\rho(\mathbf{x},t) = \int_{\Omega_v} f(\mathbf{x},\mathbf{v},t)\,d\mathbf{v}, \quad \mathbf{J}(\mathbf{x},t) = \int_{\Omega_v} f(\mathbf{x},\mathbf{v},t)\mathbf{v}\,d\mathbf{v},$$

where the equations are defined on $\Omega = \Omega_x \times \Omega_v$, $\mathbf{x} \in \Omega_x$ denotes the position in physical space, and $\mathbf{v} \in \Omega_v$ in velocity space. Here $f(\mathbf{x}, \mathbf{v}, t) \geq 0$ is the distribution function of electrons at position $\mathbf{x}$ with velocity $\mathbf{v}$ at time t, $\mathbf{E}(\mathbf{x}, t)$ is the electric field, $\mathbf{B}(\mathbf{x}, t)$ in the magnetic field, $\rho(\mathbf{x}, t)$ is the electron charge density, and $\mathbf{J}(\mathbf{x}, t)$ is the current density. The charge density of background ions is denoted by ρ_i, which is chosen to satisfy total charge neutrality, $\int_{\Omega_x} (\rho(\mathbf{x}, t) - \rho_i)\, d\mathbf{x} = 0$. Periodic boundary conditions in Ω_x and compact support in Ω_v are assumed. The VM system has wide applications in plasma physics for describing space and laboratory plasmas, with application to fusion devices, high-power microwave generators, and large scale particle accelerators.

Much work has been carried out in the literature aiming at accurate deterministic description of the probability density function for nonlinear behavior of charged particles in plasma. Califano *et al.* used a semi-Lagrangian approach to compute the streaming Weibel instability [6], current filamentation instability [7], magnetic vortices [8], magnetic reconnection [9]. Also, various methods have been proposed for the relativistic VM system [10, 11, 12, 13]. This work concerns the discontinuous Galerkin (DG) method for solving the VM system. The DG method is a class of finite element method that uses discontinuous polynomial spaces, and they have desirable properties for convection-dominated problems [14]. In particular, DG methods have been used to simulate the Vlasov-Poisson system in plasmas [15, 16, 17] and for a gravitational infinite homogeneous stellar system [18]. They have been also used to solve VM system [19, 20] and the relativistic VM system [21]. The DG methods have nice properties such as stability, charge and energy conservation and high order accuracy, which are highly desirable for long time simulations.

1.2 Accuracy enhanement and Superconvergence techniques for DG methods

The main computational challenge for any grid based solver for the VM system is the high-dimensionality of the Vlasov equation. This makes the computation extremely expensive even on modern-day exa-scale supercomputers. Post-processing techniques, which can greatly enhance the resolution of the numerical solution at any given time,

are therefore desirable because it is only applied once at the end of the simulation with negligible computational cost. Post-processing for finite element methods is a mature technology. The post-processing technique presented here takes advantage of the information contained in the negative-order norm and was originally developed by Bramble and Schatz [22] in the context of continuous finite element methods for elliptic problems. It consists of a convolution of the finite element solution with a local averaging operator. We can then establish the convergence in the negative order norm which is higher than that one obtained in the usual L^2-norm. In [23], Cockburn, Luskin, Shu and Süli applied this technique to the DG methods for solving linear hyperbolic equations. This technique was further extended to the DG methods for solving nonlinear conservational laws [24, 25] and nonlinear symmetric systems of hyperbolic conservation laws [26]. This method is currently part of a filtering family known as a Smoothness-Increasing Accuracy-Conserving (SIAC) filters [27]. This chapter will demonstrate the performance of post-processing by the SIAC filter for DG solutions to the VM system. In particular, we consider benchmark numerical tests for Vlasov-Ampére (VA) and VM systems, and study the numerical error for short and long time simulations with varying polynomial order.

In order to validate the enhanced accuracy of the post-processed solution, an important step is to establish the superconvergence of the negative order norm of the error and its divided differences. In [23], Cockburn, Luskin, Shu and Süli established a framework to prove negative-order estimates for the DG solutions to linear conservational laws of order $2k+1$ using polynomials of degree k. After this, there have been important extensions. L^2 and L^∞ superconvergence estimates were established for DG solutions for linear constant coefficient hyperbolic systems with the position-dependent SIAC filter [28]. Ji, Meng *et al* [24, 25, 26] proved superconvergence for non-linear conservation laws and nonlinear symmetric hyperbolic systems of the DG solutions of order at least $(\frac{3}{2}k+1)$. It is highly nontrivial to establish superconvergence for nonlinear problems because a suitable

dual problem has to be identified, and additionally the divided difference of the solution does not satisfy the PDE, which makes the proof highly technical [25, 26]. In Chapter 2, we aim to prove negative-order estimates of DG solutions to the VM system. Since the VM system is nonlinear, it is nontrivial to extend the proof in [23]. We identify a proper dual problem, which aids the estimates of the consistency term. In the end, we proved superconvergence of order $(2k + \frac{1}{2})$ in the negative norm for the probability distribution function and the electromagnetic fields.

1.3 A hybrid method for the radiation transport equation

The second part of this thesis will deal with the radiation transport equation (RTE) [29, 30, 31]. This equation describes the movement of particles through a material medium by means of a kinetic distribution function that gives the density of particles with respect to the local phase space measure. In a general setting, the phase space is six dimensional: three dimensions for particle position and three for particle momentum, the latter of which is typically decomposed into energy and direction (or angle) of flight. Thus in the time-dependent setting, the RTE is defined over a seven-dimensional domain.

The RTE describes two basic processes: particle advection and interactions with the material medium. These interactions can be of various types and include scattering and emission/absorption processes. The rate at which these processes occur is determined by the properties of the material, expressed via cross-sections. Material cross-sections may vary in space and depend on the particle energy and, in situations that the material evolves, the cross-sections may evolve as well. When cross-sections vary significantly, the RTE may exhibit multiscale behavior. It is the combination of this multiscale behavior with the high-dimensional phase space that makes simulating the RTE a challenging task.

A well-known multi-scale feature of the RTE is the diffusion limit. In regions where the scattering cross-section is large, the solution of the RTE can be accurately approximated by its angular average [32, 33]. Moreover this average is well-approximated by the solution of a diffusion equation. This solution to diffusion equation has long been used

as a cheap approximation the solution to the RTE in scattering dominated regimes.

Another common limit is the absorption limit, which is characterized by a complete lack of scattering. In this case, the RTE does not have a simple asymptotic approximation. However, due the abscence of scattering, there is no coupling between the angles and energies of the kinetic distribution. Thus, with a proper discretiation, the RTE solution can be easily parallelized.

In problems for which the scattering cross-section varies dramatically, both of the limits above can exist simultaneously, along with a range of transition regimes in between. A consequence of this fact is that a monolithic numerical treat of the RTE will require many degrees of freedom that are strongly coupled. In practice, the time-dependent RTE is often updated in time with an implicit scheme. In such cases, designing the linear solvers can be a challenge.

A variety of approaches have been proposed for addressing the multiscale challenges posed of the RTE. These include micro-macro decompositions [34], high order-low order (HOLO) methods [35], diffusion-based acceleration [30, 36], and preconditioned Krylov approaches [37]. In this thesis, we consider a hybrid formulation [2] that is based on the notion of first-collision source [38]. In this hybrid formulation, the RTE is split into two components: an uncollided component that tracks the particles up to point of their first material interaction and a collided component that track the particles that remain. The resulting system is then approximated with two different angular discretizations: a high-resolution discretization for the uncollided equation and a low-resolution for the collided equation. The intuition that drives this strategy is that scattering produces a smoother solution; hence the collided equation should require less resolution to recover an accurate solution. The uncollided equation, on the other, requires higher resolution; however it takes the form of a purely absorbing RTE and can therefore be solved much more efficiently the original RTE using the same number of degrees of freedom. The efficiency of the hybrid approach for the RTE has been demonstrated in several papers

[39, 40, 41], including generalizations to hybrid energy discretizations [42] and hybrid spatial discretizations [43].

A key component of the hybrid implementation for the time-dependent RTE is a relabeling procedure that, after a given time step, maps the collided numerical solution into the space of the uncollided numerical solution and then uses the sum to re-initialize the uncollided equation. Meanwhile, the collided equation is re-initialized to zero. This relabeling step is critical, since otherwise the hybrid numerical solution would eventually convergence to a low-resolution numerical solution of the collided equation.

Despite the intuitive motivation of the hybrid and the success of the hybrid approach in numerical simulations, the method still lacks rigorous justification. This is due in part to complications introduced by the multiscale behavior of the RTE. For example, spectral approximations of the RTE in angle are fairly straightforward to analyze [44], but a multiscale analysis that takes into account the degree of scattering is significantly more complicated [45]. The relabeling step of the hybrid formulation complicates the situation even further.

In Chapter 3, we take a first step in analyzing the hybrid method for the time-dependent mono-energetic version of the RTE with isotropic scattering. We focus only on the angular discretization of the RTE, comparing the standard spectral approximation (P_N) for the full system with a discretization of the hybrid that features a spectral approximation of the same resolution for the collided equation but assumes an exact solution for the uncollided equation. Clearly, the hybrid formulated in this way is more expensive than the monolithic approach. Thus the goal of the analysis is determine what is gained from the extra work involves in a high-resolution simulation of the uncollided equation, which in practice is computed with a high-fidelity collocation method or with a Monte-Carlo method.

In Chapter 4, we provide a brief conclusion and future work.

CHAPTER 2

SUPERCONVERGENCE AND ACCURACY ENHANCEMENT OF DISCONTINUOUS GALERKIN SOLUTIONS FOR VLASOV-MAXWELL EQUATIONS

In this chapter, we will provide the first step towards proving rigorously accuracy enhancement of the postprocessed DG solution to the VM system (1.1), by proving superconvergence of the negative norm of the DG solution. The remainder of the chapter is organized as follows. In Section 2.1, we introduce the DG method for the VM system as well as relevant notations that will be required for the negative order estimates. In Section 2.2 we introduce SIAC filtering. In Section 2.3 we prove the negative-order norm estimates of the DG solutions to the VM system. The superconvergence results are confirmed numerically in Section 2.4.

2.1 Discontinuous Galerkin Numerical Scheme

2.1.1 Notations, Definitions and Projections

We begin by introducing the necessary notation used in the chapter. Without loss of generality, we assume the spatial and velocity domain to be $\Omega_x = [-L_x, L_x]^{d_x}$ and $\Omega_v = [-L_v, L_v]^{d_v}$, where L_v is chosen large enough so that $f = 0$ at $\partial\Omega_v$. Through out the chapter, standard notations will be used for the Sobolev spaces. Given a bounded domain $D \in \mathbb{R}^\star$ (with $\star = d_x$,d_v, or $d_x + d_v$) and any nonnegative integer m, $H^m(D)$ denotes the L^2-Sobolev space of order m with the standard Sobolev norm $\|\cdot\|_{m,D}$, $W^{m,\infty}$ denotes the L^∞-Sobolev space of order m with the standard Sobolev norm $\|\cdot\|_{m,\infty,D}$ and the semi-norm $|\cdot|_{m,\infty,D}$. When $m = 0$, we also use $H^0(D) = L^2(D)$ and $W^{0,\infty}(D) = L^\infty(D)$.

Let $\mathcal{T}_h^x = \{K_x\}$ and $\mathcal{T}_h^v = \{K_v\}$ be partitions of Ω_x and Ω_v, respectively, with K_x and K_v being Cartesian elements or simplices; then $\mathcal{T}_h = \{K : K = K_x \times K_v, \forall K_x \in \mathcal{T}_h^x, \forall K_v \in \mathcal{T}_h^v\}$ defines a partition of Ω. Let $\mathcal{E}_x$ be the set of the edges of $\mathcal{T}_h^x$ and $\mathcal{E}_v$ the set of the edges of $\mathcal{T}_h^v$; then the edges of $\mathcal{T}_h$ will be $\mathcal{E} = \{K_x \times e_v : \forall K_x \in \mathcal{T}_h^x, \forall e_v \in \mathcal{E}_v\} \cup \{e_x \times K_v : \forall e_x \in \mathcal{E}_x, \forall K_x \in \mathcal{T}_h^x\}$.

Furthermore, $\mathcal{E}_v = \mathcal{E}_v^i \cup \mathcal{E}_v^b$ with $\mathcal{E}_v^i$ and $\mathcal{E}_v^b$ being the set of interior and boundary edges of $\mathcal{T}_h^v$ respectively. In addition, we denote the mesh size of $\mathcal{T}_h$ as $h = \max(h_x, h_v) = \max_{K\in\mathcal{T}_h} h_K$, where $h_x = \max_{K_x\in\mathcal{T}_h^x} h_{K_x}$ with $h_{K_x} = \operatorname{diam}(K_x)$, $h_v = \max_{K_v\in\mathcal{T}_h^v} h_{K_v}$ with $h_{K_v} = \operatorname{diam}(K_v)$, and $h_K = \max(h_{K_x}, h_{K_v})$ for $K = K_x \times K_v$. When the mesh is refined, we assume both $\frac{h_x}{h_{x,\min}}$ and $\frac{h_v}{h_{v,\min}}$ are uniformly bounded from above by a positive constant σ_0. Here $h_{x,\min} = \min_{K_x} h_{K_x\in\mathcal{T}_h^x}$ and $h_{v,\min} = \min_{K_v\in\mathcal{T}_h^v} h_{K_v}$. It is further assumed that $\{\mathcal{T}_h^\star\}_h$ is shape-regular with $\star = x$ or v. That is, if $\rho_{K_\star}$ denotes the diameter of the largest sphere included in $K_\star$, there is

$$\frac{h_{K_\star}}{\rho_{K_\star}} \le \sigma_\star, \quad \forall K_\star \in \mathcal{T}_h^\star$$

for a positive constant $\sigma_\star$ independent of $h_\star$. Furthermore the inner products are defined as

$$(g,h)_\Omega = \int_\Omega gh\,dx\,dv = \sum_{K\in\mathcal{T}_h}\int_K gh\,dx\,dv, \tag{2.1}$$

$$(\mathbf{U},\mathbf{W})_{\Omega_x} = \int_{\Omega_x} \mathbf{U}\cdot\mathbf{W}\,dx = \sum_{K_x\in\mathcal{T}_h^x}\int_{K_x} \mathbf{U}\cdot\mathbf{W}\,dx. \tag{2.2}$$

Now for $g \in L^2(\Omega)$, $\mathbf{U},\mathbf{W} \in (L^2(\Omega_x))^{d_x}$, we define the L^2-norm of $(g,\mathbf{U},\mathbf{W})$ as

$$\|(g,\mathbf{U},\mathbf{W})\|_{0,\Omega} = \sqrt{\|g\|_{0,\Omega}^2 + \|\mathbf{U}\|_{0,\Omega_x}^2 + \|\mathbf{W}\|_{0,\Omega_x}^2} \tag{2.3}$$

This will be helpful in the error analysis of the negative-order norm. The negative order norm of order l is defined as: given $l > 0$ and domain Ω,

$$\|(g,\mathbf{U},\mathbf{W})\|_{-l,\Omega} = \sup_{\phi\in C_0^\infty(\Omega),\mathcal{U},\mathcal{W}\in[C_0^\infty(\Omega_x)]^{d_x}} \frac{(g,\phi)_\Omega + (\mathbf{U},\mathcal{U})_{\Omega_x} + (\mathbf{W},\mathcal{W})_{\Omega_x}}{\sqrt{\|\phi\|_{l,\Omega}^2 + \|\mathcal{U}\|_{l,\Omega_x}^2 + \|\mathcal{W}\|_{l,\Omega_x}^2}}$$

Next we define the discrete spaces

$$\begin{aligned}\mathcal{G}_h^k &= \left\{g\in L^2(\Omega) : g|_{K=K_x\times K_v} \in P^k(K_x\times K_v), \forall K_x\in\mathcal{T}_h^x, \forall K_x\in\mathcal{T}_h^x, \forall K_v\in\mathcal{T}_h^v,\right\} \qquad (2.4)\\ &= \left\{g\in L^2(\Omega) : g|_K \in P^k(K), \forall K\in\mathcal{T}_h\right\},\\ \mathcal{U}_h^r &= \left\{\mathbf{U}\in\left[L^2(\Omega_x)\right]^{d_x} : \mathbf{U}|_{K_x}\in\left[P^r(K_x)\right]^{d_x}, \forall K_x\in\mathcal{T}_h^x\right\}, \qquad (2.5)\end{aligned}$$

where $P^r(D)$ denotes the set of polynomials of total degree at most r on D, and k and r are nonnegative integers.

For piecewise functions defined with respect to $\mathcal{T}_h^x$ or $\mathcal{T}_h^v$, we further introduce the jumps and averages as follows. For any edge $e = \{K_x^+ \cap K_x^-\} \in \mathcal{E}_x$, with $\mathbf{n}_x^\pm$ as the outward unit normal to $\partial K_x^\pm$, $g^\pm = g|_{K_x^\pm}$ and $\mathbf{U}^\pm = \mathbf{U}|_{K_x^\pm}$, the jump across e are defined as

$$[g]_x = g^+\mathbf{n}_x^+ + g^-\mathbf{n}_x^-, \quad [\mathbf{U}]_x = \mathbf{U}^+ \cdot \mathbf{n}_x^+ + \mathbf{U}^- \cdot \mathbf{n}_x^-, \quad [\mathbf{U}]_\tau = \mathbf{U}^+ \times \mathbf{n}_x^+ + \mathbf{U}^- \times \mathbf{n}_x^-$$

and the averages are

$$\{g\}_x = \frac{1}{2}(g^+ + g^-), \quad \{\mathbf{U}\}_x = \frac{1}{2}(\mathbf{U}^+ + \mathbf{U}^-).$$

By replacing the subscript x with v, one can define $[g]_v$, $[\mathbf{U}]_v$, $\{g\}_v$, and $\{\mathbf{U}\}_v$ for an interior edge of $\mathcal{T}_h^v$ in $\mathcal{E}_v^i$. For a boundary edge $e \in \mathcal{E}_v^b$ with $\mathbf{n}_v$ being the outward unit normal we use

$$[g]_v = g\mathbf{n}_v, \quad \{g\}_v = \frac{1}{2}g, \quad \{\mathbf{U}\}_v = \frac{1}{2}\mathbf{U}. \tag{2.6}$$

This is consistent with the fact that the exact solution f is compactly supported in $\mathbf{v}$.

For convenience, we introduce some shorthand notations, $\int_{\Omega_\star} = \int_{\mathcal{T}_h^\star} = \sum_{K_\star \in \mathcal{T}_h^\star} \int_{K_\star}$, $\int_\Omega = \int_{\mathcal{T}_h} = \sum_{K\in\mathcal{T}_h} \int_K$, $\int_{\mathcal{E}_\star} = \sum_{e\in\mathcal{E}_\star} \int_e$, where again $\star$ is x or v. In addition, $\|g\|_{0,\mathcal{E}} = (\|g\|^2_{0,\mathcal{E}_x\times\mathcal{T}_h^v} + \|g\|^2_{0,\mathcal{T}_h^x\times\mathcal{E}_v})^{1/2}$ with $\|g\|_{0,\mathcal{E}_x\times\mathcal{T}_h^v} = \left(\int_{\mathcal{E}_x}\int_{\mathcal{T}_h^v} g^2\, d\mathbf{v}\, ds_{\mathbf{x}}\right)^{1/2}$, and we have that $\|g\|_{0,\mathcal{T}_h^x\times\mathcal{E}_v} = \left(\int_{\mathcal{T}_h^x}\int_{\mathcal{E}_v} g^2\, ds_{\mathbf{v}}\, d\mathbf{x}\right)^{1/2}$. We will make use of the following equality, which can be easily verified using the definition of averages and jumps.

$$\frac{1}{2}[g^2]_\star = g_\star[g]_\star, \text{with } \star = x \text{ or } v. \tag{2.7}$$

2.1.2 The DG method for the Vlasov-Maxwell system

Now we review the DG method for the VM system proposed in [19]. The scheme seeks a numerical solution $f_h \in \mathcal{G}_h^k$ and $(\mathbf{E}_h, \mathbf{B}_h) \in \mathcal{U}_h^k \times \mathcal{U}_h^k$ such that for any $g \in \mathcal{G}_h^k$,

$\mathbf{U}, \mathbf{W} \in \mathcal{U}_h^k$,

$$\int_K \partial_t f_h g \, d\mathbf{x}d\mathbf{v} - \int_K f_h \mathbf{v} \cdot \nabla_x g \, d\mathbf{x}d\mathbf{v} - \int_K f_h(\mathbf{E}_h + \mathbf{v} \times \mathbf{B}_h) \cdot \nabla_{\mathbf{v}} g \, d\mathbf{x}d\mathbf{v} + \int_{K_v} \int_{\partial K_x} \widehat{f_h \mathbf{v} \cdot \mathbf{n}}_x g \, ds_{\mathbf{x}} d\mathbf{v} + \int_{K_x} \int_{\partial K_v} (f_h(\mathbf{E}_h + \widehat{\mathbf{v} \times \mathbf{B}}_h) \cdot \mathbf{n}_v) g \, ds_{\mathbf{v}} d\mathbf{x} = 0, \tag{2.8a}$$

$$\int_{K_x} \partial_t \mathbf{E}_h \cdot \mathbf{U} \, d\mathbf{x} = \int_{K_x} \mathbf{B}_h \cdot \nabla_{\mathbf{x}} \times \mathbf{U} \, d\mathbf{x} + \int_{\partial K_x} \widehat{\mathbf{n}_x \times \mathbf{B}}_h \cdot \mathbf{U} \, ds_{\mathbf{x}} - \int_{K_x} \mathbf{J}_h \cdot \mathbf{U} \, d\mathbf{x}, \tag{2.8b}$$

$$\int_{K_x} \partial_t \mathbf{B}_h \cdot \mathbf{W} \, d\mathbf{x} = - \int_{K_x} \mathbf{E}_h \cdot \nabla_{\mathbf{x}} \times \mathbf{W} \, d\mathbf{x} - \int_{\partial K_x} \widehat{\mathbf{n}_x \times \mathbf{E}}_h \cdot \mathbf{W} \, ds_{\mathbf{x}} \tag{2.8c}$$

with

$$\mathbf{J}_h(\mathbf{x}, t) = \int_{\mathcal{T}_h^{\mathbf{v}}} f_h(\mathbf{x}, \mathbf{v}, t) \mathbf{v} \, d\mathbf{v}. \tag{2.9}$$

Here $\mathbf{n}_x$ and $\mathbf{n}_v$ are outward unit normals of ∂K_x and ∂K_v, respectively. All "hat" functions are numerical fluxes that are determined by upwinding, i.e.,

$$\widehat{f_h \mathbf{v} \cdot \mathbf{n}}_x \coloneqq \widetilde{f_h \mathbf{v}} \cdot \mathbf{n}_x = \left(\{f_h \mathbf{v}\}_x + \frac{|\mathbf{v} \cdot \mathbf{n}_x|}{2} [f_h]_x \right) \cdot \mathbf{n}_x \tag{2.10a}$$

$$\begin{aligned} f_h(\mathbf{E}_h + \widehat{\mathbf{v} \times \mathbf{B}}_h) \cdot \mathbf{n}_v &\coloneqq f_h(\mathbf{E}_h + \widetilde{\mathbf{v} \times \mathbf{B}}_h) \cdot \mathbf{n}_v \\ &= \left(\{f_h(\mathbf{E}_h + \mathbf{v} \times \mathbf{B}_h)\}_v + \frac{|(\mathbf{E}_h + \mathbf{v} \times \mathbf{B}_h) \cdot \mathbf{n}_v|}{2} [f_h]_v \right) \cdot \mathbf{n}_v, \end{aligned} \tag{2.10b}$$

$$\widehat{\mathbf{n}_x \times \mathbf{E}}_h \coloneqq \mathbf{n}_x \times \widetilde{\mathbf{E}_h} = \mathbf{n}_x \times \left(\{\mathbf{E}_h\}_x + \frac{1}{2} [\mathbf{B}_h]_\tau \right) \tag{2.10c}$$

$$\widehat{\mathbf{n}_x \times \mathbf{B}}_h \coloneqq \mathbf{n}_x \times \widetilde{\mathbf{B}_h} = \mathbf{n}_x \times \left(\{\mathbf{B}_h\}_x - \frac{1}{2} [\mathbf{E}_h]_\tau \right) \tag{2.10d}$$

where these relations define the meaning of "tilde". In [19], alternating and central fluxes for the Maxwell's equation are also considered. The discussions will be similar to what will be presented in this chapter for the upwind flux, and thus are omitted.

Upon summing up (2.8a) with respect to $K \in \mathcal{T}_h$ and similarly summing (2.8b) and (2.8c) with respect to $K_x \in \mathcal{T}_h^x$, the scheme (2.8) becomes the following: look for $f_h \in \mathcal{G}_h^k$, $\mathbf{E}_h, \mathbf{B}_h \in \mathcal{U}_h^k$, such that

$$((f_h)_t, g)_\Omega + a_h(f_h, \mathbf{E}_h, \mathbf{B}_h; g) = 0 \tag{2.11a}$$

$$((\mathbf{E}_h)_t, \mathbf{U})_{\Omega_x} + ((\mathbf{B}_h)_t, \mathbf{W})_{\Omega_x} + b_h(\mathbf{E}_h, \mathbf{B}_h; \mathbf{U}, \mathbf{W}) = l_h(\mathbf{J}_h; \mathbf{U}), \tag{2.11b}$$

for any $g \in \mathcal{G}_h^k$, $\mathbf{U}, \mathbf{W} \in \mathcal{U}_h^k$, where

$$a_h(f_h, \mathbf{E}_h, \mathbf{B}_h; g) = a_{h,1}(f_h; g) + a_{h,2}(f_h, \mathbf{E}_h, \mathbf{B}_h; g), \quad l_h(\mathbf{J}_h; \mathbf{U}) = -\int_{\mathcal{T}_h^x} \mathbf{J}_h \cdot \mathbf{U}\, d\mathbf{x},$$

$$\begin{aligned} b_h(\mathbf{E}_h, \mathbf{B}_h; \mathbf{U}, \mathbf{W}) = &-\int_{\mathcal{T}_h^x} \mathbf{B}_h \cdot \nabla_{\mathbf{x}} \times \mathbf{U}\, d\mathbf{x} - \int_{\mathcal{E}_x} \widetilde{\mathbf{B}_h} \cdot [\mathbf{U}]_\tau \, ds_x \\ &+ \int_{\mathcal{T}_h^x} \mathbf{E}_h \cdot \nabla_{\mathbf{x}} \times \mathbf{W}\, d\mathbf{x} + \int_{\mathcal{E}_x} \widetilde{\mathbf{E}_h} \cdot [\mathbf{W}]_\tau \, ds_x, \end{aligned}$$

and

$$a_{h,1}(f_h; g) = -\int_{\mathcal{T}_h} f_h \mathbf{v} \cdot \nabla_{\mathbf{x}} g\, d\mathbf{x} d\mathbf{v} + \int_{\mathcal{T}_h^v} \int_{\mathcal{E}_x} \widetilde{f_h \mathbf{v}} \cdot [g]_x \, ds_x d\mathbf{v}$$

$$\begin{aligned} a_{h,2}(f_h, \mathbf{E}_h, \mathbf{B}_h; g) = &-\int_{\mathcal{T}_h} f_h \left(\mathbf{E}_h + \mathbf{v} \times \mathbf{B}_h\right) \cdot \nabla_{\mathbf{v}} g\, d\mathbf{x} d\mathbf{v} \\ &+ \int_{\mathcal{T}_h^x} \int_{\mathcal{E}_v} f_h (\widetilde{\mathbf{E}_h + \mathbf{v}} \times \mathbf{B}_h) \cdot [g]_v \, ds_v d\mathbf{x} \end{aligned}$$

The semi-discrete formulation (2.8) can then be solved by a numerical ODE solver, see the description in [19]. The L^2 and energy stability of (2.8) are established in [19]. The main result in [19] for the semi-discrete L^2 error estimates of the approximations f_h, $\mathbf{E}_h$, $\mathbf{B}_h$, is as follows.

Theorem 2.1.1 ([19]). *For $k \geq 2$ when $d_x = 3$ and $k \geq 1$ when $d_x = 1, 2$, the semi-discrete DG method of* (2.11a)-(2.11b)*, for the Vlasov-Maxwell equations with the upwind fluxes of* (2.10a)-(2.10d)*, has the following error estimate*

$$\|(f - f_h)(t)\|_{0,\Omega}^2 + \|(\mathbf{E} - \mathbf{E}_h)(t)\|_{0,\Omega_x}^2 + \|(\mathbf{B} - \mathbf{B}_h)(t)\|_{0,\Omega_x}^2 \leq C h^{2k+1}, \quad \forall t \in [0, T]. \quad (2.12)$$

Here the constant C is independent of h, but depends on the upper bounds of $\|\partial_t f\|_{k+1,\Omega}$, $\|f\|_{k+1,\Omega}$, $|f|_{1,\infty,\Omega}$, $\|\mathbf{E}\|_{1,\infty,\Omega_x}$, $\|\mathbf{B}\|_{1,\infty,\Omega_x}$, $\|\mathbf{E}\|_{k+1,\Omega_x}$, $\|\mathbf{B}\|_{k+1,\Omega_x}$ over the time interval $[0, T]$, and it also depends on the polynomial degree k, mesh parameters σ_0, σ_x and σ_v, and domain parameters L_x and L_v.

In this work, we also consider (1.1) when there is no magnetic field (i.e. when $\mathbf{B} = 0$). This reduced problem is called the VA system, and the DG discretizations would follow a similar discussion by setting $\mathbf{B}_h = 0$ in (2.8) at all times.

2.2 Smoothness-Increasing Accuracy-Conserving Filters

We extract the higher-order accuracy of the DG method solved over a uniform mesh contained in the negative-order norm by using the SIAC filter. This technique could also be applied over nonuniform meshes, however this would force us to compute the post-processing coefficients in each element in the mesh, increasing the computational complexity of the implementation [46]. This filter improves the order of accuracy by reducing the spurious oscillations in the error. This is done by convolving the numerical approximation with a specially chosen kernel,

$$(f_h^*(\mathbf{x},\mathbf{v}),\mathbf{E}_h^*(\mathbf{x}),\mathbf{B}_h^*(\mathbf{x})) = K_h^{2(k+1),k+1} \star (f_h,\mathbf{E}_h,\mathbf{B}_h)(\mathbf{x},\mathbf{v}), \tag{2.13}$$

where $(f_h^*,\mathbf{E}_h^*,\mathbf{B}_h^*)$ is the filtered solution, $(f_h,\mathbf{E}_h,\mathbf{B}_h)$ is an approximated solution computed at the final time, and $K_h^{2(k+1),k+1}$ is the convolution kernel. The kernel is translation-invariant and composed of a linear combination of B-splines of order $k+1$ obtained by convolving the characteristic function over the interval $(-\frac{1}{2},\frac{1}{2})$ with itself k times and scaled by the uniform mesh size. Using B-splines makes this kernel computationally efficient, provided the mesh is uniform, as the kernel is translation invariant and is locally supported in at most $2k+2$ elements. The one-dimensional convolution kernel is of the form:

$$K_h^{2(k+1),k+1}(x) = \frac{1}{h}\sum_{\gamma=-k}^{k} c_\gamma^{2(k+1),k+1}\psi^{(k+1)}\left(\frac{x}{h}-\gamma\right). \tag{2.14}$$

The weights of the B-splines, $c_\gamma^{2(k+1),k+1}$, are chosen so that accuracy is not destroyed (the kernel can reproduce polynomials of degree up to $2k$), i.e. $K_h^{2(k+1),k+1} \star p = p$ for $p = 1,\, x,\, \cdots, x^{2k}$, see [23] for details.

For the general case, assume the mesh size is uniform in each direction, given arbitrary $(\mathbf{x},\mathbf{v}) = (x_1,\cdots,x_{d_x},v_1,\cdots,v_{d_v}) \in \mathbb{R}^{d_x+d_v}$, we set

$$\psi^{(k+1)}(\mathbf{x},\mathbf{v}) = \prod_{i=1}^{d_x}\psi^{(k+1)}(x_i)\prod_{j=1}^{d_v}\psi^{(k+1)}(v_j) \tag{2.15}$$

The kernel for our case is of the form

$$K_h^{2(k+1),k+1}(\mathbf{x},\mathbf{v}) = \frac{1}{\left(\prod_{i=1}^{d_x} h_{x_i}\right)\left(\prod_{j=1}^{d_v} h_{v_j}\right)} \times \sum_{\gamma\in\{-k,\dots,k\}^{d_x+d_v}} \mathbf{c}_\gamma^{2(k+1),k+1}\psi^{(k+1)}\left(\left(\frac{x_1}{h_{x_1}},\cdots,\frac{x_{d_x}}{h_{d_x}},\frac{v_1}{h_{v_1}},\cdots,\frac{v_{d_v}}{h_{d_v}}\right)-\gamma\right) \quad (2.16)$$

where h_{x_i} and h_{v_i} denote the mesh size in x_i and v_i direction, resp. The success of the filter relies on the following results.

Theorem 2.2.1. *(Bramble and Schatz [22]) For $T>0$, let $u=(f,\mathbf{E},\mathbf{B})$ be the exact solution of the problem* (1.1)*. Let $\Omega_0 + 2supp(K_h^{2(k+1),k+1}(\mathbf{x},\mathbf{v})) \subset\subset \Omega$ and $U=(f_h,\mathbf{E}_h,\mathbf{B}_h)$ is any approximation to u, then*

$$\|u(T)-K_h^{2(k+1),k+1}\star U\|_{0,\Omega_0} \le \frac{h^{2k+2}}{(2k+2)!}|u|_{2k+2,\Omega} + C_{\mathrm{P}}\sum_{|\lambda|\le k+1}\|\partial_h^\lambda(u-U)\|_{-(k+1),\Omega}. \quad (2.17)$$

where C_{P} depends solely on $\Omega_0,\ \Omega,\ d_x,\ d_v,\ k, c_\gamma^{2(k+1),k+1}$, and it is independent of h.

In (2.17), we used the notation of the divided differences. We define

$$\partial_{h_{x_i}} w(\mathbf{x},\mathbf{v}) = \frac{1}{h_{x_i}}\left(w\left(\mathbf{x}+\frac{1}{2}h_{x_i}\mathbf{e}_i,\mathbf{v}\right)-w\left(\mathbf{x}-\frac{1}{2}h_{x_i}\mathbf{e}_i,\mathbf{v}\right)\right), \quad (2.18)$$

here $\mathbf{e}_i$ is the unit multi-index whose i-th component is 1 and all others 0. Analogously for velocity space variables v_j, the difference quotients are defined as

$$\partial_{h_{v_j}} w(\mathbf{x},\mathbf{v}) = \frac{1}{h_{v_j}}\left(w\left(\mathbf{x},\mathbf{v}+\frac{1}{2}h_{v_j}\mathbf{e}_j\right)-w\left(\mathbf{x},\mathbf{v}-\frac{1}{2}h_{v_j}\mathbf{e}_j\right)\right), \quad (2.19)$$

For any multi-index $\lambda=(\alpha_{x_1},\cdots,\alpha_{d_x},\beta_{v_1},\cdots,\beta_{d_v})$ we set λ-th order difference quotient to be

$$\partial_h^\lambda w(\mathbf{x},\mathbf{v}) = (\partial_{h_{x_1}}^{\alpha_1}\cdots\partial_{h_{x_{d_x}}}^{\alpha_{d_x}})(\partial_{h_{v_1}}^{\beta_1}\cdots\partial_{h_{v_{d_v}}}^{\beta_{d_v}})w(\mathbf{x},\mathbf{v}). \quad (2.20)$$

2.3 Superconvergent Error Estimates for the DG method

In this section, we prove the superconvergence error estimate in the negative norm of the DG solution for the VM system. In Section 2.3.1, we review basic approximation and regularity properties. Section 2.3.2 will construct the dual problem which is the key to our estimates. The main result and the proof will be given in Section 2.3.3.

2.3.1 Preliminaries

We summarize some of the standard approximation properties of the above discrete spaces, as well as some inverse inequalities [47]. For any nonnegative integer k, Let Π^k be the L^2 projection onto $\mathcal{G}_h^k$, and $\mathbf{\Pi}_x^m$ be the L^2 projection onto $\mathcal{U}_h^m$. We define $\zeta_h^g = \Pi^k g - g$ and $\zeta_h^{\mathbf{U}} = \mathbf{\Pi}_x^k \mathbf{U} - \mathbf{U}$, as the *Projection errors* of g and $\mathbf{U}$ respectively.

Lemma 2.3.1. *(Approximation properties) There exist a constant $C > 0$, such that for any $g \in H^{k+1}(\Omega)$ and $\mathbf{U} \in [H^{k+1}(\Omega)]^{d_x}$, the following hold:*

$$\begin{aligned}
\|\zeta_h^g\|_{0,K} + h_K \|\nabla_\star \zeta_h^g\|_{0,K} + h_K^{1/2} \|\zeta_h^g\|_{0,\partial K} &\le C h_K^{k+1} \|g\|_{k+1,K}, \quad \forall K \in \mathcal{T}_h \\
\|\zeta_h^{\mathbf{U}}\|_{0,K_x} + h_{K_x} \|\nabla_{\mathbf{x}} \times \zeta_h^{\mathbf{U}}\|_{0,K_x} + h_{K_x}^{1/2} \|\zeta_h^{\mathbf{U}}\|_{0,\partial K_x} &\le C h_{K_x}^{k+1} \|\mathbf{U}\|_{k+1,K_x}, \quad \forall K_x \in \mathcal{T}_h^x \\
\|\zeta_h^{\mathbf{U}}\|_{0,\infty,K_x} &\le C h_{K_x}^{k+1} \|\mathbf{U}\|_{k+1,\infty,K_x}, \quad \forall K_x \in \mathcal{T}_h^x
\end{aligned}$$

where the constant C is independent of the mesh sizes h_K and h_{K_x}, but depends on k and the shape regularity parameters σ_x and σ_v of the mesh. Here $\star = x$ or v.

Lemma 2.3.2 (Inverse inequality)**.** *There exists a constant $C > 0$, such that for any $g \in P^k(K)$ or $P^k(K_x) \times P^k(K_v)$ with $K = (K_x \times K_v) \in \mathcal{T}_h$, and for any $\mathbf{U} \in [P^k(K_x)]^{d_x}$, the following hold:*

$$\begin{aligned}
\|\nabla_{\mathbf{x}} g\|_{0,K} \le C h_{K_x}^{-1} \|g\|_{0,K}, \quad \|\nabla_{\mathbf{v}} g\|_{0,K} \le C h_{K_v}^{-1} \|g\|_{0,K}, \\
\|\mathbf{U}\|_{0,\infty,K_x} \le C h_{K_x}^{-d_x/2} \|\mathbf{U}\|_{0,K_x}, \quad \|\mathbf{U}\|_{0,\partial K_x} \le C h_{K_x}^{-1/2} \|\mathbf{U}\|_{0,K_x},
\end{aligned}$$

where the constant C is independent of the mesh sizes h_{K_x}, h_{K_v}, but depends on k and the shape regularity parameters σ_x and σ_v of the mesh.

To assist the proof, we also need a regularity result for a linear PDE system.

Lemma 2.3.3. *Consider the following system of equations with periodic boundary conditions in* $\mathbf{x}$ *and zero boundary condition in* $\mathbf{v}$ *for all* $t \in [0, T]$*:*

$$\partial_t \varphi + \mathbf{A_1}(\mathbf{x}, \mathbf{v}, t) \cdot \nabla_{\mathbf{x}} \varphi + \mathbf{A_2}(\mathbf{x}, \mathbf{v}, t) \cdot \nabla_{\mathbf{v}} \varphi + \mathbf{A_3}(\mathbf{x}, \mathbf{v}, t) \cdot \mathbf{F} = 0, \tag{2.21a}$$

$$\partial_t \mathbf{F} = \nabla_{\mathbf{x}} \times \mathbf{D} + \int_{\Omega_v} g \nabla_{\mathbf{v}} \varphi \, d\mathbf{v}, \tag{2.21b}$$

$$\partial_t \mathbf{D} = -\nabla_{\mathbf{x}} \times \mathbf{F} - \int_{\Omega_v} g(\mathbf{v} \times \nabla_{\mathbf{v}} \varphi) \, d\mathbf{v}, \tag{2.21c}$$

where the given functions $\mathbf{A_1}, \mathbf{A_2} \in W^{l+1,\infty}(\Omega)$ *satisfy the divergence free constraint* $\nabla_{\mathbf{x}} \cdot \mathbf{A_1} = 0$ *and* $\nabla_{\mathbf{v}} \cdot \mathbf{A_2} = 0$*. For any* $l \geq 0$ *and the fixed time* t*, the solution to* (2.21) *satisfy the following estimate*

$$\|\varphi(\cdot, \cdot, t)\|_{l,\Omega}^2 + \|\mathbf{F}(\cdot, t)\|_{l,\Omega_x}^2 + \|\mathbf{D}(\cdot, t)\|_{l,\Omega_x}^2 \leq C \left[\|\varphi(\cdot, \cdot, 0)\|_{l,\Omega}^2 + \|\mathbf{F}(\cdot, 0)\|_{l,\Omega_x}^2 + \|\mathbf{D}(\cdot, 0)\|_{l,\Omega_x}^2 \right]. \tag{2.22}$$

Here C *depends on* $\|\mathbf{A_3}\|_{L^\infty((0,T);W^{l,\infty}(\Omega))}$ *and* $\|g\|_{L^\infty((0,T);W^{l+1,\infty}(\Omega))}$*.*

Proof. See the appendix. □

2.3.2 The dual problem

In order to prove negative-order estimates for the system, the key is to find the dual problem associated to (1.1). We note that, for the nonlinear problem, the dual problem is not unique, see [48]. We construct the dual problem as follows: find functions $\varphi(\cdot, \cdot, t)$, $\mathbf{F}(\cdot, t)$ and $\mathbf{D}(\cdot, t)$ such that $\varphi(\cdot, \mathbf{v}, t)$ is periodic in all dimensions in space and $\varphi(\mathbf{x}, \cdot, t)$ vanishes in the boundary of the velocity region for all $t \in [0, T]$ and

$$\partial_t \varphi + \mathbf{v} \cdot \nabla_{\mathbf{x}} \varphi + (\mathbf{E} + \mathbf{v} \times \mathbf{B}) \cdot \nabla_{\mathbf{v}} \varphi - \mathbf{v} \cdot \mathbf{F} = 0 \tag{2.23a}$$

$$\partial_t \mathbf{F} = \nabla_{\mathbf{x}} \times \mathbf{D} - \int_{\Omega_v} f \nabla_{\mathbf{v}} \varphi \, d\mathbf{v}, \tag{2.23b}$$

$$\partial_t \mathbf{D} = -\nabla_{\mathbf{x}} \times \mathbf{F} + \int_{\Omega_v} f(\mathbf{v} \times \nabla_{\mathbf{v}} \varphi) \, d\mathbf{v} \tag{2.23c}$$

with final time conditions $\varphi(\mathbf{x}, \mathbf{v}, T) = \Phi(\mathbf{x})$, $\mathbf{F}(\mathbf{x}, T) = \mathfrak{F}(\mathbf{x})$ and $\mathbf{D}(\mathbf{x}, T) = \mathfrak{D}(\mathbf{x})$, $\Phi \in C_0^\infty(\Omega)$ and $\mathfrak{D}, \mathfrak{F} \in [C_0^\infty(\Omega_x)]^{d_x}$.

Notice that by multiplying $(\varphi, \mathbf{F}, \mathbf{D})$ on both sides of (1.1a)-(1.1b), and multiplying by $(f, \mathbf{E}, \mathbf{B})$ on both sides of (2.23a)-(2.23c), and then summing up and integrating over velocity and physical space, we obtain

$$\begin{aligned}
&\int_\Omega \partial_t(f\varphi)\, d\mathbf{x}\, d\mathbf{v} + \int_\Omega \nabla_x \cdot (f\varphi\mathbf{v})\, d\mathbf{x}\, d\mathbf{v} + \int_\Omega \nabla_{\mathbf{v}} \cdot (f\varphi(\mathbf{E} + \mathbf{v} \times \mathbf{B}))\, d\mathbf{x}\, d\mathbf{v} \\
&- \int_\Omega f\mathbf{v} \cdot \mathbf{F}\, d\mathbf{x}\, d\mathbf{v} = 0, \\
&\int_{\Omega_x} \partial_t(\mathbf{E} \cdot \mathbf{F} + \mathbf{B} \cdot \mathbf{D})\, d\mathbf{x} + \int_\Omega f\mathbf{v} \cdot \mathbf{F}\, d\mathbf{x}\, d\mathbf{v} \\
&= \int_{\Omega_x} \nabla_{\mathbf{x}} \cdot (\mathbf{B} \times \mathbf{F})\, d\mathbf{x} + \int_{\Omega_x} \nabla_{\mathbf{x}} \cdot (\mathbf{E} \times \mathbf{D})\, d\mathbf{x} - \int_\Omega f(\mathbf{E} + \mathbf{v} \times \mathbf{B}) \cdot \nabla_{\mathbf{v}}\varphi\, d\mathbf{x}d\mathbf{v},
\end{aligned}$$

where we used the identities

$$\nabla_\star(\phi\mathbf{U}) = \phi\nabla_\star \cdot \mathbf{U} + \mathbf{U} \cdot \nabla_\star\phi,$$

$$\nabla_\star \cdot (\mathbf{U} \times \mathbf{W}) = \mathbf{W} \cdot (\nabla_\star \times \mathbf{U}) - \mathbf{U} \cdot (\nabla_\star \times \mathbf{W}),$$

for scalar functions ϕ and vector functions $\mathbf{U}$ and $\mathbf{W}$ and the fact that $\nabla_{\mathbf{v}} \cdot (\mathbf{E} + \mathbf{v} \times \mathbf{B}) = \mathbf{0}$.

By adding all equations above and using boundary conditions, we arrive at

$$\frac{d}{dt}[(f, \varphi)_\Omega + (\mathbf{E}, \mathbf{F})_{\Omega_x} + (\mathbf{B}, \mathbf{D})_{\Omega_x}] + \mathcal{F}(f, \mathbf{E}, \mathbf{B}; \varphi) = 0, \tag{2.24}$$

where

$$\mathcal{F}(f, \mathbf{E}, \mathbf{B}; \varphi) = \int_\Omega f(\mathbf{E} + \mathbf{v} \times \mathbf{B}) \cdot \nabla_{\mathbf{v}}\varphi\, d\mathbf{x}d\mathbf{v}. \tag{2.25}$$

2.3.3 The main result

In this part, we give our main theorem on the negative-norm of the error for the DG solutions. Note that superconvergence of the negative norm of the solution itself is not sufficient in proving high order convergence of the post-processed solution according to Theorem 2.2.1. However, it is a necessary first step. As shown in [25], it is highly non-trivial to prove superconvergence of the divided difference of the solution for nonlinear problems, we will leave this to explore in our future work.

Theorem 2.3.1. *If $(f_h, \mathbf{E}_h, \mathbf{B}_h)$ is a solution to* (2.11a)-(2.11b) *with the numerical initial condition $f_h = \Pi^k f$ and $\mathbf{E}_h = \mathbf{\Pi}_x^k \mathbf{E}$, $\mathbf{B}_h = \mathbf{\Pi}_x^k \mathbf{B}$ and $k \geq (d_x + d_v)/2$, then*

$$\|(f - f_h, \mathbf{E} - \mathbf{E}_h, \mathbf{B} - \mathbf{B}_h)\|_{-(k+1),\Omega} \leq C h^{2k+1/2},$$

where C is a constant independent of h and depends on the upper bounds of $\|\partial_t f\|_{k+2,\Omega}$, $\|f\|_{k+2,\Omega}$, $|f|_{1,\infty,\Omega}$, $\|\mathbf{E}\|_{1,\infty,\Omega_x}$, $\|\mathbf{B}\|_{1,\infty,\Omega_x}$, $\|\mathbf{E}\|_{k+2,\Omega_x}$, $\|\mathbf{B}\|_{k+2,\Omega_x}$ over the time interval $[0, T]$, and it also depends on the polynomial degree k, mesh parameters σ_0, σ_x and σ_v, and domain parameters L_x and L_v.

Proof. We define $e_h^f = f - f_h = \epsilon_h^f - \zeta_h^f$, where $\epsilon_h^f = \Pi^k f - f_h$ and ζ_h^f is defined just as in Section 2.3.1. Analogously $\epsilon_h^{\mathbf{E}} = \mathbf{\Pi}_x^k \mathbf{E} - \mathbf{E}_h$, $\epsilon_h^{\mathbf{B}} = \mathbf{\Pi}_x^k \mathbf{B} - \mathbf{B}_h$, then $e_h^{\mathbf{E}} = \mathbf{E} - \mathbf{E}_h = \epsilon_h^{\mathbf{E}} - \zeta_h^{\mathbf{E}}$ and $e_h^{\mathbf{B}} = \mathbf{B} - \mathbf{B}_h = \epsilon_h^{\mathbf{B}} - \zeta_h^{\mathbf{B}}$. We follow the ideas in [23]. For any $\Phi \in C_0^\infty(\Omega), \mathfrak{F}, \mathfrak{D} \in [C_0^\infty(\Omega_x)]^{d_x}$, we estimate the term

$$\begin{aligned}
&(e_h^f(T), \Phi)_\Omega + (e_h^{\mathbf{E}}(T), \mathfrak{F})_{\Omega_x} + (e_h^{\mathbf{B}}(T), \mathfrak{D})_{\Omega_x} \\
=&(e_h^f(T), \varphi(T))_\Omega + (e_h^{\mathbf{E}}(T), \mathbf{F}(T))_{\Omega_x} + (e_h^{\mathbf{B}}(T), \mathbf{D}(T))_{\Omega_x} \\
=&(f(T), \varphi(T))_\Omega + (\mathbf{E}(T), \mathbf{F}(T))_{\Omega_x} + (\mathbf{B}(T), \mathbf{D}(T))_{\Omega_x} \\
&- [(f_h(T), \varphi(T))_\Omega + (\mathbf{E}_h(T), \mathbf{F}(T))_{\Omega_x} + (\mathbf{B}_h(T), \mathbf{D}(T))_{\Omega_x}] \\
=&(f_0, \varphi(0))_\Omega + (\mathbf{E}_0, \mathbf{F}(0))_{\Omega_x} + (\mathbf{B}_0, \mathbf{D}(0))_{\Omega_x} - \int_0^T \mathcal{F}(f, \mathbf{E}, \mathbf{B}; \varphi)\, d\tau \\
&- (f_h(0), \varphi(0))_\Omega - (\mathbf{E}_h(0), \mathbf{F}(0))_{\Omega_x} - (\mathbf{B}_h(0), \mathbf{D}(0))_{\Omega_x} \\
&- \int_0^T \frac{d}{dt}[(f_h, \varphi)_\Omega + (\mathbf{E}_h, \mathbf{F})_{\Omega_x} + (\mathbf{B}_h, \mathbf{D})_{\Omega_x}]\, d\tau \\
=&-\left[(\zeta_h^{f_0}, \varphi(0))_\Omega + (\zeta_h^{\mathbf{E}_0}, \mathbf{F}(0))_{\Omega_x} + (\zeta_h^{\mathbf{B}}, \mathbf{D}(0))_{\Omega_x}\right] \\
&- \int_0^T ((f_h)_t, \varphi)_\Omega + ((\mathbf{E}_h)_t, \mathbf{F})_{\Omega_x} + ((\mathbf{B}_h)_t, \mathbf{D})_{\Omega_x}\, d\tau \\
&- \int_0^T (f_h, \varphi_t)_\Omega + (\mathbf{E}_h, \mathbf{F}_t)_{\Omega_x} + (\mathbf{B}_h, \mathbf{D}_t)_{\Omega_x} + \mathcal{F}(f, \mathbf{E}, \mathbf{B}; \varphi)\, d\tau,
\end{aligned}$$

where for the first equality we used (2.24), and the numerical initial condition is used in the last equality.

Notice that for any $\chi \in \mathcal{G}_h^k$, $\xi, \eta \in \mathcal{U}_h^k$

$$
\begin{aligned}
&\int_0^T ((f_h)_t, \varphi)_\Omega + ((\mathbf{E}_h)_t, \mathbf{F})_{\Omega_x} + ((\mathbf{B}_h)_t, \mathbf{D})_{\Omega_x}\, d\tau \\
&= \int_0^T ((f_h)_t, \varphi - \chi)_\Omega\, d\tau + \int_0^T ((f_h)_t, \chi)_\Omega\, d\tau + \int_0^T ((\mathbf{E}_h)_t, \mathbf{F} - \xi)_{\Omega_x} + ((\mathbf{B}_h)_t, \mathbf{D} - \eta)_{\Omega_x}\, d\tau \\
&\quad + \int_0^T ((\mathbf{E}_h)_t, \xi)_{\Omega_x} + ((\mathbf{B}_h)_t, \eta)_{\Omega_x}\, d\tau \\
&= \int_0^T ((f_h)_t, \varphi - \chi)_\Omega\, d\tau - \int_0^T a_h(f_h, \mathbf{E}_h, \mathbf{B}_h; \chi)\, d\tau \\
&\quad + \int_0^T ((\mathbf{E}_h)_t, \mathbf{F} - \xi)_{\Omega_x} + ((\mathbf{B}_h)_t, \mathbf{D} - \eta)_{\Omega_x}\, d\tau \\
&\quad - \int_0^T b_h(\mathbf{E}_h, \mathbf{B}_h; \xi, \eta) - l_h(\mathbf{J}_h, \xi)\, d\tau \\
&= \int_0^T ((f_h)_t, \varphi - \chi)_\Omega + a_h(f_h, \mathbf{E}_h, \mathbf{B}_h; \varphi - \chi)\, d\tau + \int_0^T ((\mathbf{E}_h)_t, \mathbf{F} - \xi)_{\Omega_x} + ((\mathbf{B}_h)_t, \mathbf{D} - \eta)_{\Omega_x}\, d\tau \\
&\quad + \int_0^T b_h(\mathbf{E}_h, \mathbf{B}_h; \mathbf{F} - \xi, \mathbf{D} - \eta) - l_h(\mathbf{J}_h, \mathbf{F} - \xi)\, d\tau - \int_0^T a_h(f_h, \mathbf{E}_h, \mathbf{B}_h; \varphi)\, d\tau \\
&\quad - \int_0^T b_h(\mathbf{E}_h, \mathbf{B}_h; \mathbf{F}, \mathbf{D}) - l_h(\mathbf{J}_h, \mathbf{F})\, d\tau.
\end{aligned}
$$

After this calculation we can conclude that

$$
(e_h^f(T), \Phi)_\Omega + (e_h^{\mathbf{E}}(T), \mathfrak{F})_{\Omega_x} + (e_h^{\mathbf{B}}(T), \mathfrak{D})_{\Omega_x} = \Theta_{\mathrm{M}} + \Theta_{\mathrm{N}} + \Theta_{\mathrm{C}}, \tag{2.26}
$$

where

$$
\begin{aligned}
\Theta_{\mathrm{M}} &= -\left[(\zeta_h^{f_0}, \varphi(0))_\Omega + (\zeta_h^{\mathbf{E}_0}, \mathbf{F}(0))_{\Omega_x} + (\zeta_h^{\mathbf{B}_0}, \mathbf{D}(0))_{\Omega_x}\right], \\
\Theta_{\mathrm{N}} &= -\int_0^T ((f_h)_t, \varphi - \chi)_\Omega + a_h(f_h, \mathbf{E}_h, \mathbf{B}_h; \varphi - \chi)\, d\tau \\
&\quad - \int_0^T ((\mathbf{E}_h)_t, \mathbf{F} - \xi)_{\Omega_x} + ((\mathbf{B}_h)_t, \mathbf{D} - \eta)_{\Omega_x} + b_h(\mathbf{E}_h, \mathbf{B}_h; \mathbf{F} - \xi, \mathbf{D} - \eta) - l_h(\mathbf{J}_h, \mathbf{F} - \xi)\, d\tau, \\
\Theta_{\mathrm{C}} &= -\int_0^T (f_h, \varphi_t)_\Omega - a_h(f_h, \mathbf{E}_h, \mathbf{B}_h; \varphi)\, d\tau \\
&\quad - \int_0^T (\mathbf{E}_h, \mathbf{F}_t)_{\Omega_x} + (\mathbf{B}_h, \mathbf{D}_t)_{\Omega_x} - b_h(\mathbf{E}_h, \mathbf{B}_h; \mathbf{F}, \mathbf{D}) + l_h(\mathbf{J}_h, \mathbf{F})\, d\tau - \int_0^T \mathcal{F}(f, \mathbf{E}, \mathbf{B}; \varphi)\, d\tau.
\end{aligned}
$$

In the following we will estimate Θ_{M}, Θ_{N} and Θ_{C}.

Lemma 2.3.4 (Projection Estimate)**.** Θ_{M} *satisfies*

$$|\Theta_{\mathrm{M}}| \leq Ch^{2k+2}\sqrt{\|\varphi(0)\|_{k+1,\Omega}^2 + \|\mathbf{F}(0)\|_{k+1,\Omega_x}^2 + \|\mathbf{D}(0)\|_{k+1,\Omega_x}^2} \tag{2.28}$$

where C *depends on* $\|f_0\|_{k+1,\Omega}$, $\|\mathbf{E}_0\|_{k+1,\Omega_x}$ *and* $\|\mathbf{B}_0\|_{k+1,\Omega_x}$.

Proof. By the definition of Π^k,

$$\begin{aligned}(f_0 - \Pi^k f_0, \varphi(0))_\Omega &= (f_0 - \Pi^k f_0, \varphi(0) - \Pi^k\varphi(0))_\Omega \\ &\leq \|f_0 - \Pi^k f_0\|\|\varphi(0) - \Pi^k\varphi(0)\| \\ &\leq Ch^{k+1}\|f_0\|_{k+1,\Omega}h^{k+1}\|\varphi(0)\|_{k+1,\Omega}.\end{aligned}$$

The last line was an application of the first part of Lemma 2.3.1. By the same lines we obtain analogous results for the $\mathbf{E}$ and $\mathbf{B}$ parts. The conclusion follows by grouping them all together and an application of Cauchy-Schwarz inequality. □

For the second term, we have the following result:

Lemma 2.3.5 (Residual)**.** *Let* $\chi = \Pi^k f, \xi = \mathbf{\Pi}_x^k\mathbf{F}, \eta = \mathbf{\Pi}_x^k\mathbf{D}$*, we have*

$$|\Theta_{\mathrm{N}}| \leq Ch^{2k+1/2}\left[\int_0^T \|\varphi\|_{k+1,\Omega}^2 + \|\mathbf{F}\|_{k+1,\Omega_x}^2 + \|\mathbf{D}\|_{k+1,\Omega_x}^2\,dt\right]^{1/2}$$

where C *depends on the upper bounds of* $\|f\|_{k+2,\Omega}$, $\|f\|_{1,\infty,\Omega}$, $\|\mathbf{E}\|_{0,\infty,\Omega_x}$, $\|\mathbf{B}\|_{0,\infty,\Omega_x}$, $\|\mathbf{E}\|_{k+2,\Omega_x}$, $\|\mathbf{B}\|_{k+2,\Omega_x}$ *over the time interval* $[0,T]$*, and it also depends on the polynomial degree* k*, mesh parameters* σ_0*,* σ_x *and* σ_v*, and domain parameters* L_x *and* L_v*.*

Proof. Due to the definition of the projection operators, $((f_h)_t, \varphi - \chi)_\Omega = 0$, $((\mathbf{E}_h)_t, \mathbf{F} - \xi)_{\Omega_x} = 0$, and $((\mathbf{B}_h)_t, \mathbf{D} - \eta)_{\Omega_x} = 0$, and $l_h(\mathbf{J}_h; \mathbf{F} - \xi) = -(\mathbf{J}_h, \mathbf{F} - \xi)_{\Omega_x} = 0$, we have

$$\Theta_{\mathrm{N}} = \int_0^T -a_h(f_h, \mathbf{E}_h, \mathbf{B}_h; \zeta_h^\varphi) - b_h(\mathbf{E}_h, \mathbf{B}_h; \zeta_h^{\mathbf{F}}, \zeta_h^{\mathbf{D}})\,d\tau.$$

From its definition,

$$\begin{aligned} b_h(\mathbf{E}_h, \mathbf{B}_h; \zeta_h^{\mathbf{F}}, \zeta_h^{\mathbf{D}}) &= \int_{\mathcal{T}_h^x} \mathbf{E}_h \cdot \nabla_{\mathbf{x}} \times \zeta_h^{\mathbf{D}}\, d\mathbf{x} - \int_{\mathcal{T}_h^x} \mathbf{B}_h \cdot \nabla_{\mathbf{x}} \times \zeta_h^{\mathbf{F}}\, d\mathbf{x} \\ &+ \int_{\mathcal{E}_x} \widetilde{\mathbf{E}_h} \cdot [\zeta_h^{\mathbf{D}}]_\tau\, ds_{\mathbf{x}} - \int_{\mathcal{E}_x} \widetilde{\mathbf{B}_h} \cdot [\zeta_h^{\mathbf{F}}]_\tau\, ds_{\mathbf{x}} \\ &= -\int_{\mathcal{T}_h^x} e_h^{\mathbf{E}} \cdot \nabla_{\mathbf{x}} \times \zeta_h^{\mathbf{D}}\, d\mathbf{x} + \int_{\mathcal{T}_h^x} e_h^{\mathbf{B}} \cdot \nabla_{\mathbf{x}} \times \zeta_h^{\mathbf{F}}\, d\mathbf{x} \\ &- \int_{\mathcal{E}_x} \widetilde{e_h^{\mathbf{E}}} \cdot [\zeta_h^{\mathbf{D}}]_\tau\, ds_{\mathbf{x}} + \int_{\mathcal{E}_x} \widetilde{e_h^{\mathbf{B}}} \cdot [\zeta_h^{\mathbf{F}}]_\tau\, ds_{\mathbf{x}} \\ &+ \int_{\mathcal{T}_h^x} (\nabla_{\mathbf{x}} \times \mathbf{E}) \cdot \zeta_h^{\mathbf{D}}\, d\mathbf{x} - \int_{\mathcal{T}_h^x} (\nabla_{\mathbf{x}} \times \mathbf{B}) \cdot \zeta_h^{\mathbf{F}}\, d\mathbf{x}. \end{aligned}$$

By Lemma 2.3.1,

$$\left| \int_{\mathcal{T}_h^x} (e_h^{\mathbf{E}}) \cdot \nabla_{\mathbf{x}} \times \zeta_h^{\mathbf{D}}\, d\mathbf{x} \right| \leq C h^k \|e_h^{\mathbf{E}}\|_{0,\Omega_x} \|\mathbf{D}\|_{k+1,\Omega_x},$$

$$\left| \int_{\mathcal{T}_h^x} e_h^{\mathbf{B}} \cdot \nabla_{\mathbf{x}} \times \zeta_h^{\mathbf{F}}\, d\mathbf{x} \right| \leq C h^k \|e_h^{\mathbf{B}}\|_{0,\Omega_x} \|\mathbf{F}\|_{k+1,\Omega_x},$$

$$\begin{aligned} \left| \int_{\mathcal{E}_x} (\widetilde{e_h^{\mathbf{E}}}) \cdot [\zeta_h^{\mathbf{D}}]_\tau - (\widetilde{e_h^{\mathbf{B}}}) \cdot [\zeta_h^{\mathbf{F}}]_\tau\, ds_{\mathbf{x}} \right| &\leq C h^{k+1/2} \left(\|\mathbf{D}\|_{k+1,\Omega_x} + \|\mathbf{F}\|_{k+1,\Omega_x} \right) \\ &\times \left(\|e_h^{\mathbf{E}}\|_{0,\mathcal{E}_x} + \|e_h^{\mathbf{B}}\|_{0,\mathcal{E}_x} \right). \end{aligned}$$

Now notice that

$$\begin{aligned} \|e_h^{\mathbf{E}}\|_{0,\mathcal{E}_x} &\leq \|\epsilon_h^{\mathbf{E}}\|_{0,\mathcal{E}_x} + \|\zeta_h^{\mathbf{E}}\|_{0,\mathcal{E}_x} \\ &\leq C[h^{-1/2}\|\epsilon_h^{\mathbf{E}}\|_{0,\Omega_x} + h^{k+1/2}] \\ &\leq C h^{-1/2}[\|e_h^{\mathbf{E}}\|_{0,\Omega_x} + h^{k+1}]. \end{aligned}$$

Analogously

$$\|e_h^{\mathbf{B}}\|_{0,\mathcal{E}_x} \leq C h^{-1/2}[\|e_h^{\mathbf{B}}\|_{0,\Omega_x} + h^{k+1}].$$

Therefore,

$$\begin{aligned} \left| \int_{\mathcal{E}_x} (\widetilde{e_h^{\mathbf{E}}}) \cdot [\zeta_h^{\mathbf{D}}]_\tau - (\widetilde{e_h^{\mathbf{B}}}) \cdot [\zeta_h^{\mathbf{F}}]_\tau\, ds_{\mathbf{x}} \right| &\leq C h^k \left(\|\mathbf{D}\|_{k+1,\Omega_x} + \|\mathbf{F}\|_{k+1,\Omega_x} \right) \\ &\times \left(\|e_h^{\mathbf{E}}\|_{0,\Omega_x} + \|e_h^{\mathbf{B}}\|_{0,\Omega_x} + h^{k+1} \right). \end{aligned}$$

Now by the properties of the orthogonal projection $\mathbf{\Pi}_x^k$

$$\left|\int_{\mathcal{T}_h^x}(\nabla_{\mathbf{x}}\times\mathbf{E})\cdot\zeta_h^{\mathbf{D}}\,d\mathbf{x}\right| = \left|\int_{\mathcal{T}_h^x}(\nabla_{\mathbf{x}}\times\mathbf{E}-\mathbf{\Pi}_x^k(\nabla_{\mathbf{x}}\times\mathbf{E}))\cdot\zeta_h^{\mathbf{D}}\,d\mathbf{x}\right| \leq Ch^{2k+2}\|\mathbf{D}\|_{k+1,\Omega_x},$$

where C depends on $\|\mathbf{E}\|_{k+2,\Omega_x}$. By an analogous procedure

$$\left|\int_{\mathcal{T}_h^x}(\nabla_{\mathbf{x}}\times\mathbf{B})\cdot\zeta_h^{\mathbf{F}}\,d\mathbf{x}\right| \leq Ch^{2k+2}\|\mathbf{F}\|_{k+1,\Omega_x},$$

where C depends on $\|\mathbf{B}\|_{k+2,\Omega_x}$. Putting all the above calculations together, we arrive at,

$$|b_h(\mathbf{E}_h,\mathbf{B}_h;\zeta_h^F,\zeta_h^D)| \leq Ch^k\left(\|\mathbf{D}\|_{k+1,\Omega_x}+\|\mathbf{F}\|_{k+1,\Omega_x}\right)\left(\|e_h^{\mathbf{E}}\|_{0,\Omega_x}+\|e_h^{\mathbf{B}}\|_{0,\Omega_x}+h^{k+1}\right), \quad (2.29)$$

where C depends on $\|\mathbf{E}\|_{k+2,\Omega_x}$, $\|\mathbf{B}\|_{k+2,\Omega_x}$.

We will deal now with the term a_h, which is

$$a_h(f_h,\mathbf{E}_h,\mathbf{B}_h,\zeta_h^{\varphi}) = a_{h,1}(f_h,\zeta_h^{\varphi})+a_{h,2}(f_h,\mathbf{E}_h,\mathbf{B}_h;\zeta_h^{\varphi}). \quad (2.30)$$

First, we have

$$a_{h,1}(f_h;\zeta_h^{\varphi}) = \int_{\mathcal{T}_h} e_h^f\mathbf{v}\cdot\nabla_{\mathbf{x}}\zeta_h^{\varphi}\,d\mathbf{x}d\mathbf{v}+\int_{\mathcal{T}_h^v}\int_{\mathcal{E}_x}\widetilde{e_h^f\mathbf{v}}[\zeta_h^{\varphi}]_x\,ds_{\mathbf{x}}d\mathbf{v}-\int_{\mathcal{T}_h}\nabla_{\mathbf{x}}f\cdot\mathbf{v}\zeta_h^{\varphi}\,d\mathbf{x}d\mathbf{v}$$

The first term can be easily bounded, by using Lemma 2.3.1.

$$\left|\int_{\mathcal{T}_h} e_h^f\mathbf{v}\cdot\nabla_{\mathbf{x}}\zeta_h^{\varphi}\,d\mathbf{x}d\mathbf{v}\right| \leq Ch^k\|e_h^f\|_{0,\Omega}\|\varphi\|_{k+1,\Omega}.$$

Similarly,

$$\begin{aligned}\left|\int_{\mathcal{T}_h^v}\int_{\mathcal{E}_x}\widetilde{e_h^f\mathbf{v}}[\zeta_h^{\varphi}]_x\,ds_{\mathbf{x}}d\mathbf{v}\right| &\leq C\|e_h^f\|_{\mathcal{T}_h^v\times\mathcal{E}_x}\|\zeta_h^{\varphi}\|_{\mathcal{T}_h^v\times\mathcal{E}_x}\\ &\leq Ch^{k+1/2}\|e_h^f\|_{\mathcal{T}_h^v\times\mathcal{E}_x}\|\varphi\|_{k+1,\Omega}\\ &\leq Ch^{k+1/2}(\|\epsilon_h^f\|_{\mathcal{T}_h^v\times\mathcal{E}_x}+\|\zeta_h^f\|_{\mathcal{T}_h^v\times\mathcal{E}_x})\|\varphi\|_{k+1,\Omega}\\ &\leq Ch^k(\|e_h^f\|_{0,\Omega}+h^{k+1})\|\varphi\|_{k+1,\Omega}.\end{aligned}$$

For the last term notice that by the properties of the projection Π^k and the fact that $\Pi^k(\nabla_{\mathbf{x}}f\cdot\mathbf{v})$ is a polynomial of degree k,

$$\begin{aligned}\int_{\mathcal{T}_h}\nabla_{\mathbf{x}}f\cdot\mathbf{v}\zeta_h^{\varphi}\,d\mathbf{x}d\mathbf{v} &= \int_{\mathcal{T}_h}(\nabla_{\mathbf{x}}f\cdot\mathbf{v}-\Pi^k(\nabla_{\mathbf{x}}f\cdot\mathbf{v}))\zeta_h^{\varphi}\,d\mathbf{x}d\mathbf{v}\\ &\leq Ch^{2k+2}\|\varphi\|_{k+1,\Omega},\end{aligned}$$

where C depends on $\|f\|_{k+2,\Omega}$. By using all the calculations above, we can conclude that

$$|a_{h,1}(f_h;\zeta_h^\varphi)| \leq Ch^k\|e_h^f\|_{0,\Omega}\|\varphi\|_{k+1,\Omega} + Ch^{2k+1}\|\varphi\|_{k+1,\Omega}, \tag{2.31}$$

where C depends on $\|f\|_{k+2,\Omega}$. To conclude our proof, we only need to bound $a_{h,2}$, this time we will do things a little bit different, notice that

$$a_{h,2}(f_h,\mathbf{E}_h,\mathbf{B}_h,\zeta_h^\varphi) = a_{h,2}(f,\mathbf{E}_h,\mathbf{B}_h,\zeta_h^\varphi) - a_{h,2}(e_h^f,\mathbf{E}_h,\mathbf{B}_h,\zeta_h^\varphi).$$

We will get started by noting that $f(\widetilde{\mathbf{E}_h + \mathbf{v}\times\mathbf{B}_h}) = f\{\mathbf{E}_h + \mathbf{v}\times\mathbf{B}_h\}_v = f\left(\mathbf{E}_h + \mathbf{v}\times\mathbf{B}_h\right)$, then

$$\begin{aligned}
a_{h,2}(f,\mathbf{E}_h,\mathbf{B}_h,\zeta_h^\varphi) &= -\int_{\mathcal{T}_h} f(\mathbf{E}_h + \mathbf{v}\times\mathbf{B}_h)\cdot\nabla_{\mathbf{v}}\zeta_h^\varphi\, d\mathbf{x}d\mathbf{v} \\
&\quad + \int_{\mathcal{T}_h^x}\int_{\mathcal{E}_v} f(\mathbf{E}_h + \mathbf{v}\times\mathbf{B}_h)\cdot[\zeta_h^\varphi]_v\, d\mathbf{x}d\mathbf{v} \\
&= \int_{\mathcal{T}_h} f(e_h^{\mathbf{E}} + \mathbf{v}\times e_h^{\mathbf{B}})\cdot\nabla_{\mathbf{v}}\zeta_h^\varphi\, d\mathbf{x}d\mathbf{v} - \int_{\mathcal{T}_h^x}\int_{\mathcal{E}_v} f(e_h^{\mathbf{E}} + \mathbf{v}\times e_h^{\mathbf{B}})\cdot[\zeta_h^\varphi]_v\, d\mathbf{x}d\mathbf{v} \\
&\quad + \int_{\mathcal{T}_h} \nabla_{\mathbf{v}} f\cdot(\mathbf{E} + \mathbf{v}\times\mathbf{B})\zeta_h^\varphi\, d\mathbf{x}d\mathbf{v}.
\end{aligned}$$

We obtained the last inequality by adding and subtracting $\int_{\mathcal{T}_h} f(\mathbf{E}+\mathbf{v}\times\mathbf{B})\cdot\nabla_{\mathbf{v}}\zeta_h^\varphi\, d\mathbf{x}d\mathbf{v}$, integration by parts, and the fact that $\nabla_{\mathbf{v}}\cdot(\mathbf{E}+\mathbf{v}\times\mathbf{B}) = 0$. in this way

$$\left|\int_{\mathcal{T}_h} f(e_h^{\mathbf{E}} + \mathbf{v}\times e_h^{\mathbf{B}})\cdot\nabla_{\mathbf{v}}\zeta_h^\varphi\, d\mathbf{x}d\mathbf{v}\right| \leq Ch^k(\|e_h^{\mathbf{E}}\|_{0,\Omega_x} + \|e_h^{\mathbf{B}}\|_{0,\Omega_x})\|\varphi\|_{k+1,\Omega},$$

and

$$\left|\int_{\mathcal{T}_h^x}\int_{\mathcal{E}_v} f(e_h^{\mathbf{E}} + \mathbf{v}\times e_h^{\mathbf{B}})\cdot[\zeta_h^\varphi]_v\, d\mathbf{v}d\mathbf{x}\right| \leq Ch^{k+1/2}(\|e_h^{\mathbf{E}}\|_{0,\Omega_x} + \|e_h^{\mathbf{B}}\|_{0,\Omega_x})\|\varphi\|_{k+1,\Omega}.$$

Last but not least by the same arguments as previous estimates

$$\begin{aligned}
\int_{\mathcal{T}_h} \nabla_{\mathbf{v}} f\cdot(\mathbf{E}+\mathbf{v}\times\mathbf{B})\zeta_h^\varphi\, d\mathbf{x}d\mathbf{v} &= \int_{\mathcal{T}_h} (\nabla_{\mathbf{v}} f\cdot(\mathbf{E}+\mathbf{v}\times\mathbf{B}) - \Pi^k\nabla_{\mathbf{v}} f\cdot(\mathbf{E}+\mathbf{v}\times\mathbf{B}))\zeta_h^\varphi\, d\mathbf{x}d\mathbf{v} \\
&\leq Ch^{2k+2}\|\varphi\|_{k+1,\Omega},
\end{aligned}$$

where C depends on $\|f\|_{k+2,\Omega}, \|\mathbf{E}\|_{k+1,\Omega_x}, \|\mathbf{B}\|_{k+1,\Omega_x}$. We can conclude that

$$|a_{h,2}(f,\mathbf{E}_h,\mathbf{B}_h;\zeta_h^\varphi)| \leq Ch^k(\|e_h^{\mathbf{E}}\|_{0,\Omega_x} + \|e_h^{\mathbf{B}}\|_{0,\Omega_x})\|\varphi\|_{k+1,\Omega} + Ch^{2k+2}\|\varphi\|_{k+1,\Omega}. \tag{2.32}$$

Finally we just need to estimate

$$a_{h,2}(e_h^f, \mathbf{E}_h, \mathbf{B}_h; \zeta_h^\varphi) = -\int_{\mathcal{T}_h} e_h^f \left(\mathbf{E}_h + \mathbf{v} \times \mathbf{B}_h\right) \cdot \nabla_{\mathbf{v}} \zeta_h^\varphi \, d\mathbf{x} d\mathbf{v} + \int_{\mathcal{T}_h^x} \int_{\mathcal{E}_v} e_h^f (\widetilde{\mathbf{E}_h + \mathbf{v}} \times \mathbf{B}_h) \cdot [\zeta_h^\varphi]_v \, ds_v d\mathbf{x}$$

We have

$$\left| \int_{\mathcal{T}_h} e_h^f \left(\mathbf{E}_h + \mathbf{v} \times \mathbf{B}_h\right) \cdot \nabla_{\mathbf{v}} \zeta_h^\varphi \, d\mathbf{x} d\mathbf{v} \right| \leq C h^k \|e_h^f\|_{0,\Omega} (\|\mathbf{E}_h\|_{0,\infty,\Omega_x} + \|\mathbf{B}_h\|_{0,\infty,\Omega_x}) \|\varphi\|_{k+1,\Omega}$$

$$\leq C h^k \|e_h^f\|_{0,\Omega} (\|\epsilon_h^{\mathbf{E}}\|_{0,\infty,\Omega_x} + \|\epsilon_h^{\mathbf{B}}\|_{0,\infty,\Omega_x} + \|\Pi_x^k \mathbf{E}\|_{0,\infty,\Omega_x} + \|\Pi_x^k \mathbf{B}\|_{0,\infty,\Omega_x}) \|\varphi\|_{k+1,\Omega}$$

$$\leq C h^{k-d_x/2} \|e_h^f\|_{0,\Omega} (\|\epsilon_h^{\mathbf{E}}\|_{0,\Omega_x} + \|\epsilon_h^{\mathbf{B}}\|_{0,\Omega_x}) \|\varphi\|_{k+1,\Omega}$$

$$+ C h^k \|e_h^f\|_{0,\Omega} (\|\mathbf{E}\|_{0,\infty,\Omega_x} + \|\mathbf{B}\|_{0,\infty,\Omega_x}) \|\varphi\|_{k+1,\Omega}$$

$$\leq C h^{k-d_x/2} \|e_h^f\|_{0,\Omega} (\|e_h^{\mathbf{E}}\|_{0,\Omega_x} + \|e_h^{\mathbf{B}}\|_{0,\Omega_x} + h^{k+1}) \|\varphi\|_{k+1,\Omega} + C h^k \|e_h^f\|_{0,\Omega} \|\varphi\|_{k+1,\Omega}$$

$$\leq C h^k \|e_h^f\|_{0,\Omega} (h^{-d_x/2} \|e_h^{\mathbf{E}}\|_{0,\Omega_x} + h^{-d_x/2} \|e_h^{\mathbf{B}}\|_{0,\Omega_x} + 1) \|\varphi\|_{k+1,\Omega},$$

Here we used the fact that whenever $d_x = 1,\ 2,\ 3$, $k + 1 - d_x/2 > 0$, Lemma 2.3.2 and the fact that Π_x is bounded in any L^p-norm $(1 \leq p \leq \infty)$ [49, 50],

$$\|\Pi_x \mathbf{E}\|_{0,\infty,\Omega_x} \leq C \|\mathbf{E}\|_{0,\infty,\Omega_x}, \|\Pi_x \mathbf{B}\|_{0,\infty,\Omega_x} \leq C \|\mathbf{B}\|_{0,\infty,\Omega_x}.$$

Finally

$$\int_{\mathcal{T}_h^x} \int_{\mathcal{E}_v} e_h^f (\widetilde{\mathbf{E}_h + \mathbf{v}} \times \mathbf{B}_h) \cdot [\zeta_h^\varphi]_v \, ds_v d\mathbf{x}$$

$$\leq C h^{k+1/2} (\|\mathbf{E}_h\|_{0,\infty,\Omega_x} + \|\mathbf{B}_h\|_{0,\infty,\Omega_x}) \|e_h^f\|_{0,\mathcal{T}_h^x \times \mathcal{E}_v} \|\varphi\|_{k+1,\Omega}$$

$$\leq C h^{k+1/2} (\|\mathbf{E}_h\|_{0,\infty,\Omega_x} + \|\mathbf{B}_h\|_{0,\infty,\Omega_x}) h^{-1/2} (\|e_h^f\|_{0,\mathcal{T}_h} + h^{k+1} \|f\|_{k+1,\Omega}) \|\varphi\|_{k+1,\Omega}$$

$$\leq C h^k (\|\mathbf{E}_h\|_{0,\infty,\Omega_x} + \|\mathbf{B}_h\|_{0,\infty,\Omega_x}) (\|e_h^f\|_{0,\mathcal{T}_h} + h^{k+1} \|f\|_{k+1,\Omega}) \|\varphi\|_{k+1,\Omega}$$

$$\leq C h^k (\|e_h^f\|_{0,\mathcal{T}_h} + h^{k+1}) (h^{-d_x/2} \|e_h^{\mathbf{E}}\|_{0,\Omega_x} + h^{-d_x/2} \|e_h^{\mathbf{B}}\|_{0,\Omega_x} + 1) \|\varphi\|_{k+1,\Omega}.$$

In this way we conclude that

$$|a_{h,2}(e_h, \mathbf{E}_h, \mathbf{B}_h; \zeta_h^\varphi)| \leq C h^k (\|e_h^f\|_{0,\mathcal{T}_h} + h^{k+1}) (h^{-d_x/2} \|e_h^{\mathbf{E}}\|_{0,\Omega_x} + h^{-d_x/2} \|e_h^{\mathbf{B}}\|_{0,\Omega_x} + 1) \|\varphi\|_{k+1,\Omega} \tag{2.33}$$

Then by putting together (2.29), (2.31), (2.32), (2.33), and using Theorem 2.1.1, we have

$$
\begin{aligned}
&|a_h(f_h, \mathbf{E}_h, \mathbf{B}_h; \zeta_h^{\varphi}) + b_h(\mathbf{E}_h, \mathbf{B}_h; \zeta_h^{\mathbf{F}}, \zeta_h^{\mathbf{D}})| \\
&\leq Ch^k \left(\|\mathbf{D}\|_{k+1,\Omega_x} + \|\mathbf{F}\|_{k+1,\Omega_x} + \|\varphi\|_{k+1,\Omega}\right) \left(\|e_h^{\mathbf{E}}\|_{0,\Omega_x} + \|e_h^{\mathbf{B}}\|_{0,\Omega_x} + h^{k+1}\right) \\
&+ Ch^k (\|e_h^f\|_{0,\mathcal{T}_h} + h^{k+1})(h^{-d_x/2}\|e_h^{\mathbf{E}}\|_{0,\Omega_x} + h^{-d_x/2}\|e_h^{\mathbf{B}}\|_{0,\Omega_x} + 1)\|\varphi\|_{k+1,\Omega} \\
&\leq Ch^{2k+1/2} \left(\|\mathbf{D}\|_{k+1,\Omega_x} + \|\mathbf{F}\|_{k+1,\Omega_x} + \|\varphi\|_{k+1,\Omega}\right).
\end{aligned}
$$

where we have used $k + 1/2 - d_x/2 > 0$. An application of Cauchy-Schwarz inequality concludes the proof. □

Lastly, we need to estimate the third term, Θ_{C}.

Lemma 2.3.6 (Consistency)**.** *We have*

$$
|\Theta_{\mathrm{C}}| \leq Ch^{2k+1} \left[\int_0^T \|\varphi\|_{k+1,\Omega}^2 \, dt\right]^{1/2} \tag{2.34}
$$

where C depends on the upper bounds of $\|\partial_t f\|_{k+1,\Omega}$,$\|f\|_{k+1,\Omega}$, $|f|_{1,\infty,\Omega}$, $\|\mathbf{E}\|_{1,\infty,\Omega_x}$, $\|\mathbf{B}\|_{1,\infty,\Omega_x}$, $\|\mathbf{E}\|_{k+1,\Omega_x}$, $\|\mathbf{B}\|_{k+1,\Omega_x}$ over the time interval $[0, T]$, and it also depends on the polynomial degree k, mesh parameters σ_0, σ_x and σ_v, and domain parameters L_x and L_v.

Proof. The terms inside the integral of Θ_{C} can be split in $I + II$, where

$$
\begin{aligned}
I &= (f_h, \varphi_t)_\Omega - a_h(f_h, \mathbf{E}_h, \mathbf{B}_h; \varphi) + l_h(\mathbf{J}_h, \mathbf{F}) \\
II &= (\mathbf{E}_h, \mathbf{F}_t)_{\Omega_x} + (\mathbf{B}_h, \mathbf{D}_t)_{\Omega_x} - b_h(\mathbf{E}_h, \mathbf{B}_h; \mathbf{F}, \mathbf{D}) + \mathcal{F}(f, \mathbf{E}, \mathbf{B}; \varphi)
\end{aligned}
$$

since φ is a smooth function, $[\varphi]_x = 0$ and $[\varphi]_v = 0$, in this way, by using (2.23a), and the definition of l_h, we conclude that,

$$
\begin{aligned}
I &= (f_h, -\mathbf{v} \cdot \nabla_{\mathbf{x}} \varphi - (\mathbf{E} + \mathbf{v} \times \mathbf{B}) \cdot \nabla_{\mathbf{v}} \varphi + \mathbf{v} \cdot \mathbf{F})_\Omega - a_h(f_h, \mathbf{E}_h, \mathbf{B}_h; \varphi) + l_h(\mathbf{J}_h; \mathbf{F}) \\
&= -\int_{\mathcal{T}_h} f_h \mathbf{v} \cdot \nabla_{\mathbf{x}} \varphi \, d\mathbf{x} d\mathbf{v} - \int_\Omega f_h (\mathbf{E} + \mathbf{v} \times \mathbf{B}) \cdot \nabla_{\mathbf{v}} \varphi \, d\mathbf{x} d\mathbf{v} - l_h(\mathbf{J}_h; \mathbf{F}) \\
&+ \int_{\mathcal{T}_h} f_h \mathbf{v} \cdot \nabla_{\mathbf{x}} \varphi \, d\mathbf{x} + \int_\Omega f_h (\mathbf{E}_h + \mathbf{v} \times \mathbf{B}_h) \cdot \nabla_{\mathbf{v}} \varphi \, d\mathbf{x} d\mathbf{v} + l_h(\mathbf{J}_h; \mathbf{F}) \\
&= -\int_\Omega f_h (e_h^{\mathbf{E}} + \mathbf{v} \times e_h^{\mathbf{B}}) \cdot \nabla_{\mathbf{v}} \varphi \, d\mathbf{x} \, d\mathbf{v}.
\end{aligned}
$$

On the other hand, by using (2.23b) and (2.23c), since $\mathbf{F}$ and $\mathbf{D}$ are smooth functions $[\mathbf{F}]_\tau = [\mathbf{D}]_\tau = 0$,we have that

$$\begin{aligned}
II &= (\mathbf{E}_h, \nabla_{\mathbf{x}} \times \mathbf{D})_{\mathcal{T}_h^x} - (\mathbf{B}_h, \nabla_{\mathbf{x}} \times \mathbf{F})_{\mathcal{T}_h^x} - b_h(\mathbf{E}_h, \mathbf{B}_h; \mathbf{F}, \mathbf{D}) + \mathcal{F}(f, \mathbf{E}, \mathbf{B}; \varphi) \\
&\quad - \int_\Omega f\mathbf{E}_h \cdot \nabla_v \varphi \, d\mathbf{x}d\mathbf{v} + \int_\Omega f\mathbf{B}_h \cdot (\mathbf{v} \times \nabla_v \varphi) \, d\mathbf{x}d\mathbf{v} \\
&= (\mathbf{E}_h, \nabla_{\mathbf{x}} \times \mathbf{D})_{\mathcal{T}_h^x} - (\mathbf{B}_h, \nabla_{\mathbf{x}} \times \mathbf{F})_{\mathcal{T}_h^x} - (\mathbf{E}_h, \nabla_{\mathbf{x}} \times \mathbf{D})_{\mathcal{T}_h^x} + (\mathbf{B}_h, \nabla_{\mathbf{x}} \times \mathbf{F})_{\mathcal{T}_h^x} \\
&\quad - \int_\Omega f(\mathbf{E}_h + \mathbf{v} \times \mathbf{B}_h) \cdot \nabla_{\mathbf{v}} \varphi \, d\mathbf{x}d\mathbf{v} + \int_\Omega f(\mathbf{E} + \mathbf{v} \times \mathbf{B}) \cdot \nabla_{\mathbf{v}} \varphi \, d\mathbf{x}d\mathbf{v} \\
&= \int_\Omega f(e_h^{\mathbf{E}} + \mathbf{v} \times e_h^{\mathbf{B}}) \cdot \nabla_{\mathbf{v}} \varphi \, d\mathbf{x}d\mathbf{v}.
\end{aligned}$$

We obtain

$$\begin{aligned}
I + II &= \int_\Omega e_h^f (e_h^{\mathbf{E}} + \mathbf{v} \times e_h^{\mathbf{B}}) \cdot \nabla_{\mathbf{v}} \varphi \, d\mathbf{x}d\mathbf{v} \\
&\le C\|e_h^f\|_\Omega (\|e_h^{\mathbf{E}}\|_{\Omega_x} + \|e_h^{\mathbf{B}}\|_{\Omega_x}) \|\nabla_{\mathbf{v}} \varphi\|_{\infty,\Omega} \\
&\le C\|e_h^f\|_\Omega (\|e_h^{\mathbf{E}}\|_{\Omega_x} + \|e_h^{\mathbf{B}}\|_{\Omega_x}) \|\varphi\|_{k+1,\Omega}
\end{aligned}$$

where we used the Sobolev inequality [51], $\|\nabla_{\mathbf{v}} \varphi\|_{\infty,\Omega} \le C\|\varphi\|_{k+1,\Omega}$, which requires $k > (d_x + d_v)/2$. Using Theorem 2.1.1, we conclude the proof. □

It is easy to transform the dual problem (2.23) to an initial value problem (2.21) by changing time $t' = T - t$. Then using Lemma 2.3.3, where $\mathbf{A_1}(\mathbf{x}, \mathbf{v}, t) = -\mathbf{v}$, $\mathbf{A_2}(\mathbf{x}, \mathbf{v}, t) = -(\mathbf{E} + \mathbf{v} \times \mathbf{B})$, $\mathbf{A_3}(\mathbf{x}, \mathbf{v}, t) = \mathbf{v}$, $g = f$ and $l = k + 1$,

$$\|\varphi\|_{k+1,\Omega}^2 + \|\mathbf{F}\|_{k+1,\Omega_x}^2 + \|\mathbf{D}\|_{k+1,\Omega_x}^2 \le C[\|\Phi\|_{k+1,\Omega}^2 + \|\mathfrak{F}\|_{k+1,\Omega_x}^2 + \|\mathfrak{D}\|_{k+1,\Omega_x}^2] \tag{2.35}$$

where C depends on $\|f\|_{L^\infty((0,T);W^{k+2,\infty}(\Omega))}$. Then an application of Theorem 2.1.1 gives us

$$|(e_h^f(T), \Phi)_\Omega + (e_h^{\mathbf{E}}(T), \mathfrak{F})_{\Omega_x} + (e_h^{\mathbf{B}}(T), \mathfrak{D})_{\Omega_x}| \le Ch^{2k+1/2} \sqrt{\|\Phi\|_{k+1,\Omega}^2 + \|\mathfrak{F}\|_{k+1,\Omega_x}^2 + \|\mathfrak{D}\|_{k+1,\Omega_x}^2} \tag{2.36}$$

Therefore the estimate for the zero-divided difference negative-order norm is given by

$$\begin{aligned}
&\|(f - f_h, \mathbf{E} - \mathbf{E}_h, \mathbf{B} - \mathbf{B}_h)\|_{-(k+1),\Omega} \\
= &\sup_{\phi \in C_0^\infty(\Omega), \mathfrak{F}, \mathfrak{D} \in [C^\infty(\Omega_x)]^{d_x}} \frac{(f - f_h, \Phi)_\Omega + (\mathbf{E} - \mathbf{E}_h, \mathfrak{F})_{\Omega_x} + (\mathbf{B} - \mathbf{B}_h, \mathfrak{D})_{\Omega_x}}{\sqrt{\|\Phi\|_{k+1,\Omega}^2 + \|\mathfrak{F}\|_{k+1,\Omega_x}^2 + \|\mathfrak{D}\|_{k+1,\Omega_x}^2}} \le Ch^{2k+1/2}.
\end{aligned}$$ □

2.4 Numerical Experiments

In this section, we validate our theoretical results using several numerical tests. In particular, we want to demonstrate the performance of the post-processing technique for the VA system and the VM system. We heavily use the fact that the VM (VA) system is time reversible to provide quantitative measurements of the errors. In particular, let $f(\mathbf{x}, \mathbf{v}, 0)$, $\mathbf{E}(\mathbf{x}, 0)$, $\mathbf{B}(\mathbf{x}, 0)$ denote the initial conditions and $f(\mathbf{x}, \mathbf{v}, T)$, $\mathbf{E}(\mathbf{x}, T)$, $\mathbf{B}(\mathbf{x}, T)$ be the solution of the VM system at $t = T$. If we choose $f(\mathbf{x}, -\mathbf{v}, T)$, $\mathbf{E}(\mathbf{x}, T)$, $-\mathbf{B}(\mathbf{x}, T)$ as the initial condition at $t = 0$, then evolving the VM system to $t = T$, we will recover $f(\mathbf{x}, -\mathbf{v}, 0)$, $\mathbf{E}(\mathbf{x}, 0)$, $-\mathbf{B}(\mathbf{x}, 0)$.

2.4.1 Vlasov-Ampére examples

We consider two classical benchmark examples.

- Landau damping:

$$f(x, v, 0) = f_M(v)(1 + A\cos(kx)), \quad x \in [0, L],\ v \in [-V_c, V_c], \tag{2.37}$$

  where $A = 0.5$, $k = 0.5$, $L = 4\pi$, $V_c = 6\pi$, and $f_M(v) = \frac{1}{\sqrt{2\pi}} e^{-v^2/2}$.

- Two-stream instability:

$$f(x, v, 0) = f_{TS}(v)(1 + A\cos(kx)), \quad x \in [0, L],\ v \in [-V_c, V_c], \tag{2.38}$$

  where $A = 0.05$, $k = 0.5$, $L = 4\pi$, $V_c = 6\pi$, and $f_{TS}(v) = \frac{1}{\sqrt{2\pi}} v^2 e^{-v^2/2}$.

Notice that in both examples we have taken V_c to be larger than the usual values in the literature in order to completely eliminate the boundary effects and accurately reflect the accuracy enhancement property.

In Tables 2.1, we run the VA system with initial condition from Landau damping to $T = 1$ and then back to $T = 0$ and then we apply the SIAC filter, and compare it with the initial conditions. We use the third order TVD-RK method as the time integrator [52]. To make sure the spatial error dominates, we take $\Delta t = \mathrm{CFL}/(V_c/\Delta x + \mathbf{E}_{\max}/\Delta v)$ for $\mathbb{P}^1$,

$\mathbf{E}_{\max}$ denotes the maximum value of $\mathbf{E}(\cdot, T)$ in Ω_x, for $\mathbb{P}^2$ we take $\Delta t = \mathrm{CFL}/(V_c/(\Delta x)^{5/3} + E_{max}/(\Delta v)^{5/3})$, and $\Delta t = \mathrm{CFL}/(V_c/(\Delta x)^{7/3} + E_{max}/(\Delta v)^{7/3})$ for $\mathbb{P}^3$. For $\mathbb{P}^1$ and $\mathbb{P}^3$ we take the $\mathrm{CFL} = 0.1$, and we take the $\mathrm{CFL} = 0.2$ for $\mathbb{P}^2$. From the table, we observe $(k+1)$-th order of convergence for the DG solution before post-processing for both f and $\mathbf{E}$. We can clearly see that we improve the order of the error to at least $O(h^{2k+1/2})$ after post-processing.

In Figure 2.1 we plot the errors of the numerical solution before and after post-processing for $\mathbb{P}^1$ and using 128×128 elements. We can see that the errors before post-processing are highly oscillatory, and that the post-processing smooths out the error and greatly reduces its magnitude. In Figure 2.2, we plot the errors of the approximations for $\mathbf{E}$ obtained when solving using a 128×128 mesh with $\mathbb{P}^1$ and 32×32 mesh with $\mathbb{P}^3$. We can clearly see that the errors before post-processing are highly oscillatory, and the post-processing gets rid of the oscillations and dramatically reduces the magnitude of the error.

Another point that we want to make is the following: if we look at Table 2.1, for $k = 2$ and a mesh of 64×64, the L^2-errors before and after post-processing are similar in magnitude. However, if we look at Figure 2.3 which plots the absolute value of the error in f in this case, we can clearly see that the L^∞-norm of the error of the filtered solution is much smaller than the unfiltered solution. Therefore, by removing the spurious oscillations, even if the L^2-error is comparable, the L^∞ error is further reduced by the post-processor. This is probably due to the high oscillatory nature of the solution.

In Tables 2.2, we run the VA system with initial condition from two stream instability to $T = 1$ and then back to $T = 0$ and then we apply the SIAC filter, and compare it with the initial conditions. To integrate in time we used Fourth order Runge-Kutta for $\mathbb{P}^1$ and $\mathbb{P}^3$ and third order TVD-RK method for $\mathbb{P}^2$. For $\mathbb{P}^1$ and $\mathbb{P}^2$ we take Δt just as in the Landau damping example, and $\Delta t = \mathrm{CFL}/(V_c/(\Delta x)^{7/4} + E_{max}/(\Delta v)^{7/4})$ for $\mathbb{P}^3$. We use $\mathrm{CFL} = 0.2$ for all cases. We can observe $(k + 1/2)$-order of convergence for the DG solution before post-processing for both f and $\mathbf{E}$. Just as in the Landau damping example we can see the

	Before post-processing				After post-processing			
mesh	error f	order	error $\mathbf{E}$	order	error f^*	order	error $\mathbf{E}^*$	order
$\mathbb{P}^1$								
16×16	1.42E-02	-	1.19E-02	-	2.28E-02	-	1.04E-02	-
32×32	6.22E-03	1.19	3.16E-03	1.91	6.16E-03	1.89	2.84E-03	1.88
64×64	1.59E-03	1.97	5.65E-04	2.48	8.74E-04	2.82	4.36E-04	2.70
128×128	4.08E-04	1.96	1.12E-04	2.33	1.10E-04	2.99	6.31E-05	2.79
256×256	1.03E-04	1.98	2.51E-05	2.16	1.37E-05	3.00	9.01E-06	2.81
512×512	2.60E-05	1.99	6.14E-06	2.03	1.71E-06	3.00	1.71E-06	2.39
$\mathbb{P}^2$								
16×16	7.08E-03	-	1.97E-03	-	2.09E-02		1.88E-03	-
32×32	1.08E-03	2.71	1.13E-04	4.12	2.87E-03	2.87	1.08E-04	4.12
64×64	1.35E-04	3.00	6.62E-06	4.10	1.20E-04	4.58	5.15E-06	4.39
128×128	1.63E-05	3.04	5.59E-07	3.57	2.70E-06	5.47	2.04E-07	4.66
256×256	2.01E-06	3.03	6.57E-08	3.09	5.29E-08	5.67	5.75E-09	5.15
$\mathbb{P}^3$								
16×16	1.73E-03	-	2.19E-04	-	2.16E-02	-	9.71E-05	-
32×32	1.52E-04	3.51	7.18E-06	4.93	2.60E-03	3.05	3.09E-06	4.97
64×64	1.06E-05	3.84	1.30E-07	5.79	5.65E-05	5.52	7.52E-08	5.36
128×128	6.45E-07	4.04	3.42E-09	5.25	3.95E-07	7.16	8.24E-10	6.51

Table 2.1 L^2 errors for the numerical solution and the post-processed solution for Landau Damping.

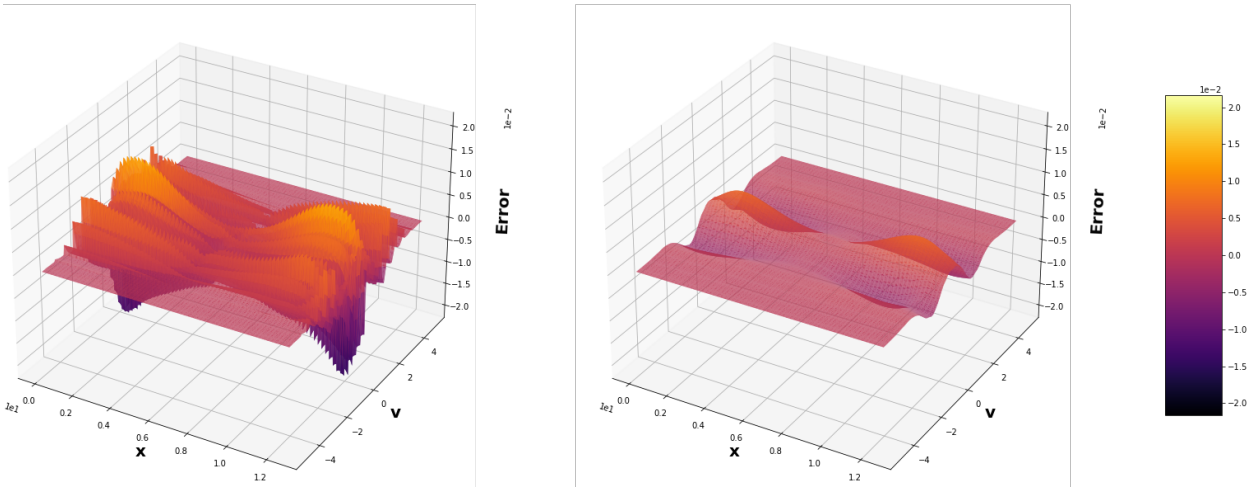
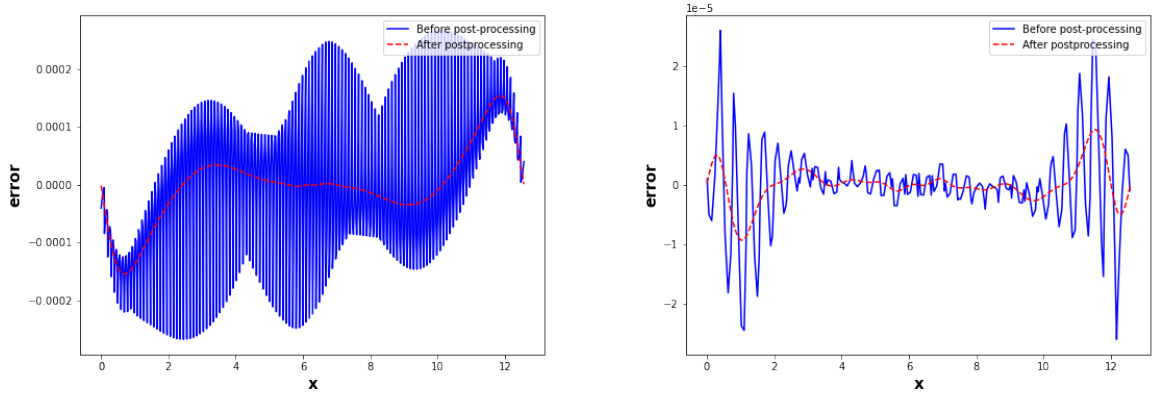


Figure 2.1 Errors for f before (on the left) and after post-processing (on the right) for 128×128 elements and $\mathbb{P}^1$. Landau damping.



(a) 128×128 and $\mathbb{P}^1$ (b) 32×32 and $\mathbb{P}^3$

Figure 2.2 Errors before (solid line) and after post-processing (dashed line) for $\mathbf{E}$ for different mesh sizes and $\mathbb{P}^k$. Landau damping. $T = 2$.

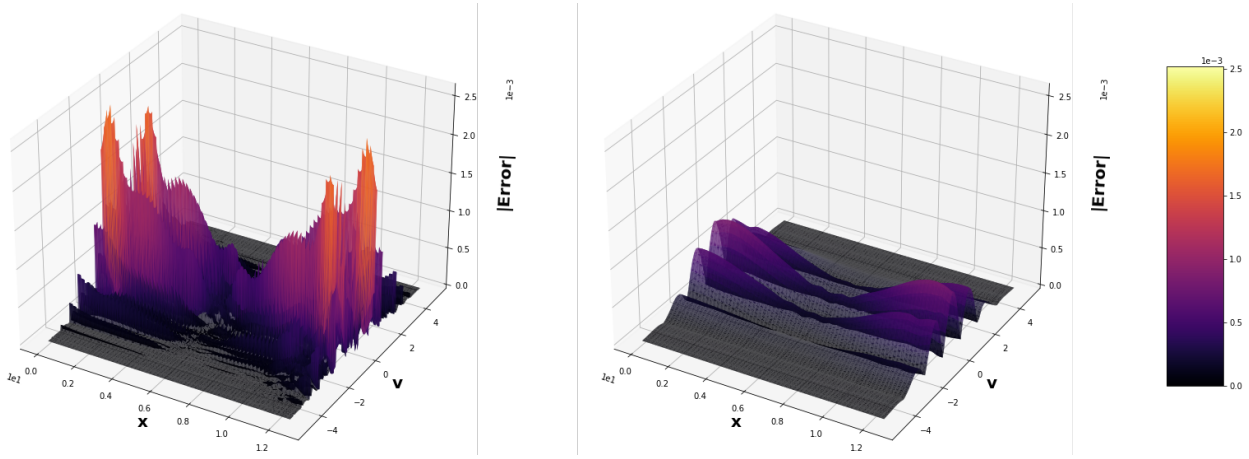


Figure 2.3 Absolute value of errors for f before (on the left) and after post-processing (on the right) for 64×64 elements and $\mathbb{P}^2$. Landau damping.

improvement of the error to $O(h^{2k+1/2})$ after post-processing.

In Figure 2.4 we plot the errors of the numerical solution before and after post-processing for $\mathbb{P}^1$ and using 128×128 elements. We can arrrive to similar conclusions as in the Landau damping case.

Figure 2.5, we plot the errors of the approximations for $\mathbf{E}$ obtained when solving using a 16×16 mesh with $\mathbb{P}^2$ and 32×32 mesh with $\mathbb{P}^3$. We can clearly see how the SIAC filter, gets rid of the spurious oscillations and dramatically reduce the magnitude of the error.

Now we provide plots comparing the solution profile before and after post-processing

	Before post-processing				After post-processing			
mesh	error f	order	error $\mathbf{E}$	order	error f^*	order	error $\mathbf{E}^*$	order
				$\mathbb{P}^1$				
16×16	2.63E-02	-	3.24E-03	-	3.46E-02	-	3.12E-03	-
32×32	5.58E-03	2.23	2.13E-04	3.93	9.83E-03	1.82	1.50E-04	4.38
64×64	2.23E-03	1.32	4.54E-05	2.23	1.19E-03	3.05	2.45E-05	2.62
128×128	6.08E-04	1.88	1.04E-05	2.13	8.91E-05	3.74	3.66E-06	2.74
256×256	1.78E-04	1.77	2.49E-06	2.06	7.06E-06	3.66	5.15E-07	2.83
512×512	4.98E-05	1.84	6.13E-07	2.02	6.80E-07	3.38	6.30E-08	3.03
				$\mathbb{P}^2$				
16×16	7.60E-03	-	1.07E-04		3.54E-02		3.91E-05	-
32×32	2.36E-03	1.69	5.45E-06	4.30	9.81E-03	1.85	4.88E-06	3.00
64×64	2.96E-04	2.99	4.18E-07	3.70	5.48E-04	4.16	3.22E-07	3.92
128×128	4.69E-05	2.66	3.67E-08	3.51	1.28E-05	5.42	1.14E-08	4.81
256×256	7.15E-06	2.71	4.72E-09	2.96	2.23E-07	5.84	5.16E-10	4.47
				$\mathbb{P}^3$				
16×16	5.06E-03	-	2.31E-05	-	3.68E-02	-	1.88E-05	-
32×32	1.09E-04	5.54	1.18E-06	4.29	1.02E-02	1.85	8.97E-08	7.71
64×64	2.69E-05	2.01	1.99E-08	5.89	3.47E-04	4.88	6.56E-09	3.77
128×128	2.01E-06	3.75	2.74E-10	6.18	2.83E-06	6.94	9.68E-11	6.08

Table 2.2 L^2 errors for the numerical solution and the post-processed solution. Two stream instability.

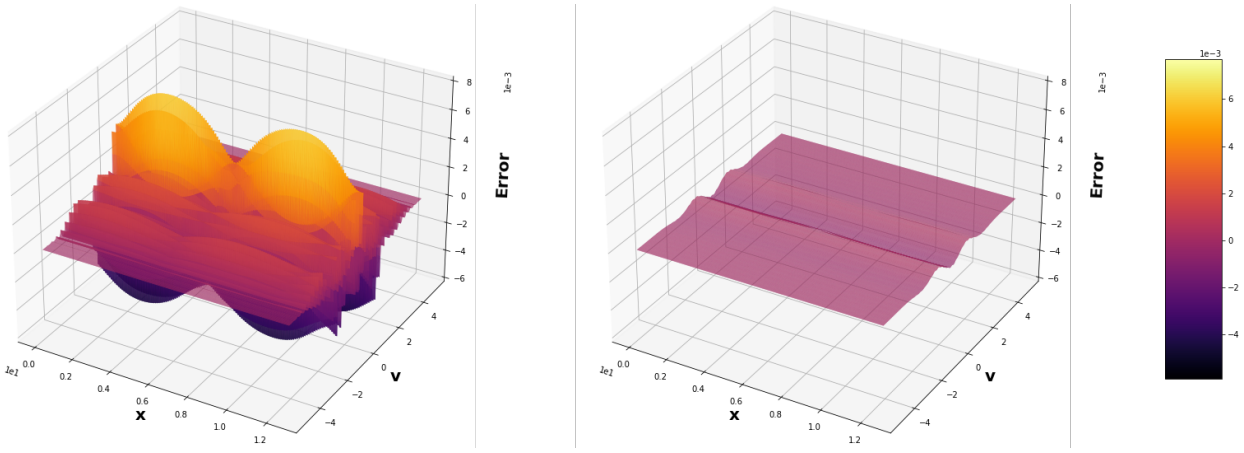
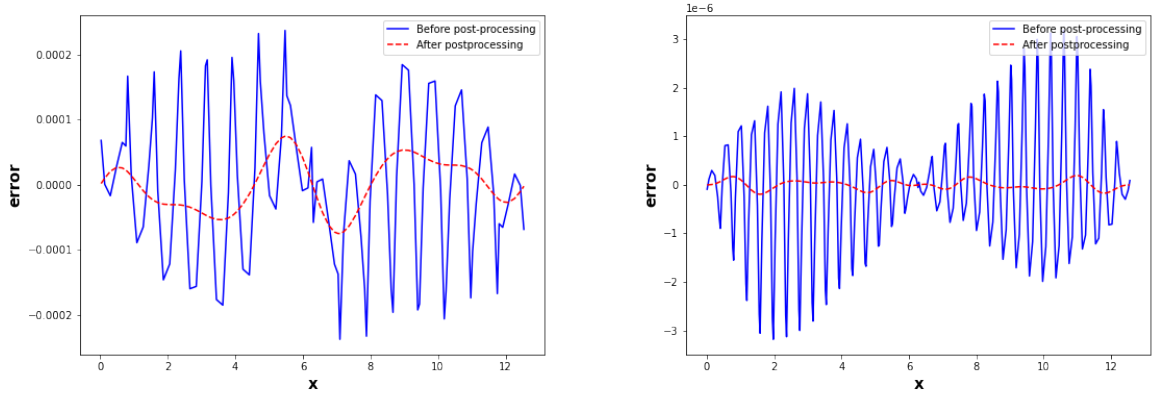


Figure 2.4 Errors for f before (on the left) and after post-processing (on the right) for 128 elements and $\mathbb{P}^1$, using two-stream instability.



(a) 16×16 and $\mathbb{P}^2$ (b) 32×32 and $\mathbb{P}^3$

Figure 2.5 Errors before (solid line) and after post-processing (dashed line) for $\mathbf{E}$ for different mesh sizes and $\mathbb{P}^k$. Two stream instability. $T = 2$.

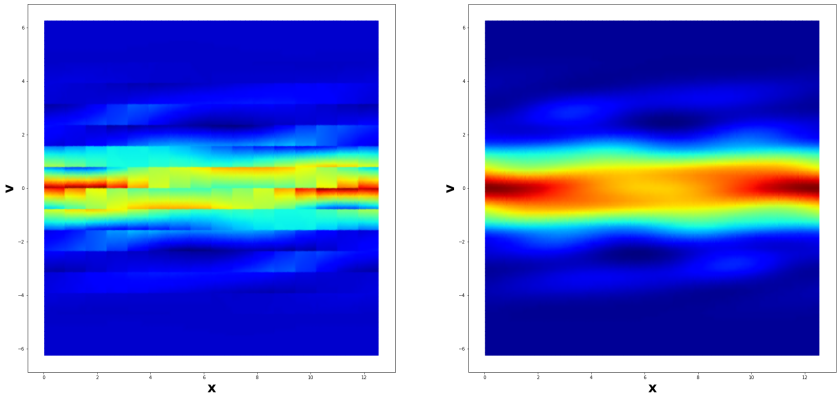
for a longer computational time. To compute those plots, we use a third-order Runge-Kutta method with $\Delta t = \mathrm{CFL}/(V_c/\Delta x + E_{max}/\Delta v))$ and $\mathrm{CFL} = 0.1$. In Figures 2.6 to 2.9, we show a comparison of contour plots of the numerical solution for f before and after post-processing with different mesh size and $k = 1, 2$. There is visible improvement of the resolution of the solution, particularly for $k = 1$.

2.4.2 Vlasov-Maxwell example

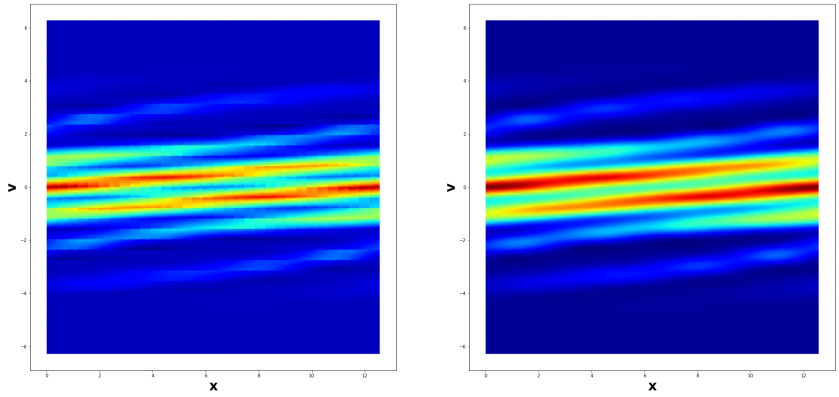
In this part, we will test our post-processor for the VM system. Specifically we will use the streaming Weibel (SW) instability as an example. This is a reduced version of the VM equations with one spatial variable, x_2, and two velocity variables v_1 and v_2. The variables under consideration are the distribution function $f(x_2, v_1, v_2, t)$, a 2D electric field $\mathbf{E} = (E_1(x_2, t), E_2(x_2, t), 0)$ and a 1D magnetic field $\mathbf{B} = (0, 0, B_3(x_2, t))$ and the reduced VM system reads as

$$\partial_t f + v_2 f_{x_2} + (E_1 + v_2 B_3) f_{v_1} + (E_2 - v_1 B_3) f_{v_2} = 0, \tag{2.39a}$$

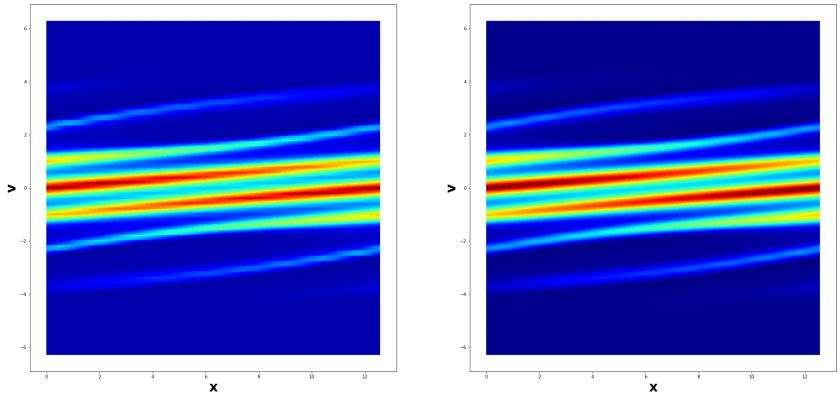
$$\frac{\partial B_3}{\partial t} = \frac{\partial E_1}{\partial x_2}, \quad \frac{\partial E_1}{\partial t} = \frac{\partial B_3}{\partial x_2} - j_1, \quad \frac{\partial E_2}{\partial t} = -j_2, \tag{2.39b}$$



(a) 16×16

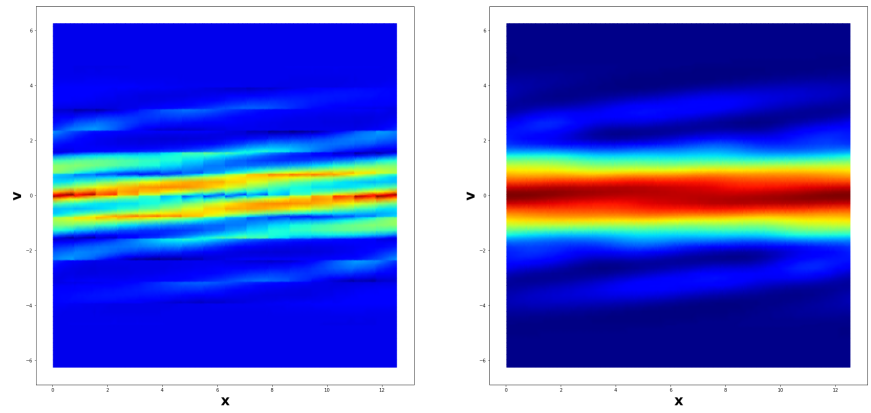


(b) 32×32

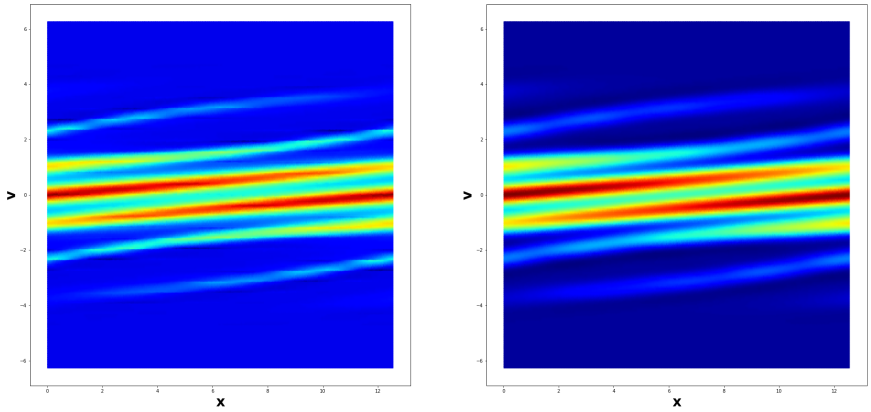


(c) 64×64

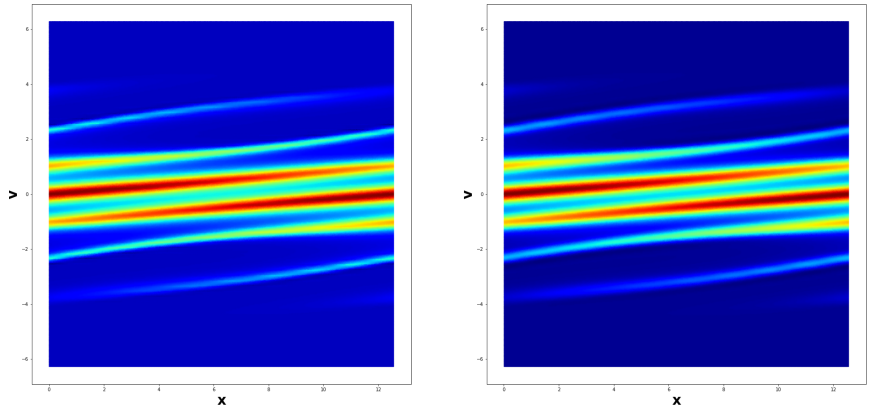
Figure 2.6 Comparison of contour plots before (left) and after post-processing (right) for different mesh-sizes. Landau damping, $k = 1$ and $T = 10$.



(a) 16×16

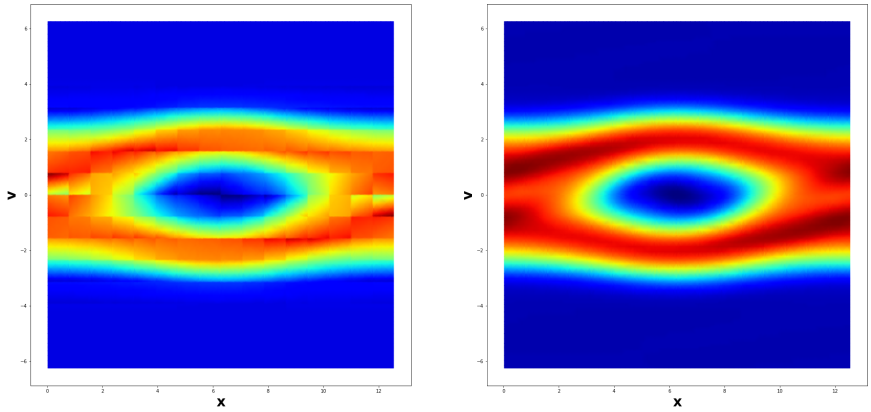


(b) 32×32

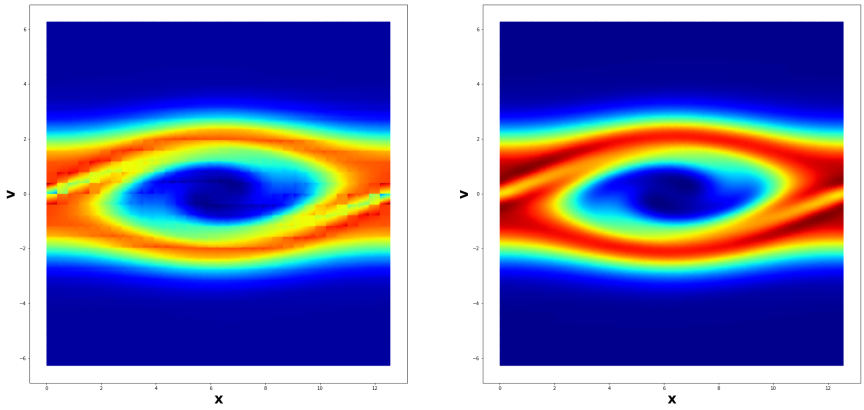


(c) 64×64

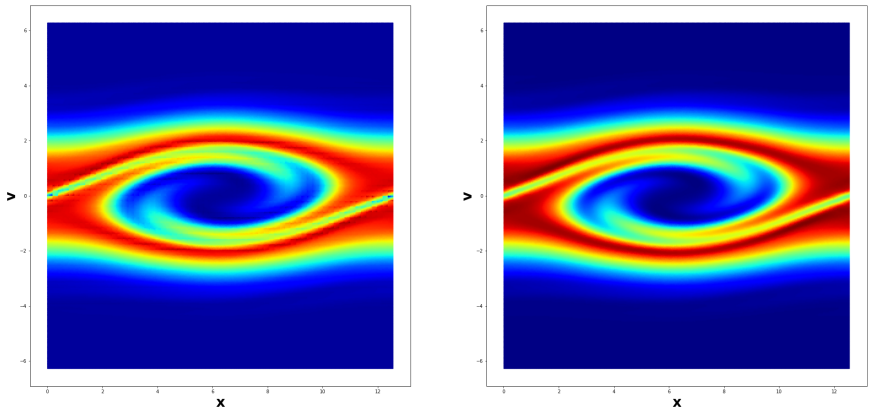
Figure 2.7 Comparison of the definition of the contour plots before (left) and after post-processing (right) for different mesh-sizes. Landau damping, $k = 2$ and $T = 10$.



(a) 16×16

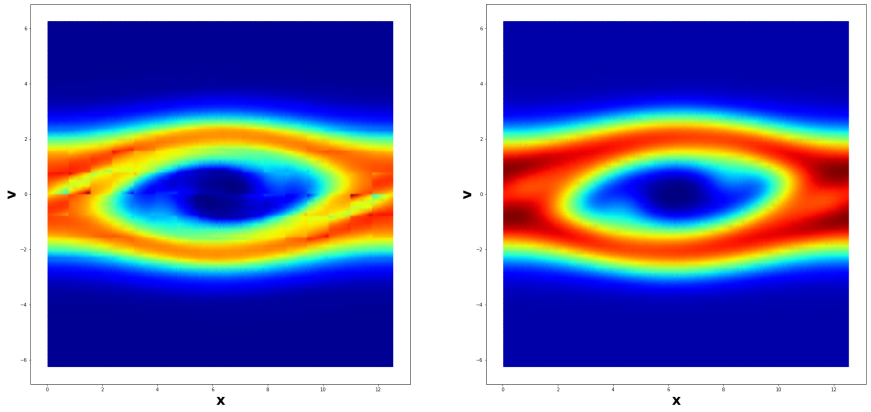


(b) 32×32

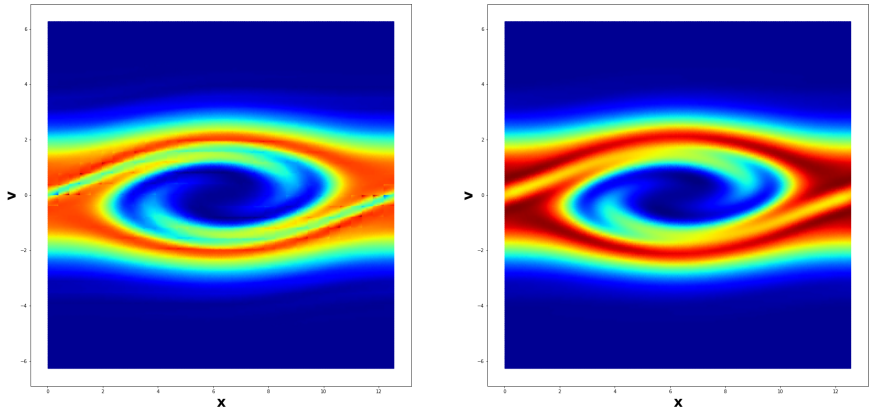


(c) 64×64

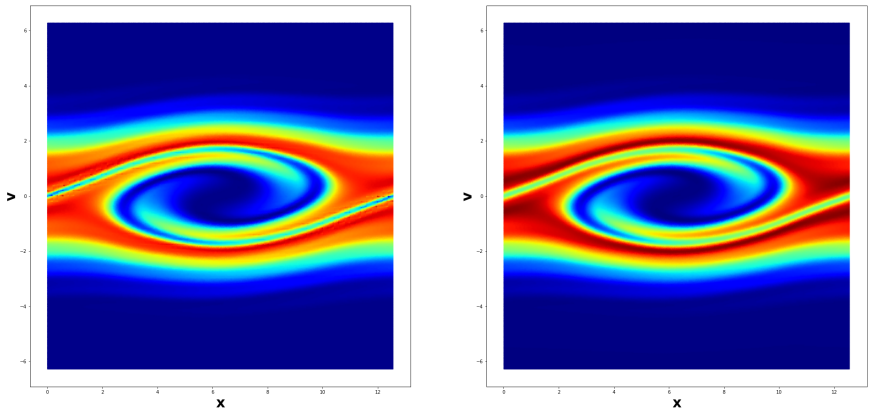
Figure 2.8 Comparison of contour plots before (left) and after post-processing (right) for different mesh-sizes. Two stream instability, $k = 1$ and $T = 20$.



(a) 16×16



(b) 32×32



(c) 64×64

Figure 2.9 Comparison of the definition of the contour plots before (left) and after post-processing (right) for different mesh-sizes. Two stream instability, $k = 2$ and $T = 20$.

where

$$j_1 = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} f(x_2, v_1, v_2, t) v_1 \, dv_1 dv_2, \quad j_2 = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} f(x_2, v_1, v_2, t) v_2 \, dv_1 dv_2. \tag{2.40}$$

The initial conditions are given by

$$f(x_2, v_1, v_2, 0) = \frac{1}{\pi\beta} e^{-v_2^2/\beta} [\delta e^{-(v_1-\omega_{0,1})^2/\beta} + (1-\delta) e^{-(v_1+\omega_{0,2})^2/\beta}], \tag{2.41a}$$

$$E_1(x_2, v_1, v_2, 0) = E_2(x_2, v_1, v_2, 0) = 0, \quad B_3(x_2, v_1, v_2, 0) = b \sin(\kappa_0 x_2), \tag{2.41b}$$

which for $b = 0$ is an equilibrium state composed of counter-streaming beams propagating perpendicular to the direction of inhomogeneity. Following [6, 19], we trigger the instability by taking $\beta = 0.01$, $b = 0.001$ (the amplitude of the initial perturbation of the magnetic field). Here, $\Omega_x = [0, L_y]$, where $L_y = 2\pi/\kappa_0$, and we set $\Omega_v = [-1.8, 1.8]^2$. We consider the following set of parameters,

$$\delta = 0.5, \, \omega_{0,1} = \omega_{0,2} = 0.3, \, \kappa_0 = 0.2.$$

In Table 2.3, we run the VM system with initial condition from SW instability to $T = 5$ and then back to $T = 0$, we then apply the SIAC filter and compare it with the initial conditions. We use a third order TVD-RK method as the time integrator. To make sure the spatial error dominates, we take $\Delta t = O(\Delta x)$ for $\mathbb{P}^1$ and $\Delta t = O(\Delta x^{5/3})$ for $\mathbb{P}^2$, in both cases we used $\mathrm{CFL} = 0.1$. From the table we can observe $(k+1)$-th order of convergence for the DG solution before post-processing for f, E_1, E_2 and B_3. After post-processing we can see overall the order of convergence improves to $O(h^{2k+1/2})$.

In Figure 2.10 we plot a cross-section of the errors of the numerical solution at $x_2 \approx 0.15\pi$ before and after post-processing for $\mathbb{P}^1$ using $80 \times 80 \times 80$ elements. We can see that before post-processing that the errors are highly oscillatory, and after post-processing the error surface is smooth out and the error is much smaller in magnitude. In Figure 2.11 we plot the errors of E_1, E_2 and B_3, we used the same number of elements as in Figure 2.10, We can clearly see similar conclusions.

Before post-processing								
Mesh	Error f	Order	Error $\mathbf{B}_3$	Order	Error $\mathbf{E}_1$	Order	Error $\mathbf{E}_2$	Order
$\mathbb{P}^1$								
$20 \times 20 \times 20$	2.20E-01	-	2.61E-06	-	2.12E-06	-	5.31E-06	-
$40 \times 40 \times 40$	7.17E-02	1.61	6.54E-07	2.00	7.06E-07	1.58	5.46E-07	3.28
$80 \times 80 \times 80$	1.92E-02	1.90	1.63E-07	2.00	1.96E-07	1.85	7.05E-08	2.95
$160 \times 160 \times 160$	4.89E-03	1.98	4.07E-08	2.00	5.13E-08	1.94	6.40E-09	3.46
$\mathbb{P}^2$								
$20 \times 20 \times 20$	1.07E-01	-	2.56E-07	-	2.49E-07	-	1.02E-06	
$40 \times 40 \times 40$	1.64E-02	2.70	3.14E-08	3.03	2.93E-08	3.09	9.72E-08	3.40
$80 \times 80 \times 80$	2.23E-03	2.88	1.63E-09	4.27	1.90E-09	3.95	6.93E-09	3.81
$160 \times 160 \times 160$	2.92E-04	2.93	1.41E-10	3.52	1.72E-10	3.46	2.46E-10	4.81
After post-processing								
Mesh	Error f^*	Order	Error $\mathbf{B}_3^*$	Order	Error $\mathbf{E}_1^*$	Order	Error $\mathbf{E}_2^*$	Order
$\mathbb{P}^1$								
$20 \times 20 \times 20$	2.95E-01	-	3.17E-07	-	1.08E-07	-	5.08E-06	-
$40 \times 40 \times 40$	6.13E-02	2.27	7.16E-08	2.14	1.49E-08	2.87	4.38E-07	3.54
$80 \times 80 \times 80$	5.87E-03	3.38	1.12E-08	2.68	3.11E-09	2.26	6.33E-08	2.79
$160 \times 160 \times 160$	4.19E-04	3.81	2.01E-09	2.48	7.47E-10	2.06	6.22E-09	3.35
$\mathbb{P}^2$								
$20 \times 20 \times 20$	2.89E-01	-	1.24E-08	-	9.06E-09	-	4.41E-07	-
$40 \times 40 \times 40$	4.58E-02	2.66	5.61E-10	4.46	2.97E-10	4.93	2.63E-08	4.07
$80 \times 80 \times 80$	2.03E-03	4.49	2.94E-11	4.25	1.31E-11	4.50	2.57E-09	3.36
$160 \times 160 \times 160$	4.43E-05	5.52	1.65E-12	4.15	5.55E-13	4.56	1.12E-10	4.53

Table 2.3 L^2 errors for the numerical solution (Above) and the post-processed solution (Below). SW instability.

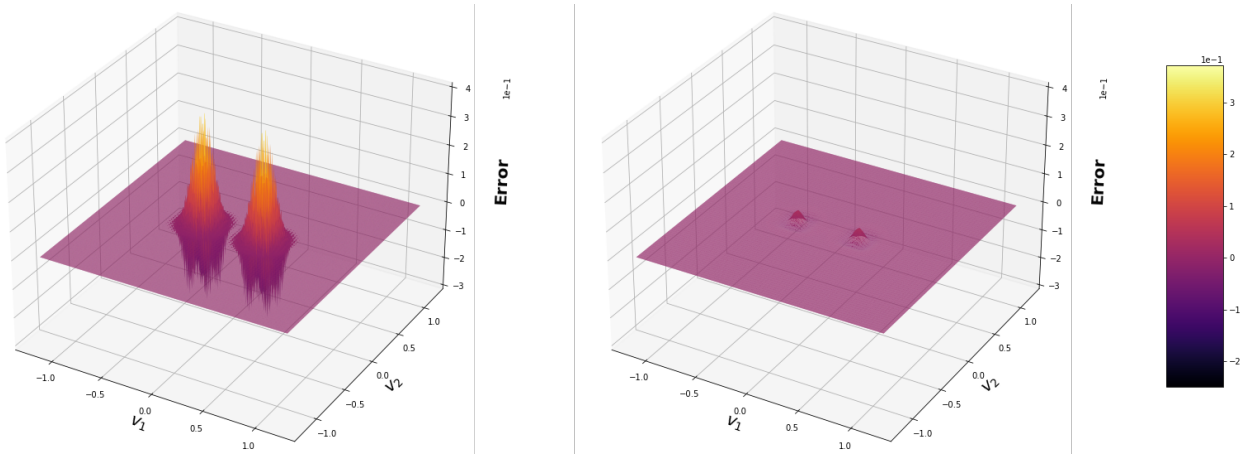
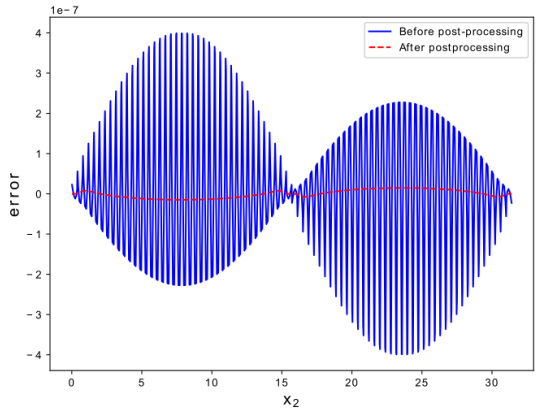
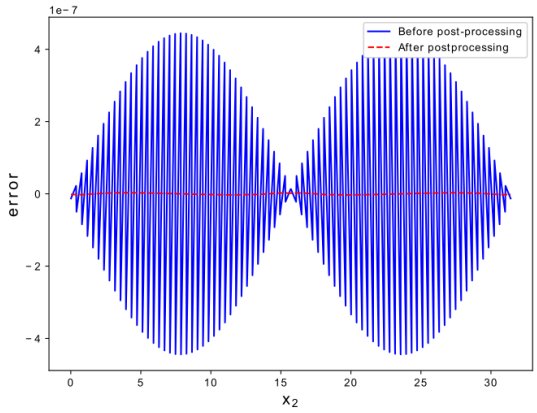


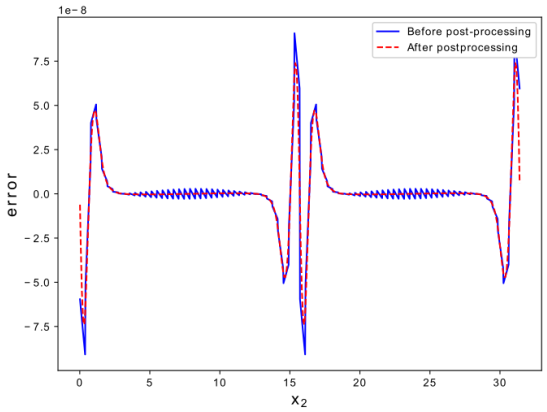
Figure 2.10 Cross-sectional plot of the error for f at $x_2 \approx 0.15\pi$, before (on the left) and after post-processing (on the right) for 80^3 elements and $\mathbb{P}^1$. SW instability.



(a) Error for $\mathbf{B}_3$.



(b) Error for $\mathbf{E}_1$.



(c) Error for $\mathbf{E}_2$.

Figure 2.11 Errors before (solid line) and after post-processing (dashed line) for the different fields using mesh size of $80 \times 80 \times 80$ and $\mathbb{P}^1$. $T = 10$. SW instability.

CHAPTER 3

NUMERICAL ANALYSIS OF A HYBRID METHOD FOR RADIATION TRANSPORT

In this Chapter we will analyze a Hybrid method for Radiation transport introduced in [2]. The remaining of the chapter is organized as follows: In Section 3.1, we introduce the RTE, reduce it to the purely scattering problem, recall the P_N method, and then describe the setup of the hybrid. Having established the setting of of the problem, we then summarize the main results of this chapter . In Section 3.2, we derive error estimates for the P_N equations. In Section 3.3, we analyze the hybrid problem. In Section 3.4, we generalize results back to the original RTE with non-zero absorption. The appendix contains some generic results used for the estimates of this chapter.

3.1 Background

3.1.1 The radiation transport equation

We consider a time-dependent transport equation with periodic boundaries, isotropic scattering, unit-speed particles, and diffusion scaling:

$$\varepsilon\partial_t\Psi^\varepsilon + \Omega\cdot\nabla_x\Psi^\varepsilon + \frac{\sigma_\mathrm{t}}{\varepsilon}\Psi^\varepsilon = \left(\frac{\sigma_\mathrm{t}}{\varepsilon} - \varepsilon\sigma_\mathrm{a}\right)\overline{\Psi^\varepsilon} + \varepsilon Q, \qquad \overline{\Psi^\varepsilon} = \frac{1}{4\pi}\int_{\mathbb{S}^2}\Psi^\varepsilon\, d\Omega, \tag{3.1a}$$

$$\Psi^\varepsilon|_{t=0} = g. \tag{3.1b}$$

Here $\Psi^\varepsilon = \Psi^\varepsilon(x,\Omega,t)$ is a function of position $x \in X = [0,2\pi)^3$, direction of flight $\Omega \in \mathbb{S}^2$, and time $t > 0$. It can be interpreted physically as the density of particles at time t with respect to the measure $d\Omega dx$. Particles interact with a material background characterized by an absorption cross-section $\sigma_\mathrm{a} \geq 0$, total cross-section $\sigma_\mathrm{t} \geq \sigma_\mathrm{a}$ (which accounts for scattering and absorption), and a known source $Q = Q(x,\Omega,t)$. The quantity $\sigma_\mathrm{t} - \varepsilon^2\sigma_\mathrm{a}$ is the scaled scattering cross-section, where the non-dimensional parameter $\epsilon > 0$ characterizes the strength of the scattering as well as the relevant time scale. Indeed, it is well-known [32] that in the limit $\varepsilon \to 0$, $\Psi^\varepsilon \to \Psi^0$ where Ψ^0 is independent of angle and satisfies the

diffusion equation of the form

$$\partial_t \Psi^0 - \nabla_x \cdot \left(\frac{1}{3\sigma_{\mathrm{t}}} \nabla_x \Psi^0 \right) + \sigma_{\mathrm{a}} \Psi^0 = 0, \qquad \Psi^0\big|_{t=0} = \frac{1}{4\pi} \int_{\mathbb{S}^2} g \, d\Omega. \tag{3.2a}$$

For $g \in L^2(X \times \mathbb{S}^2)$, $\sigma_{\mathrm{t}}, \sigma_{\mathrm{a}} \in L^\infty(X)$, $Q \in L^2(X \times \mathbb{S}^2 \times [0,\infty))$ (3.1), is known to have a semi-group solution $\Psi^\varepsilon \in C([0,\infty); L^2(X \times \mathbb{S}^2))$ [53, Theorem XXI.2.3]. If in addition, $\Omega \cdot \nabla_x g \in L^2(X \times \mathbb{S}^2)$, then $\Psi^\varepsilon \in C^1([0,\infty); L^2(X \times \mathbb{S}^2))$. We assume this is the case for remainder of the chapter.

In order to facilitate a clear stability and error analysis, we assume that the cross-sections σ_{a} and σ_{t} are constant in space. This assumption on σ_{a} allows us to convert (3.1) to a purely scattering system for the function $\psi = e^{\sigma_{\mathrm{a}} t} \Psi^\varepsilon$:[1]

$$\varepsilon \partial_t \psi + \Omega \cdot \nabla_x \psi + \frac{\sigma}{\varepsilon} \psi = \frac{\sigma}{\varepsilon} \overline{\psi} + \varepsilon q, \qquad \overline{\psi} = \frac{1}{4\pi} \int_{\mathbb{S}^2} \psi \, d\Omega, \tag{3.3a}$$

$$\psi|_{t=0} = g, \tag{3.3b}$$

where $q = e^{\sigma_{\mathrm{a}} t} Q$ and $\sigma := \sigma_{\mathrm{t}}$. Henceforth, we focus our analysis on (3.3). The results can then be translated back to the case of non-zero absorption by undoing the transformation, which gives exponential decay if $\sigma_{\mathrm{a}} > 0$. This assumption is made for simplicity, but it does introduce a measure of regularity into the solution that is not typical in applications. Indeed, a more reasonable assumption is that the cross-sections are piece-wise smooth and that the boundaries are equipped with inflow data. Hence the analysis here can be viewed as a localized proxy for a more realistic scenario. A more sophisticated analysis to include boundary and interior layers is the subject of future work.

3.1.2 The P_N approximation

Given $N \in \mathbb{N}_{\geq 0}$, the P_N method is a spectral discretization of the transport equation with respect to the angular variable Ω. Let $\{m_{\ell,k}\}_{\ell,k}$ be the real-valued, orthonormal basis of spherical harmonics, where $\ell \geq 0$ denotes the degree and $k \in \{-\ell, \ldots, \ell\}$ denotes the

[1] To reduce notation, we suppress the dependence of ψ on ε.

order. For any $u \in L^2(\mathbb{S}^2)$, the angular moment $u_{\ell,k}$ is given by

$$u_{\ell,k} = \int_{\mathbb{S}^2} m_{\ell,k} u \, d\Omega. \tag{3.4}$$

For convenience, we collect the basis elements of degree ℓ into vectors $\mathbf{m}_\ell = (m_{\ell,-\ell}, \dots, m_{\ell,\ell})^\mathsf{T}$, and we denote by $\mathbf{u}_\ell = (u_{\ell,-\ell}, \dots, u_{\ell,\ell})^\mathsf{T}$ the vector of corresponding moments. Let $\mathbb{P}_N(\mathbb{S}^2) \subset L^2(\mathbb{S}^2)$ to be the span of all spherical harmonics with degree at most N. Then the orthogonal projections $\mathcal{P}_N \colon L^2(\mathbb{S}^2) \to \mathbb{P}_N(\mathbb{S}^2)$ and $\widetilde{\mathcal{P}}_N \colon L^2(\mathbb{S}^2) \to \mathbb{P}_N(\mathbb{S}^2)$ are given by

$$\mathcal{P}_N u = \sum_{\ell=0}^{N} \mathbf{m}_\ell^\mathsf{T} \mathbf{u}_\ell = \sum_{\ell=0}^{N} \sum_{k=-\ell}^{\ell} m_{\ell,k} u_{\ell,k} \quad \text{and} \quad \widetilde{\mathcal{P}}_N u = (\mathcal{I} - \mathcal{P}_N) u = \sum_{\ell=N+1}^{\infty} \sum_{k=-\ell}^{\ell} m_{\ell,k}, \tag{3.5}$$

where $\mathcal{I}$ is the identity operator.

The P_N approximation of (3.3) seeks a function

$$\psi^N(x, \Omega, t) = \sum_{\ell=0}^{N} \mathbf{m}_\ell^\mathsf{T}(\Omega) \boldsymbol{\psi}_\ell^N(x,t) = \sum_{\ell=0}^{N} \sum_{k=-\ell}^{\ell} m_{\ell,k}(\Omega) \boldsymbol{\psi}_{\ell,k}^N(x,t) \tag{3.6}$$

such that

$$\varepsilon \partial_t \psi^N + \mathcal{P}_N(\Omega \cdot \nabla_x \psi^N) + \frac{\sigma}{\varepsilon} \psi^N = \frac{\sigma}{\varepsilon} \overline{\psi^N} + \varepsilon \mathcal{P}_N q, \qquad \overline{\psi^N} = \frac{1}{4\pi} \int_{\mathbb{S}^2} \psi^N \, d\Omega, \tag{3.7a}$$

$$\psi^N|_{t=0} = \mathcal{P}_N g. \tag{3.7b}$$

When expressed in terms of the moments $\boldsymbol{\psi}_\ell^N$, the P_N method yields the following linear, symmetric hyperbolic system:

$$\varepsilon \partial_t \boldsymbol{\psi}_0^N + \sum_{i=1}^{3} a_1^{(i)} \partial_{x_i} \boldsymbol{\psi}_1^N = \varepsilon \mathbf{q}_0, \quad \text{for } \ell = 0, \tag{3.8a}$$

$$\varepsilon \partial_t \boldsymbol{\psi}_\ell^N + \sum_{i=1}^{3} (a_\ell^{(i)})^\mathsf{T} \partial_{x_i} \boldsymbol{\psi}_{\ell-1}^N + \sum_{i=1}^{3} a_{\ell+1}^{(i)} \partial_{x_i} \boldsymbol{\psi}_{\ell+1}^N + \frac{\sigma}{\varepsilon} \boldsymbol{\psi}_\ell^N = \varepsilon \mathbf{q}_\ell, \quad \text{for } 1 \le \ell \le N-1, \tag{3.8b}$$

$$\varepsilon \partial_t \boldsymbol{\psi}_N^N + \sum_{i=1}^{3} (a_N^{(i)})^\mathsf{T} \partial_{x_i} \boldsymbol{\psi}_{N-1}^N + \frac{\sigma}{\varepsilon} \boldsymbol{\psi}_N^N = \varepsilon \mathbf{q}_N, \quad \text{for } \ell = N. \tag{3.8c}$$

Formulas for the elements in the matrices $a_\ell^{(i)} \in \mathbb{R}^{(2\ell-1)\times(2\ell+1)}$ can be found in the appendix of [44]. In the current work, we rely only on the fact the they are bounded in the operator norm, specifically that $\|a_\ell^{(i)}\|_2 \le 4$.

The exact moments $\psi_\ell = \int_{\mathbb{S}^2} \mathbf{m}_\ell \psi d\Omega$ satisfy an infinite system of equations with a structure similar to (3.8):

$$\varepsilon\partial_t\psi_0 + \sum_{i=1}^{3} a_1^{(i)}\partial_{x_i}\psi_1 = \varepsilon\mathbf{q}_0, \qquad \text{for } \ell = 0, \quad (3.9)$$

$$\varepsilon\partial_t\psi_\ell + \sum_{i=1}^{3}(a_\ell^{(i)})^{\mathsf{T}}\partial_{x_i}\psi_{\ell-1} + \sum_{i=1}^{3} a_{\ell+1}^{(i)}\partial_{x_i}\psi_{\ell+1} + \frac{\sigma}{\varepsilon}\psi_\ell = \varepsilon\mathbf{q}_\ell, \qquad \text{for } \ell \geq 1. \quad (3.10)$$

In particular, the P_N equations for ψ^N can be obtained by truncating (3.9) at $\ell = N$ and then neglecting the moment ψ_{N+1}^N that would otherwise appear in (3.8c).

3.1.3 The hybrid method

The hybrid method is based on a separation of ψ into a *collided component* ψ_{c} and an *uncollided component* ψ_{u}. These components satisfy the coupled system

$$\varepsilon\partial_t\psi_{\mathrm{u}} + \Omega\cdot\nabla_x\psi_{\mathrm{u}} + \frac{\sigma}{\varepsilon}\psi_{\mathrm{u}} = \varepsilon q, \quad (3.11a)$$

$$\varepsilon\partial_t\psi_{\mathrm{c}} + \Omega\cdot\nabla_x\psi_{\mathrm{c}} + \frac{\sigma}{\varepsilon}\psi_{\mathrm{c}} = \frac{\sigma}{\varepsilon}\overline{\psi_{\mathrm{c}}} + \frac{\sigma}{\varepsilon}\overline{\psi_{\mathrm{u}}}, \quad (3.11b)$$

where, as before, a bar denotes the angular average of $\mathbb{S}^2$. The idea of the hybrid is to solve (3.11) using a high-resolution angular discretization for ψ_{u} and a low-resolution angular discretization for ψ_{c} over a time step $\Delta t = T/M$, where $M \in \mathbb{N}_{>0}$ and then perform a reconstruction to reinitialize ψ_{u} and ψ_{c} for the next time step. To formalize this procedure, define a set of temporal grid points $0 = t_0 < t_1 < \ldots < t_M = T$, and for $m \in \{1, 2, \ldots, M\}$, let $f(t_m^-) = \lim_{\delta\to 0^+} f(t_m - \delta)$ for any function f of t that is continuous on $[t_{m-1}, t_m)$. Then for $m \in \{1, 2, \ldots, M\}$, $\psi_{\mathrm{u},m}$ and $\psi_{\mathrm{c},m}$ satisfy the following system of equations over the interval $[t_{m-1}, t_m)$

$$\varepsilon\partial_t\psi_{\mathrm{u},m} + \Omega\cdot\nabla_x\psi_{\mathrm{u},m} + \frac{\sigma}{\varepsilon}\psi_{\mathrm{u},m} = \varepsilon q, \quad (3.12a)$$

$$\varepsilon\partial_t\psi_{\mathrm{c},m} + \Omega\cdot\nabla_x\psi_{\mathrm{c},m} + \frac{\sigma}{\varepsilon}\psi_{\mathrm{c},m} = \frac{\sigma}{\varepsilon}(\overline{\psi_{\mathrm{u},m}} + \overline{\psi_{\mathrm{c},m}}), \quad (3.12b)$$

$$\psi_{\mathrm{u},m}|_{t=t_{m-1}} = \begin{cases} g, & m = 1, \\ \psi_{\mathrm{u},m-1}(t_{m-1}^-) + \psi_{\mathrm{c},m-1}(t_{m-1}^-) & m > 1. \end{cases} \quad (3.12c)$$

$$\psi_{\mathrm{c},m}|_{t=t_{m-1}} = 0. \quad (3.12d)$$

The intuition behind this splitting is that (3.12a) can be discretized with a high-resolution angular discretization but solved more efficiently than (3.3) since angular unknowns are no longer coupled. Although (3.12b) features the same type of angular coupling as (3.3), it can be solved with fewer degrees of freedom because the source $\overline{\psi_{\mathrm{u},m}}$ is, in general, more regular than q. However, because $\psi_{\mathrm{u},m}$ decays exponentially whenever $\sigma > 0$, the hybrid is only solved for a time step Δt before the relabeling in (3.12d)- (3.12c) is implemented. The enables the hybrid to capture more high-resolution features than (3.12b) can do alone.

3.1.4 Angular discretization of the hybrid

We focus now on the angular discretization of (3.12). The strategy of the hybrid is to discretize (3.12a) in angle with a high-resolution method and (3.12b) in angle with a low-resolution method. In practice, there are a variety of strategies and combinations available to do so. For the purposes of analysis, we assume that (3.12a) is solved exactly and that (3.12b) is discretized with a P_N method. That is, we seek $\psi^N = \psi^N_{\mathrm{u},m} + \psi^N_{\mathrm{c},m}$ where for each $m \in \{1, 2, \dots M\}$

$$\left(\psi^N_{\mathrm{u},m}, \psi^N_{\mathrm{c},m}\right) \in C([t_{m-1}, t_m); X \times L^2(\mathbb{S}^2)) \times C([t_{m-1}, t_m); X \times \mathbb{P}_N(\mathbb{S}^2)) \tag{3.13}$$

satisfies

$$\varepsilon \partial_t \psi^N_{\mathrm{u},m} + \Omega \cdot \nabla_x \psi^N_{\mathrm{u},m} + \frac{\sigma}{\varepsilon} \psi^N_{\mathrm{u},m} = \varepsilon q, \tag{3.14a}$$

$$\varepsilon \partial_t \psi^N_{\mathrm{c},m} + \mathcal{P}_N\left(\Omega \cdot \nabla_x \psi^N_{\mathrm{c},m}\right) + \frac{\sigma}{\varepsilon} \psi^N_{\mathrm{c},m} = \frac{\sigma}{\varepsilon}(\overline{\psi^N_{\mathrm{u},m}} + \overline{\psi^N_{\mathrm{c},m}}), \tag{3.14b}$$

$$\tag{3.14c}$$

$$\psi^N_{\mathrm{c},m}\big|_{t=t_{m-1}} = 0, \quad \psi^N_{\mathrm{u},m}\big|_{t=t_{m-1}} = \begin{cases} g, & m = 1, \\ \psi^N_{\mathrm{u},m-1}(t^-_{m-1}) + \psi^N_{\mathrm{c},m-1}(t^-_{m-1}) & m > 1. \end{cases} \tag{3.14d}$$

We compare the accuracy of the solution defined in (3.14) with the monolithic P_N method (3.7a), using the same value of N. To simplify the numerical analysis of these two models, we keep the time and space variables continuous.

Since (3.7a) and (3.14b) have the same computational complexity, the goal is to asses the additional benefit of solving (3.14a). Clearly the additional cost of solving (3.14a) involves both memory and run-time; both are fairly easy to quantify. However, assessing the gains in accuracy is not as simple. Thus is important to understand these gains in order to better quantify observed improvements in run-time efficiency provided by the hybrid.

3.1.5 Preview of main results.

Let $s \geq 1$ be the number of angular L^2 derivatives in the solution ψ and let N be an integer such that $N \geq s-1$. Let $e^N = \psi - \psi^N$ be the error in the P_N approximation (3.7a), and let $e^N_M = \psi - (\psi^N_{\mathrm{u},M} + \psi^N_{\mathrm{c},M})$ be the error in the hybrid at the M-th time step. The main results of the chapter are the P_N error estimate in Theorem 3.2.1 and the hybrid error estimate in Theorem 3.3.2. We compare these errors for two different regimes: first, when $\sigma \asymp 1$[2] and $\varepsilon \to 0$ (the diffusion regime) and second, when $\varepsilon \asymp (1)$ and $\sigma \to 0$ (the purely absorbing regime).[3]

When $\varepsilon \ll 1$ and $\sigma \asymp 1$, Theorem 3.2.1 implies that

$$\|e^N\|_{L^2(X\times\mathbb{S}^2)}(T) \lesssim \frac{T}{(N+1)^s}\left[e^{-\sigma T/\varepsilon^2}\sum_{i=0}^{s-1}\frac{T^i}{\varepsilon^{i+1}} + \frac{\varepsilon^{s-1}}{\sigma^s} + O\left(\frac{\varepsilon}{\sigma}\right)\right]. \tag{3.15}$$

Thus with sufficient regularity, the P_N approximation is spectrally accurate and grows linearly in time for large T. The first and third term in brackets depend on whether or not q and g are isotropic (independent of angle). The first term is due to initial layers when g is non-isotropic, and the third terms arises when q is non-isotropic. When g and q are isotropic, the P_N error reduces to (see Corollary 3.2.2)

$$\|e^N\|_{L^2(X\times\mathbb{S}^2)}(T) \lesssim \frac{\varepsilon^{s-1}T}{\sigma^s(N+1)^s}. \tag{3.16}$$

For the hybrid method the initial condition and source are always isotropic. Thus

[2]Recall that $a \asymp b$ if and only if $a = O(b)$ and $b = O(a)$

[3]Technically, this is the streaming regime for Ψ, since there is no absorption.

Theorem 3.3.2 gives the following compact bound

$$\|e_M^N\|_{L^2(X\times\mathbb{S}^2)}(T) \lesssim \frac{\varepsilon^{s-1}T}{\sigma^s(N+1)^s}. \tag{3.17}$$

Thus the hybrid method comes equipped with a better error estimate than the P_N method, and the estimates agree when the data is isotropic. This suggests that the hybrid method is at least as accurate as the P_N approximation when $\varepsilon \ll 1$ and $\sigma \asymp 1$. In addition, the hybrid estimate is independent of the time step Δt used for re-initialization in this regime.

For both the hybrid method and P_N approximation, the errors converge to zero as $\varepsilon \to 0$ whenever $s > 1$ (modulo the initial layer in (3.15)). This fact is consistent with the fact that the P_N method recovers the diffusion limit (3.2a) whenever $N \geq 1$ [54].

For the second regime of interest, $\sigma \ll 1$ and $\varepsilon \asymp 1$, Theorem 3.2.1 gives the P_N error estimate

$$\|e^N\|_{L^2(X\times\mathbb{S}^2)}(T) \lesssim \frac{\mathfrak{p}_s(T/\varepsilon)}{(N+1)^s}, \tag{3.18}$$

where $\mathfrak{p}_s(\omega) = \sum_{i=0}^{s+1} c_i(T)\omega^i$ is a polynomial of degree $s+1$ with non-negative coefficients $c_i(T) = a_iT + b_i$, $a_i,\ b_i \geq 0$. Here the hybrid estimate provides a significant improvement:

$$\|e_M^N\|_{L^2(X\times\mathbb{S}^2)}(T) \lesssim \frac{\Delta t^s T}{\varepsilon^{s+1}(N+1)^s}\min(1, \frac{\Delta t\sigma}{\varepsilon^2}) \tag{3.19}$$

In particular, $\|e_M^N\|_{L^2(X\times\mathbb{S}^2)}(T) = 0$, when $\sigma = 0$. This result is expected since in that case the uncollided solution and the transport solution agree. As expected, the error is monotonic in Δt; however, small time steps require more evaluations of the uncollided equation and more reintializions. In practice, this additional cost must be taken into account.

3.2 P_N Analysis

3.2.1 Spherical harmonics preliminaries

A natural space to analyze the transport equation and the P_N approximation is the Sobolev space $H_\circ^s(\mathbb{S}^2)$. To describe this space, we recall some elementary facts about spherical harmonics which can be found, for example, in [55, 56]. For $u, v \in L^2(\mathbb{S}^2)$, let

$$(u,v)_{L^2(\mathbb{S}^2)} = \sum_{\ell=0}^{\infty} \mathbf{u}_\ell^\top \mathbf{v}_\ell \qquad \text{and} \qquad \|u\|_{L^2(\mathbb{S}^2)}^2 = \sum_{\ell=0}^{\infty} \|\mathbf{u}_\ell\|^2, \tag{3.20}$$

where

$$\mathbf{u}_\ell^\top \mathbf{v}_\ell = \sum_{k=-\ell}^{\ell} u_{\ell,k} v_{\ell,k} \qquad \text{and} \qquad \|\mathbf{u}_\ell\|^2 = \mathbf{u}_\ell^\top \mathbf{u}_\ell = \sum_{k=-\ell}^{\ell} |u_{\ell,k}|^2. \tag{3.21}$$

A standard way to define Sobolev spaces on the sphere is via the Laplace-Beltrami operator $\Delta_\circ$, which is the spherical component of the Laplacian and for which the spherical harmonics are eigenfunctions: $-\Delta_\circ Y_{\ell,j} = \ell(\ell+1) Y_{\ell,j}$. For even integers s, the usual norm is

$$\|u\|_{H^s_\circ(\mathbb{S}^2)} := \left\| \left(\frac{1}{4} + \Delta_\circ \right)^{s/2} u \right\| = \left(\sum_{\ell=0}^{\infty} b_\ell \|\mathbf{u}_\ell\|^2 \right)^{1/2}, \qquad b_\ell = \left(\frac{1}{2} + \ell \right)^{2s}. \tag{3.22}$$

The definition of this inner product extends naturally to all $s \in \mathbb{R}$, and the space $H^s_\circ(\mathbb{S}^2)$ is then the completion of smooth functions under the $H^s_\circ(\mathbb{S}^2)$ norm [55].

Rather than working directly with the $H^s_\circ(\mathbb{S}^2)$ norm in (3.22), it is convenient in the analysis below to use an equivalent norm. For $s \geq 0$, define the $H^s(\mathbb{S}^2)$ semi-norm and norm by

$$|u|_{H^s(\mathbb{S}^2)} := \left(\sum_{\ell=s}^{\infty} b_\ell \|\mathbf{u}_\ell\|^2 \right)^{1/2} \quad \text{and} \quad \|u\|_{H^s(\mathbb{S}^2)} = \left(s\|u\|^2_{L^2(\mathbb{S}^2)} + |u|^2_{H^s(\mathbb{S}^2)} \right)^{1/2}, \tag{3.23}$$

respectively, where the sum in the semi-norm definition in (3.23) begins at s for technical arguments that are used in the proof of Lemma (3.2.5) below. When $s = 0$, the norms coincide: $\|u\|_{H^0(\mathbb{S}^2)} = \|u\|_{H^0_\circ(\mathbb{S}^2)} = \|u\|_{L^2(\mathbb{S}^2)}$. More generally, the following equivalence holds.

Lemma 3.2.1 (Norm equivalence)**.** *For any $s \geq 0$,*

$$c_1(s)\|u\|_{H^s(\mathbb{S}^2)} \leq \|u\|_{H^s_\circ(\mathbb{S}^2)} \leq c_2(s)\|u\|_{H^s(\mathbb{S}^2)}. \tag{3.24}$$

where

$$c_1(s) = \begin{cases} 1 & \text{if} \quad s = 0, \\ \frac{1}{\sqrt{3s}} & \text{if} \quad s \geq 1 \end{cases} \quad \text{and} \quad c_2(s) = \begin{cases} 1 & \text{if} \quad s = 0, \\ \sqrt{\frac{5}{s}} \left(s - \frac{1}{2} \right)^s & \text{if} \quad s \geq 1 \end{cases}. \tag{3.25}$$

Proof. When $s = 0$, the norms are equal, so (3.24) holds trivially. Thus assume that $s \geq 1$. The first inequality in (3.24) follows from the fact that

$$\|u\|^2_{H^s(\mathbb{S}^2)} = s\sum_{\ell=0}^{\infty}\|\mathbf{u}_\ell\|^2 + \sum_{\ell=s}^{\infty}\left(\frac{1}{2}+\ell\right)^{2s}\|\mathbf{u}_\ell\|^2 \leq 3s\sum_{\ell=0}^{\infty}\left(\frac{1}{2}+\ell\right)^{2s}\|\mathbf{u}_\ell\|^2 = \frac{1}{[c_1(s)]^2}\|u\|^2_{H^s_\diamond(\mathbb{S}^2)}. \tag{3.26}$$

To prove the second inequality in (3.24), we use the elementary inequality

$$\frac{s}{4} \leq \left(s-\frac{1}{2}\right)^{2s}, \quad s \geq 1, \tag{3.27}$$

to conclude that

$$\begin{aligned}\|u\|^2_{H^s_\diamond(\mathbb{S}^2)} &= \sum_{\ell=0}^{s-1}\left(\frac{1}{2}+\ell\right)^{2s}\|\mathbf{u}_\ell\|^2 + |u|^2_{H^s(\mathbb{S}^2)} \leq \frac{1}{s}\left(s-\frac{1}{2}\right)^{2s} s\|u\|^2_{L^2(\mathbb{S}^2)} + |u|^2_{H^s(\mathbb{S}^2)} \\ &\leq \frac{1}{s}\left(s-\frac{1}{2}\right)^{2s} s\|u\|^2_{L^2(\mathbb{S}^2)} + \frac{4}{s}\left(s-\frac{1}{2}\right)^{2s}|u|^2_{H^s(\mathbb{S}^2)} \\ &\leq \frac{5}{s}\left(s-\frac{1}{2}\right)^{2s}\left(s\|u\|^2_{L^2(\mathbb{S}^2)} + |u|^2_{H^s(\mathbb{S}^2)}\right) = [c_2(s)]^2\|u\|^2_{H^s(\mathbb{S}^2)}.\end{aligned} \tag{3.28}$$

□

Lemma 3.2.2 (Approximation property)**.** *For $s \geq 0$ and $N \geq \max\{0, s-1\}$,*

$$\|(\mathcal{I}-\mathcal{P}_N)u\|_{L^2(\mathbb{S}^2)} \leq \frac{1}{(N+1)^s}|(\mathcal{I}-\mathcal{P}_N)u|_{H^s(\mathbb{S}^2)} \leq \frac{1}{(N+1)^s}|u|_{H^s(\mathbb{S}^2)}. \tag{3.29}$$

Proof. From the definition of the projection $\mathcal{P}_N$ in (3.5),

$$\underbrace{\sum_{\ell=N+1}^{\infty}\|\mathbf{u}_\ell\|^2}_{=\|(\mathcal{I}-\mathcal{P}_N)u\|^2_{L^2(\mathbb{S}^2)}} \leq \frac{1}{(N+1)^{2s}}\underbrace{\sum_{\ell=N+1}^{\infty}\left(\frac{1}{2}+\ell\right)^{2s}\|\mathbf{u}_\ell\|^2}_{=|(\mathcal{I}-\mathcal{P}_N)u|^2_{H^s(\mathbb{S}^2)}} \leq \frac{1}{(N+1)^{2s}}\underbrace{\sum_{\ell=s}^{\infty}\left(\frac{1}{2}+\ell\right)^{2s}\|\mathbf{u}_\ell\|^2}_{=|u|^2_{H^s(\mathbb{S}^2)}} \tag{3.30}$$

Taking square roots of each term above yields the desired result. □

For vector-valued functions of space, we define the usual $L^2(X)$ inner-product and norm by

$$(\mathbf{v},\mathbf{w})_{L^2(X)} = \int_X \mathbf{v}(x)^\top\mathbf{w}(x)\,dx \quad \text{and} \quad \|\mathbf{v}\|^2_{L^2(X)} = (\mathbf{v},\mathbf{v})_{L^2(X)} = \int_X \|\mathbf{v}(x)\|^2\,dx. \tag{3.31}$$

The space $H^{0,s}(X\times\mathbb{S}^2) = L^2(X; H^s(\mathbb{S}^2))$ is the space of measurable functions $u\colon X\times\mathbb{S}^2 \to \mathbb{R}$ with the semi-inner product

$$(u,v)_{H^{0,s}(X\times\mathbb{S}^2)} = \int_X \sum_{\ell=s}^{\infty} \left(\frac{1}{2}+\ell\right)^{2s} \mathbf{u}_\ell(x)^\top \mathbf{v}_\ell(x)\, dx = \sum_{\ell=s}^{\infty} \left(\frac{1}{2}+\ell\right)^{2s} (\mathbf{u}_\ell, \mathbf{v}_\ell)_{L^2(X)} \tag{3.32}$$

such that the semi-norm

$$|u|^2_{H^{0,s}(X\times\mathbb{S}^2)} = \int_X \sum_{\ell=s}^{\infty} \left(\frac{1}{2}+\ell\right)^{2s} |\mathbf{u}_\ell(x)|^2 dx = \sum_{\ell=s}^{\infty} \left(\frac{1}{2}+\ell\right)^{2s} \|\mathbf{u}_\ell\|^2_{L^2(X)} \tag{3.33}$$

is bounded.

For $r \in \mathbb{N}_{\geq 0}$ and $s \geq 0$ we define $H^{r,s}(X\times\mathbb{S}^2)$ to be the space of functions $u\colon X\times\mathbb{S}^2 \to \mathbb{R}$ such that the semi-norms.

$$|u|_{H^{\varsigma,s}(X\times\mathbb{S}^2)} := \sum_{i_1,i_2,\ldots,i_\varsigma=1}^{3} |\partial_{x_{i_1}x_{i_2}\ldots x_{i_\varsigma}} u|_{H^{0,s}(X\times\mathbb{S}^2)}. \tag{3.34}$$

are bounded for all positive integers $\varsigma \leq r$. These semi-norms above are equivalent to the standard semi-norms, but are more convenient in the context of the RTE since

$$|\Omega\cdot\nabla_x u|_{H^{r,s}(X\times\mathbb{S}^2)} \leq \sum_{i=1}^{3} |\partial_{x_i} u|_{H^{r,s}(X\times\mathbb{S}^2)} = |u|_{H^{r+1,s}(X\times\mathbb{S}^2)}, \tag{3.35}$$

which will be used in the analysis below. Henceforth, the domain of integration for H^s and $H^{r,s}$ will be left off when there is no ambiguity, i.e.,

$$|u|_{H^s} := |u|_{H^s(\mathbb{S}^2)} \quad \text{and} \quad |u|_{H^{r,s}} := |u|_{H^{r,s}(X\times\mathbb{S}^2)} \tag{3.36}$$

Finally, for $p \geq 1$ and measurable functions $u\colon X\times\mathbb{S}^2\times[\alpha,\beta]\to\mathbb{R}$, we denote the space-time semi-norms by

$$|u|_{L^p([\alpha,\beta];H^{r,s})} = \left(\int_\alpha^\beta |u|^p_{H^{r,s}}\, d\tau\right)^{1/p} \quad \text{and} \quad |u|_{L^\infty([\alpha,\beta]);H^{r,s})} = \operatorname*{ess\,sup}_{t\in[\alpha,\beta]} |u|_{H^{r,s}}. \tag{3.37}$$

3.2.2 Stability of the $\mathbf{P}_N$ system

In this section, we derive estimates on high-order semi-norms that arise in the subsequent error analysis. The analysis requires iterated inequalities of Grönwall type (see

Lemma .0.1 in the Appendix), so we define the following notation to simplify integrals that arise. Let $0 \leq \alpha \leq t \leq \beta < \infty$, and for any function $f \in L^1([\alpha, \beta])$, define the bounded linear operator $\mathcal{A}_\alpha \colon L^1([\alpha, \beta]) \to C^0([\alpha, \beta])$ and its powers $\mathcal{A}_\alpha^k, k \in \mathbb{N}^{\geq 0}$, by

$$\mathcal{A}_\alpha[f](t) := \frac{1}{\varepsilon} \int_\alpha^t e^{-\sigma(t-\tau)/\varepsilon^2} f(\tau) d\tau \quad \text{and} \quad \mathcal{A}_\alpha^k[f](t) = \begin{cases} \mathcal{I}, & k = 0, \\ \mathcal{A}_\alpha[\mathcal{A}_\alpha^{k-1}[f]](t), & k \geq 1. \end{cases} \tag{3.38}$$

In addition, let $\mathbb{1} \in L^1([\alpha, \beta])$ be the function that is identically one, and let

$$F_\alpha(t) = e^{-\sigma(t-\alpha)/\varepsilon^2}, \quad t \in [\alpha, \beta]. \tag{3.39}$$

It is clear from the definition in (3.38) $\mathcal{A}_\alpha$ is monotonic; that is, if $0 \leq f(t) \leq g(t)$ for a.e. $t \in [\alpha, \beta]$, then $\mathcal{A}_\alpha[f](t) \leq \mathcal{A}_\alpha[g](t)$ for all $t \in [\alpha, \beta]$.

Lemma 3.2.3. *Let F_α be given as in* (3.39)*. Then for all $t \geq \alpha$ and every $k \in \mathbb{N}_{\geq 0}$,*

$$\mathcal{A}_\alpha^k[\mathbb{1}](t) \leq \min\left(\frac{\varepsilon^k}{\sigma^k}, \frac{1}{k!} \left(\frac{t-\alpha}{\varepsilon} \right)^k \right) \quad \textit{and} \quad \mathcal{A}_\alpha^k[F_\alpha](t) = \frac{(t-\alpha)^k e^{-\sigma(t-\alpha)/\varepsilon^2}}{k! \varepsilon^k} \tag{3.40}$$

Proof. We first prove the bound in (3.40). Since $e^{-\sigma(t-\tau)/\varepsilon^2} \leq 1$, a direct calculation gives

$$\mathcal{A}_\alpha^k[\mathbb{1}](t) \leq \frac{1}{\varepsilon^k} \int_\alpha^t \int_\alpha^{\tau_{k-1}} \cdots \int_\alpha^{\tau_1} \mathbb{1} d\tau_0 \cdots d\tau_{k-1} = \frac{1}{k!} \left(\frac{t-\alpha}{\varepsilon} \right)^k. \tag{3.41}$$

On the other hand, it follows directly from the definition of $\mathcal{A}_\alpha$ that

$$\mathcal{A}_\alpha[\mathbb{1}](t) = \frac{\varepsilon}{\sigma} \left(1 - F_\alpha(t) \right) \leq \frac{\varepsilon}{\sigma} \quad \Longrightarrow \quad \mathcal{A}_\alpha^k[\mathbb{1}](t) \leq \frac{\varepsilon^k}{\sigma^k} \tag{3.42}$$

Together (3.41) and (3.42) yield (3.40).

We prove the second statement in (3.40) by induction on k. When $k = 0$, the statement follows trivially,

$$\mathcal{A}_\alpha^0[F_\alpha](t) = F_\alpha(t) = \frac{(t-\alpha)^0 e^{-\sigma(t-\alpha)/\varepsilon^2}}{0! e^0}, \tag{3.43}$$

Now let us assume the statement is true for arbitrary k, and

$$\mathcal{A}_\alpha^{k+1}[F_\alpha](t) = \mathcal{A}_\alpha \left[\mathcal{A}_\alpha^k[F_\alpha] \right](t) = \frac{1}{\varepsilon} \int_\alpha^t e^{-\sigma(t-\tau)/\varepsilon^2} \frac{(\tau-\alpha)^k e^{-\sigma(\tau-\alpha)/\varepsilon^2}}{k! \varepsilon^k} d\tau, \tag{3.44}$$

direct integration leads to,

$$\mathcal{A}_\alpha^{k+1}[F_\alpha](t) = \frac{(t-\alpha)^{k+1} e^{-\sigma(t-\alpha)/\varepsilon^2}}{(k+1)! \varepsilon^{k+1}} \tag{3.45}$$

□

3.2.2.1 Estimates for the P_N system

We derive evolution equations for the $H^{r,s}$ semi-norms, defined in (3.34). For $r = 0$, we test each equation in (3.8) by $b_\ell \psi_\ell^N$, integrate by parts, and then sum over $\ell \in \{s, s+1, ..., N\}$. This gives

$$\frac{\varepsilon}{2}\partial_t |\psi^N|_{H^{0,s}}^2 + \frac{\sigma}{\varepsilon}|\psi^N|_{H^{0,s}}^2 = \varepsilon(\mathcal{P}_N q, \psi^N)_{H^{0,s}} - \sum_{i=1}^{3}\sum_{\ell=s}^{N}\left(\left(\ell+\frac{1}{2}\right)^{2s} - \gamma_{s,\ell}\left(\ell-\frac{1}{2}\right)^{2s}\right)\left(\psi_\ell^N, \left(a_\ell^{(i)}\right)^T \partial_{x_i}\psi_{\ell-1}^N\right)_{L^2(X)}, \tag{3.46}$$

where $\gamma_{s,\ell} = (1 - \delta_{s,\ell})$ is used to handle the first non-zero term in the sum over ℓ. For the special case $s = 0$, (3.46) recovers the usual L^2 energy equation:

$$\frac{\varepsilon}{2}\frac{d}{dt}\|\psi^N\|_{L^2(X\times\mathbb{S}^2)}^2 + \frac{\sigma}{\varepsilon}\|\psi^N - \overline{\psi^N}\|_{L^2(X\times\mathbb{S}^2)}^2 = \varepsilon(\mathcal{P}_N q, \psi^N)_{L^2(X\times\mathbb{S}^2)}^2. \tag{3.47}$$

To find a closed estimate with respect to the $H^{0,s}$ semi-norms, we focus on the summation in (3.46).

Lemma 3.2.4. *Let $s \geq 1$ and $\ell \geq s$. Then*

$$\left(\ell+\frac{1}{2}\right)^{2s} - \gamma_{s,\ell}\left(\ell-\frac{1}{2}\right)^{2s} \leq 2es\left(\ell+\frac{1}{2}\right)^{s}\left(\ell-\frac{1}{2}\right)^{s-1}. \tag{3.48}$$

Proof. We first establish an elementary inequality. Since $\ell \geq s$,

$$\ell+\frac{1}{2} = \left(\ell-\frac{1}{2}\right)\left(1+\frac{1}{\ell-\frac{1}{2}}\right) \leq \left(\ell-\frac{1}{2}\right)\left(1+\frac{1}{s-\frac{1}{2}}\right) \leq \left(\ell-\frac{1}{2}\right)e^{1/(s-1/2)} \tag{3.49}$$

Therefore

$$\left(\ell+\frac{1}{2}\right)^{s-1} \leq e^{\frac{s-1}{s-1/2}}\left(\ell-\frac{1}{2}\right)^{s-1} \leq e\left(\ell-\frac{1}{2}\right)^{s-1}. \tag{3.50}$$

We use (3.50) to show (3.48). There are two cases:

Case 1 ($\ell = s$): In this case, $\gamma_{s,\ell} = 0$. Since $2s > s + 1/2$ and by (3.50),

$$\left(\ell+\frac{1}{2}\right)^{2s} = \left(s+\frac{1}{2}\right)\left(\ell+\frac{1}{2}\right)^{s}\left(\ell+\frac{1}{2}\right)^{s-1} \leq 2es\left(\ell+\frac{1}{2}\right)^{s}\left(\ell-\frac{1}{2}\right)^{s-1} \tag{3.51}$$

Case 2 ($\ell > s$): In this case, $\gamma_{s,\ell} = 1$. Applying a binomial expansion and then (3.50) gives

$$\begin{aligned}
\left(\ell+\frac{1}{2}\right)^{2s} - \left(\ell-\frac{1}{2}\right)^{2s} &= \sum_{k=0}^{2s}\binom{2s}{k}\left(\ell-\frac{1}{2}\right)^{k} - \left(\ell-\frac{1}{2}\right)^{2s} = \sum_{k=0}^{2s-1}\binom{2s}{k}\left(\ell-\frac{1}{2}\right)^{k} \\
&= \sum_{k=0}^{2s-1}\frac{2s}{2s-k}\binom{2s-1}{k}\left(\ell-\frac{1}{2}\right)^{k} \leq 2s\sum_{k=0}^{2s-1}\binom{2s-1}{k}\left(\ell-\frac{1}{2}\right)^{k} \\
&= 2s\left(\ell+\frac{1}{2}\right)^{2s-1} = 2s\left(\ell+\frac{1}{2}\right)^{s}\left(\ell+\frac{1}{2}\right)^{s-1} \leq 2es\left(\ell+\frac{1}{2}\right)^{s}\left(\ell-\frac{1}{2}\right)^{s-1}
\end{aligned} \tag{3.52}$$

□

Lemma 3.2.5 (Semi-norm recurrence)**.** *Let $s \geq 1$, $q \in L^1([0,T]; H^{r,s})$ and $g \in H^{r,s}$. Then for all $t \in [0,T]$,*

$$|\psi^N|_{H^{r,s}}(t) \leq Cs\mathcal{A}_0[|\psi^N|_{H^{r+1,s-1}}](t) + |\mathcal{P}_N g|_{H^{r,s}}F_0(t) + \varepsilon\mathcal{A}_0[|\mathcal{P}_N q|_{H^{r,s}}](t), \tag{3.53}$$

where F_0 is defined in (3.39) *and C is a constant independent of the data.*

Proof. We assume first that $r = 0$ and focus on the last term in (3.46). It follows from (i) the induced norm bound $\|a_\ell^{(i)}\|_2 \leq 4$ [44], (ii) the bounds in Lemma 3.2.4, (iii) the Cauchy-Schwarz inequality, and (iv) the $H^{r,s}$ semi-norm definitions in (3.33) and (3.34) that

$$\begin{aligned}
&\sum_{\ell=s}^{N}\left(\left(\ell+\frac{1}{2}\right)^{2s} - \gamma_{s,\ell}\left(\ell-\frac{1}{2}\right)^{2s}\right)\left(\psi_\ell^N, \left(a_\ell^{(i)}\right)^T\partial_{x_i}\psi_{\ell-1}^N\right)_{L^2(X)} \\
&\quad\leq 2es\sum_{\ell=s}^{N}\left(\ell+\frac{1}{2}\right)^{s}\left(\ell-\frac{1}{2}\right)^{s-1}\|a_\ell^{(i)}\|_2\ \ \|\psi_\ell^N\|_{L^2(X)}\ \ \|\partial_{x_i}\psi_{\ell-1}^N\|_{L^2(X)} \\
&\quad\leq Cs\left(\sum_{\ell=s}^{N}\left(\ell+\frac{1}{2}\right)^{2s}|\psi_\ell^N|_{L^2(X)}^2\right)^{1/2}\left(\sum_{\ell=s}^{N}\left(\ell-\frac{1}{2}\right)^{2(s-1)}|\partial_{x_i}\psi_{\ell-1}^N|_{L^2(X)}^2\right)^{1/2} \\
&\quad\leq Cs|\psi^N|_{H^{0,s}}|\partial_{x_i}\psi^N|_{H^{0,s-1}}.
\end{aligned} \tag{3.54}$$

Applying the bound above to the right-hand side of (3.46) and applying Lemma .0.1 gives (3.53). The case $r > 1$ can be handled by differentiating the P_N equations in space and then repeating the arguments above. □

For a general function $\phi \in L^p([\alpha,\beta];H^{r,s})$, let $|\phi|_{L^p([\alpha,\bullet];H^{0,r+s})} : [\alpha,\beta] \to \mathbb{R}$ be the map defined

$$|\phi|_{L^p([\alpha,\bullet];H^{r,s})}(\tau) = |\phi|_{L^p([\alpha,\tau];H^{r,s})}.$$

Lemma 3.2.6 (Stability of higher-order semi-norms)**.** *Let $q \in L^1([0,T];H^{r,0})$ and $g \in H^{r,0}$. If $t \in [0,T]$, then*

$$|\psi^N|_{H^{r,0}}(t) \leq |\mathcal{P}_N g|_{H^{r,0}} + |\mathcal{P}_N q|_{L^1([0,t];H^{r,0})}. \tag{3.55}$$

If, in addition, $s \geq 1$, $q \in L^1([0,T];H^{i,j})$ and $g \in H^{i,j}$ for each i,j such that $0 \leq i \leq r$, $0 \leq j \leq s$, and $i+j = r+s$,

$$\begin{aligned}|\psi^N|_{H^{r,s}}(t) \leq{}& C_s s! \mathcal{A}_0^s \left[|\mathcal{P}_N g|_{H^{r+s,0}} + |\mathcal{P}_N q|_{L^1([0,\bullet];H^{r+s,0})}\right](t) \\ &+ C_s \sum_{i=0}^{s-1} \frac{s!}{(s-i)!} \mathcal{A}_0^i [F_0 |\mathcal{P}_N g|_{H^{r+i,s-i}}](t) + C_s \varepsilon \sum_{i=0}^{s-1} \frac{s!}{(s-i)!} \mathcal{A}_0^{i+1} [|\mathcal{P}_N q|_{H^{r+i,s-i}}](t),\end{aligned} \tag{3.56}$$

where F_0 is given in (3.39) *where C_s is a constant depending only on s.*

Proof. First we will prove (3.55) for $s = 0$, in which case $H^{r,0} = L^2(X \times \mathbb{S}^2)$. From (3.47) and the Cauchy-Schwarz inequality,

$$\frac{\varepsilon}{2} \frac{d}{dt} \|\psi^N\|^2_{L^2(X\times\mathbb{S}^2)} \leq \varepsilon (\mathcal{P}_N q, \psi^N)_{L^2(X\times\mathbb{S}^2)} \leq \varepsilon \|\mathcal{P}_N q\|_{L^2(X\times\mathbb{S}^2)} \|\psi^N\|_{L^2(X\times\mathbb{S}^2)} \tag{3.57}$$

an application of Lemma .0.1, gives

$$\|\psi^N\|_{L^2(X\times\mathbb{S}^2)}(t) \leq \|\mathcal{P}_N g\|_{L^2(X\times\mathbb{S}^2)} + \|\mathcal{P}_N q\|_{L^1([0,t];L^2(X\times\mathbb{S}^2))}. \tag{3.58}$$

For $r > 0$, $\phi^N = \partial_{x_{i_1} x_{i_2} \ldots x_{i_r}} \psi^N$ satisfies (3.7a), with initial condition given by the following $\phi^N\big|_{t=0} = \mathcal{P}_N(\partial_{x_{i_1} x_{i_2} \ldots x_{i_r}} g) = \partial_{x_{i_1} x_{i_2} \ldots x_{i_r}} \mathcal{P}_N g$ and source $\mathcal{P}_N(\partial_{x_{i_1} x_{i_2} \ldots x_{i_r}} q) = \partial_{x_{i_1} x_{i_2} \ldots x_{i_r}} \mathcal{P}_N q$. Repeating the argument above gives, in analogy with (3.58),

$$\|\partial_{x_{i_1} x_{i_2} \ldots x_{i_r}} \psi^N\|_{L^2(X\times\mathbb{S}^2)}(t) \leq \|\partial_{x_{i_1} x_{i_2} \ldots x_{i_r}} \mathcal{P}_N g\|_{L^2(X\times\mathbb{S}^2)} + \|\partial_{x_{i_1} x_{i_2} \ldots x_{i_r}} \mathcal{P}_N q\|_{L^1([0,t];L^2(X\times\mathbb{S}^2))}, \tag{3.59}$$

Summing (3.59) over each permutation of the r derivatives $\partial_{x_{i_1}}\partial_{x_{i_2}}\cdots\partial_{x_{i_r}}$ and using the definitions in (3.34) recovers (3.55).

We next prove (3.56). Let

$$b_{r,s}(t) = |\mathcal{P}_N g|_{H^{r,s}} F_0(t) + \varepsilon \mathcal{A}_0[|\mathcal{P}_N q|_{H^{r,s}}](t) \quad \text{and} \quad c_{r,s}(t) = |\psi^N|_{H^{r,s}}(t), \tag{3.60}$$

where F_0 is defined in (3.39). Then Lemma 3.2.5 gives the following recursion relation:

$$c_{r,s}(t) \leq Cs\mathcal{A}_0[c_{r+1,s-1}](t) + b_{r,s}(t), \qquad s \geq 1. \tag{3.61}$$

Unrolling this recursion in s gives

$$c_{r,s}(t) \leq s! C^s \mathcal{A}_0^s[c_{r+s,0}](t) + \sum_{i=0}^{s-1} \frac{s!}{(s-i)!} C^i \mathcal{A}_0^i[b_{r+i,s-i}](t), \tag{3.62}$$

and translating back to the semi-norms with (3.60) gives

$$\begin{aligned} |\psi^N|_{H^{r,s}}(t) \leq\ & s! C^s \mathcal{A}_0^s[|\psi^N|_{H^{r+s,0}}](t) \\ & + \sum_{i=0}^{s-1} \frac{s!}{(s-i)!} C^i |\mathcal{P}_N g|_{H^{r+i,s-i}} \mathcal{A}_0^i[F_0](t) + \varepsilon \sum_{i=0}^{s-1} \frac{s!}{(s-i)!} C^{i+1} \mathcal{A}_0^{i+1}[|\mathcal{P}_N q|_{H^{r+i,s-i}}](t). \end{aligned} \tag{3.63}$$

We then apply (3.55) (with r replaced with $r+s$) to the first term on the right hand side of (3.63).

□

3.2.2.2 Estimates for continuous system

In this section we extend the stability results for $|\psi^N|_{H^{r,s}}$ to $|\psi|_{H^{r,s}}$, using the infinite moment hierarchy in (3.9). These estimates will be useful in deriving consistency estimates. For the case $r = 0$, as in the previous section, we test each equation in (3.9) by $b_\ell \psi_\ell$, integrate by parts, and sum over $\ell \geq s$. The result is analogous to (3.46), namely

$$\begin{aligned} \frac{\varepsilon}{2}\partial_t |\psi|^2_{H^{0,s}} + \frac{\sigma}{\varepsilon}|\psi|^2_{H^{0,s}} &= \varepsilon (q,\psi)_{H^{0,s}} \\ &- \sum_{i=1}^{3}\sum_{\ell=s}^{\infty} \left(\left(\ell + \frac{1}{2}\right)^{2s} - \gamma_{s,\ell}\left(\ell - \frac{1}{2}\right)^{2s} \right) \left(\psi_\ell, \left(a_\ell^{(i)}\right)^T \partial_{x_i} \psi_{\ell-1} \right)_{L^2(X)}, \end{aligned} \tag{3.64}$$

As before, when $s = 0$, (3.64) recovers the usual L^2 stability result:

$$\frac{\varepsilon}{2}\frac{d}{dt}\|\psi\|^2_{L^2(X\times\mathbb{S}^2)} + \frac{\sigma}{\varepsilon}\|\psi - \overline{\psi}\|^2_{L^2(X\times\mathbb{S}^2)} = \varepsilon(q,\psi)^2_{L^2(X\times\mathbb{S}^2)} \tag{3.65}$$

The following results are continuous analogues to Lemmas 3.2.5 and 3.2.6. Their proofs are nearly identical, so we only give a brief summary.

Lemma 3.2.7 (Semi-norm recurrence for the continuous system)**.** *Let $s \geq 1$, $q \in L^1([0,t]; H^{r,s})$ and $g \in H^{r,s}$. Then for all $t \in [0,T]$,*

$$|\psi|_{H^{r,s}}(t) \leq Cs\mathcal{A}_0[|\psi|_{H^{r+1,s-1}}](t) + |g|_{H^{r,s}}F_0(t) + \varepsilon\mathcal{A}_0[|q|_{H^{r,s}}](t), \tag{3.66}$$

where C is a constant independent of the data.

Summary of Proof. The proof follows the same lines and with same constants as in Lemma 3.2.5 after changing ψ_ℓ^N by ψ_ℓ and then taking all the sums to infinity. To generalize the proof for $r > 1$, we differentiate the system (3.9) in space and repeat the process. □

Lemma 3.2.8 (Stability of higher order semi-norms for the continuous system)**.** *If $q \in L^1([0,T]; H^{r,0})$ and $g \in H^{r,0}$, then for all $t \in [0,T]$,*

$$|\psi|_{H^{r,0}}(t) \leq |g|_{H^{r,0}} + |q|_{L^1([0,t];H^{r,0})}. \tag{3.67}$$

If, in addition, $q \in L^1([0,T]; H^{i,j})$ and $g \in H^{i,j}$ for each i, j such that $0 \leq i \leq r$, $0 \leq j \leq s$, and $i + j = r + s$,

$$\begin{aligned} |\psi|_{H^{r,s}}(t) &\leq C_s s!\mathcal{A}_0^s\left[|g|_{H^{r+s,0}} + |q|_{L^1([0,\bullet];H^{r+s,0})}\right](t) \\ &\quad + C_s\sum_{i=0}^{s-1}\frac{s!}{(s-i)!}\mathcal{A}_0^i[F_0|g|_{H^{r+i,s-i}}](t) + C_s\varepsilon\sum_{i=0}^{s-1}\frac{s!}{(s-i)!}\mathcal{A}_0^{i+1}[|q|_{H^{r+i,s-i}}](t), \end{aligned} \tag{3.68}$$

where C_s is a constant depending only on s.

Corollary 3.2.1 (Isotropic data and zero initial condition for the continuous systems)**.** *Let $s \geq 1$. In the special case that $g = 0$ and q is isotropic, then if $t \in [0,T]$,*

$$|\psi|_{H^{r,s}}(t) \leq C_s s!\mathcal{A}_0^s[|q|_{L^1([0,\bullet];H^{r+s,0})}](t). \tag{3.69}$$

Proof of Lemma 3.2.8. With the obvious changes, the proof is the same line by line as proof in Lemma 3.2.6, with some key steps replace by their continuous counterparts, namely, in the initial step we use (3.65) instead of (3.47) and we invoke Lemma 3.2.7 instead of Lemma 3.2.5. □

3.2.3 P_N error analysis

In this section we will analyze the error produced by the solution of (3.3) when the P_N approximation is used.

Definition 3.2.1 (P_N Error)**.** *The* P_N error *is*

$$e^N(t) = \psi(t) - \psi^N(t) = \eta^N(t) + \xi^N(t), \tag{3.70}$$

where $\eta^N = \psi - \mathcal{P}_N\psi$ *is the* consistency error *and* $\xi^N = \mathcal{P}_N\psi - \psi^N$ *is the* stability error.

Lemma 3.2.9. *For all* $t \in [0,T]$, ξ^N *is controlled by* η^N *via the following estimate:*

$$\|\xi^N\|_{L^2(X\times\mathbb{S}^2)}(t) \leq \frac{1}{\varepsilon}\int_0^t \|\mathcal{P}_N(\Omega\cdot\nabla_x\eta^N\|_{L^2(X\times\mathbb{S}^2)}(\tau)\,d\tau. \tag{3.71}$$

Proof. Applying the projection $\mathcal{P}_N$ to (3.3) and subtracting (3.7a) yields a P_N equation for ξ^N with a source that depends on η^N:

$$\varepsilon\partial_t\xi^N + \mathcal{P}_N(\Omega\cdot\nabla_x\xi^N) + \frac{\sigma}{\varepsilon}\xi^N = \frac{\sigma}{\varepsilon}\overline{\xi^N} - \mathcal{P}_N(\Omega\cdot\nabla_x\eta^N), \qquad \xi^N\big|_{t=0} = 0. \tag{3.72}$$

Thus (3.72) follows immediately from the bound (3.58), replacing g by zero and q by $\varepsilon^{-1}\mathcal{P}_N(\Omega\cdot\nabla_x\eta^N)$. □

Lemma 3.2.10. *Let* $t \in [\alpha,\beta) \subseteq [0,T]$, *then*

$$\begin{aligned}\|\eta^N\|_{L^2(X\times\mathbb{S}^2)}(t) \leq e^{-\frac{\sigma(t-\alpha)}{\varepsilon^2}}\|\eta^N\|_{L^2(X\times\mathbb{S}^2)}(\alpha) + \varepsilon\mathcal{A}_\alpha[\|\widetilde{\mathcal{P}}_N q\|_{L^2(X\times\mathbb{S}^2)}](t) \\ +\mathcal{A}_\alpha[\|\widetilde{\mathcal{P}}_N(\Omega\cdot\nabla_x\psi)\|_{L^2(X\times\mathbb{S}^2)}](t)\end{aligned} \tag{3.73}$$

Proof. Applying the projection $\mathcal{P}_N$ to (3.3), and subtracting it from (3.3), we see that η^N satisfies

$$\partial_t\eta^N - \frac{1}{\varepsilon}\mathcal{P}_N(\Omega\cdot\nabla_x\eta^N) + \frac{\sigma}{\varepsilon^2}\eta^N = \widetilde{\mathcal{P}}_N q - \frac{1}{\varepsilon}\left[\Omega\cdot\nabla_x\psi - \mathcal{P}_N(\Omega\cdot\nabla_x\mathcal{P}_N\psi)\right] \tag{3.74}$$

Since for any $\phi \in L^2(\mathbb{S}^2)$, $(\mathcal{P}_N\phi, \eta^N)_{L^2(\mathbb{S}^2)} = 0$, testing the equation above against η^N gives

$$\begin{aligned}\frac{1}{2}\frac{d}{dt}\|\eta^N\|^2_{L^2(X\times\mathbb{S}^2)} + \frac{\sigma}{\varepsilon^2}\|\eta^N\|^2_{L^2(X\times\mathbb{S}^2)} &= (\widetilde{\mathcal{P}}_N q, \eta^N)_{L^2(X\times\mathbb{S}^2)} - \frac{1}{\varepsilon}(\Omega\cdot\nabla_x\psi, \eta^N)_{L^2(X\times\mathbb{S}^2)} \\ &= (\widetilde{\mathcal{P}}_N q, \eta^N)_{L^2(X\times\mathbb{S}^2)} - \frac{1}{\varepsilon}(\widetilde{\mathcal{P}}_N(\Omega\cdot\nabla_x\psi), \eta^N)_{L^2(X\times\mathbb{S}^2)} \\ &\leq (\|\widetilde{\mathcal{P}}_N q\|_{L^2(X\times\mathbb{S}^2)} + \frac{1}{\varepsilon}\|\widetilde{\mathcal{P}}_N(\Omega\cdot\nabla_x\psi)\|_{L^2(X\times\mathbb{S}^2)})\|\eta^N\|_{L^2(X\times\mathbb{S}^2)}\end{aligned} \tag{3.75}$$

The conclusion then follows from Lemma .0.1. □

An immediate corollary of Lemma 3.2.10 is the following:

Lemma 3.2.11. *Let $s \geq 0$ and $N \geq \max\{0, s-1\}$. If $g \in H^{0,s}$ and $q \in L^\infty([0,T]; H^{0,s})$. Then we have*

$$\|\eta^N\|_{L^2(X\times\mathbb{S}^2)}(T) \leq \frac{1}{(N+1)^s}\left[e^{-\sigma T/\varepsilon^2}|g|_{H^{0,s}} + \varepsilon|q|_{L^\infty([0,T];H^{0,s})}\mathcal{A}_0[\mathbb{1}](T) + \frac{T}{\varepsilon}\sup_{\tau\in[0,T]}|\psi|_{H^{1,s}}(\tau)\right]. \tag{3.76}$$

Proof. We apply Lemma 3.2.10 with $\alpha = 0$. In this case $\|\eta^N\|_{L^2(X\times\mathbb{S}^2)}(\alpha) = \|\widetilde{\mathcal{P}}_N g\|$. Meanwhile, by Lemma 3.2.2,

$$\|\widetilde{\mathcal{P}}_N g\|_{L^2(X\times\mathbb{S}^2)} \leq \frac{1}{(N+1)^s}|g|_{H^{0,s}} \quad \text{and} \quad \|\widetilde{\mathcal{P}}_N q\|_{L^2(X\times\mathbb{S}^2)} \leq \frac{1}{(N+1)^s}|q|_{H^{0,s}}. \tag{3.77}$$

Thus since $\mathcal{A}_\alpha[f](t) \leq \varepsilon^{-1}(t-\alpha)\sup_{\tau\in[\alpha,t]} f(\tau)$ (cf. (3.40) with $k=1$), another application of Lemma 3.2.2 gives

$$\begin{aligned}\mathcal{A}_0[\|\widetilde{\mathcal{P}}_N(\Omega\cdot\nabla_x\psi)\|_{L^2(X\times\mathbb{S}^2)}](T) &\leq \frac{T}{\varepsilon}\sup_{\tau\in[0,T]}\|\widetilde{\mathcal{P}}_N(\Omega\cdot\nabla_x\psi)\|_{L^2}(\tau) \\ &\leq \frac{T}{\varepsilon(N+1)^s}\sup_{\tau\in[0,T]}|\Omega\cdot\nabla_x\psi|_{H^{0,s}}(\tau) \leq \frac{T}{\varepsilon(N+1)^s}\sup_{\tau\in[0,T]}|\psi|_{H^{1,s}}(\tau)\end{aligned} \tag{3.78}$$

Plugging the preceding bound into Lemma 3.2.10 gives the result. □

Lemma 3.2.12 (a priori estimate)**.** *Let $s \geq 1$ and $N \geq s-1$. Let $q \in L^\infty([0,T]; H^{i,j})$ and $g \in H^{i,j}$ for each i, j such that $0 \leq i \leq r$, $0 \leq j \leq s$, and $i + j = r + s$. Then for all $t \in [0,T]$,*

$$
\begin{aligned}
|\psi|_{H^{r,s}}(t) \leq C_s \Bigg\{ & \left[|g|_{H^{r+s,}} + t|q|_{L^\infty([0,t);H^{r+s,0})} \right] \min\left(\frac{\varepsilon^s s!}{\sigma^s}, \left(\frac{t}{\varepsilon}\right)^s \right) \\
& + e^{-\sigma t/\varepsilon^2} \sum_{i=0}^{s-1} |g|_{H^{r+i,s-i}} \binom{s}{i} \frac{t^i}{\varepsilon^i} \\
& + \varepsilon \sum_{i=0}^{s-1} |q|_{L^\infty([0,t];H^{r+i,s-i})} \frac{s!}{(s-i)!} \min\left(\frac{\varepsilon^{i+1}}{\sigma^{i+1}}, \frac{1}{(i+1)!} \left(\frac{t}{\varepsilon}\right)^{i+1} \right) \Bigg\}.
\end{aligned} \tag{3.79}
$$

Proof. Recall the stability estimate from Lemma 3.2.8:

$$
\begin{aligned}
|\psi|_{H^{r,s}}(t) \leq{}& C_s s! \mathcal{A}_0^s \left[|g|_{H^{r+s,0}} + |q|_{L^1([0,\bullet];H^{r+s,0})} \right](t) \\
& + C_s \sum_{i=0}^{s-1} \frac{s!}{(s-i)!} \mathcal{A}_0^i [F_0 |g|_{H^{r+i,s-i}}](t) + C_s \varepsilon \sum_{i=0}^{s-1} \frac{s!}{(s-i)!} \mathcal{A}_0^{i+1} [|q|_{H^{r+i,s-i}}](t).
\end{aligned} \tag{3.80}
$$

Substituting the bounds for $\mathcal{A}_0[\mathbb{1}]$ and the formula for $\mathcal{A}_0[F_\alpha]$ from Lemma 3.2.3 into the above estimate yields the stated result. □

Theorem 3.2.1 (P_N error)**.** *Let $s \geq 1$ and $N \geq s-1$. Let $q \in L^\infty([0,T]; H^{i,j})$ and $g \in H^{i,j}$ for each i, j such that $0 \leq j \leq s$, $i + j \leq s+1$. Then*

$$
\begin{aligned}
\|e^N\|_{L^2(X\times\mathbb{S}^2)}(T) \leq{}& \frac{e^{-\sigma T/\varepsilon^2}}{(N+1)^s} |g|_{H^{0,s}} + \frac{1}{(N+1)^s} |q|_{L^\infty([0,T];H^{0,s})} \min\left(\frac{\varepsilon^2}{\sigma}, T \right) \\
& + \frac{2C_s}{(N+1)^s} \Bigg\{ \left(|g|_{H^{s+1,0}} + T|q|_{L^\infty([0,T];H^{s+1,0})} \right) \min\left(\frac{\varepsilon^{s-1} s! T}{\sigma^s}, \left(\frac{T}{\varepsilon}\right)^{s+1} \right) \\
& \quad + e^{-\sigma T/\varepsilon^2} \sum_{i=0}^{s-1} |g|_{H^{1+i,s-i}} \binom{s}{i} \frac{T^{i+1}}{\varepsilon^{i+1}} \\
& \quad + \sum_{i=0}^{s-1} |q|_{L^\infty([0,T];H^{1+i,s-i})} \frac{s!}{(s-i)!} \min\left(\frac{\varepsilon^{i+1} T}{\sigma^{i+1}}, \frac{1}{(i+1)!} \frac{T^{i+2}}{\varepsilon^{i+1}} \right) \Bigg\}
\end{aligned} \tag{3.81}
$$

Proof. By the triangle inequality, Lemma 3.2.9,

$$\begin{aligned}
\|e^N\|_{L^2(X\times\mathbb{S}^2)}(T) &\le \|\eta^N\|_{L^2(X\times\mathbb{S}^2)}(T) + \|\xi^N\|_{L^2(X\times\mathbb{S}^2)}(T) \\
&\le \|\eta^N\|_{L^2(X\times\mathbb{S}^2)}(T) + \frac{1}{\varepsilon}\int_0^T \|\mathcal{P}_N(\Omega\cdot\nabla_x\eta^N)\|_{L^2(X\times\mathbb{S}^2)}(\tau)\,d\tau \\
&\le \|\eta^N\|_{L^2(X\times\mathbb{S}^2)}(T) + \frac{1}{\varepsilon}\int_0^T \|\nabla_x\eta^N\|_{L^2(X\times\mathbb{S}^2)}(\tau)\,d\tau \\
&\le \|\eta^N\|_{L^2(X\times\mathbb{S}^2)}(T) + \frac{1}{(N+1)^s}\frac{T}{\varepsilon}\sup_{\tau\in[0,T]}|\psi|_{H^{1,s}}(\tau) \\
&\le \frac{e^{-\sigma T/\varepsilon^2}}{(N+1)^s}|g|_{H^{0,s}} + \frac{\varepsilon}{(N+1)^s}|q|_{L^\infty([0,T];H^{0,s})}\mathcal{A}_0[\mathbb{1}](T) \\
&+ \frac{2T}{\varepsilon(N+1)^s}\sup_{\tau\in[0,T]}|\psi|_{H^{1,s}}(\tau)
\end{aligned} \tag{3.82}$$

In the last two lines, we applied spectral estimate in Lemma 3.2.2. Inthe last line we used Lemma 3.2.11. Applying Lemma 3.2.12 with $r = 1$ yields the result. □

Corollary 3.2.2 (P_N error for Isotropic data)**.** *Let $s \ge 1$ and $N \ge s-1$. Let $q \in L^\infty([0,T];H^{s+1,0})$ and $g \in H^{s+1,0}$. If g and q are isotropic, then*

$$\|e^N\|_{L^2(X\times\mathbb{S}^2)}(T) \le \frac{2C_s}{(N+1)^s}\Big(|g|_{H^{s+1,0}} + T|q|_{L^\infty([0,T];H^{s+1,0})}\Big)\min\left(\frac{\varepsilon^{s-1}s!T}{\sigma^s}, \left(\frac{T}{\varepsilon}\right)^{s+1}\right) \tag{3.83}$$

3.3 Hybrid error analysis

3.3.1 A priori estimates of the uncollided component

Since our goal is to derive error estimates which only depend on data, and $\psi_{\mathrm{u},m}$ appears as a source in the collided equation, we require the following a priori estimates on $\psi_{\mathrm{u},m}$ to bound $|\psi_{\mathrm{c},m}|_{H^{1,s}}$ in the proof of Theorem 3.3.2.

Lemma 3.3.1 (Stability of the uncollided component)**.** *Let $1 \le m \le M$, $q \in L^\infty([0,T];H^{r,0})$, and $g \in H^{r,0}$ for some $r \ge 0$. Then for all $t \in [t_{m-1}, t_m)$,*

$$|\psi_{\mathrm{u},m}|_{H^{r,0}}(t) \le e^{-\sigma(t-t_{m-1})/\varepsilon^2}|g|_{H^{r,0}} + e^{-\sigma(t-t_{m-1})/\varepsilon^2}|q|_{L^1([0,t_{m-1}];H^{r,0})} + \varepsilon\mathcal{A}_{t_{m-1}}[|q|_{H^{r,0}}](t), \tag{3.84a}$$

$$|\psi_{\mathrm{u},m}|_{H^{r,0}}(t_m) + \frac{\sigma}{\varepsilon^2}|\psi_{\mathrm{u},m}|_{L^1([t_{m-1},t_m];H^{r,0})} \le |g|_{H^{r,0}} + |q|_{L^1([0,t_m];H^{r,0})}, \tag{3.84b}$$

$$|\psi_{\mathrm{u},m}|_{L^1([t_{m-1},t];H^{r,0})}(t) \le \varepsilon(|g|_{H^{r,0}} + t|q|_{L^\infty([0,t];H^{r,0})})\mathcal{A}_{t_{m-1}}[\mathbb{1}](t). \tag{3.84c}$$

Proof. We prove the result only for $r = 0$ since the other cases are obtained by applying the same techniques to (3.12a) differentiated r times in space. Testing (3.12a) with $\psi_{\mathrm{u},m}$ and applying Cauchy-Schwarz inequality, we obtain the following differential inequality

$$\frac{1}{2}\frac{d}{dt}\|\psi_{\mathrm{u},m}\|^2 + \frac{\sigma}{\varepsilon^2}\|\psi_{\mathrm{u},m}\|^2 \leq \|q\|\|\psi_{\mathrm{u},m}\|. \tag{3.85}$$

An application of (.4) in Lemma .0.1, and the fact that $\psi_{\mathrm{u},m}(t_{m-1}) = \psi(t_{m-1})$ gives

$$\|\psi_{\mathrm{u},m}\|_{L^2(X\times\mathbb{S}^2)}(t) \leq e^{-\sigma(t-t_{m-1})/\varepsilon^2}\|\psi\|_{L^2(X\times\mathbb{S}^2)}(t_{m-1}) + \varepsilon\mathcal{A}_{t_{m-1}}[\|q\|_{L^2(X\times\mathbb{S}^2)}](t). \tag{3.86}$$

Using Lemma 3.2.8 on $\|\psi\|_{L^2(X\times\mathbb{S}^2)}(t_{m-1})$ gives the first result (3.84a). An application of (.3) in Lemma .0.1 over $[t_{m-1}, t_m)$ gives

$$\|\psi_{\mathrm{u},m}\|_{L^2(X\times\mathbb{S}^2)}(t_m)+\frac{\sigma}{\varepsilon^2}\|\psi_{\mathrm{u},m}\|_{L^1([t_{m-1},t_m],L^2(X\times\mathbb{S}^2))} \leq \|q\|_{L^1([t_{m-1},t_m],L^2(X\times\mathbb{S}^2))}+\|\psi\|_{L^2(X\times\mathbb{S}^2)}(t_{m-1}), \tag{3.87}$$

and then another application of Lemma 3.2.8 to bound $\|\psi\|_{L^2(X\times\mathbb{S}^2)}(t_{m-1})$ gives (3.84b). Estimate (3.84c) is obtained from integrating the first estimate (3.84a) from t_{m-1} to t. □

3.3.2 Hybrid error analysis

In this section we will analyze the error in the hybrid method using the formulation in (3.14).

Definition 3.3.1 (Hybrid errors)**.** *Let $1 \leq m \leq M$ and $t \in [t_{m-1}, t_m)$. The m-th* hybrid error *is*

$$e_m^N(t) = e_{\mathrm{u},m}^N(t) + e_{\mathrm{c},m}^N(t), \tag{3.88}$$

where $e_{\mathrm{u},m}^N(t) = \psi_{\mathrm{u},m}(t) - \psi_{\mathrm{u},m}^N(t)$ and $e_{\mathrm{c},m}^N(t) = \psi_{\mathrm{c},m}(t) - \psi_{\mathrm{c},m}^N(t)$ are the m-th errors in the uncollided and collided components. The collided error can be further decomposed as

$$\eta_{\mathrm{c},m}^N(t) = \psi_{\mathrm{c},m}(t) - \mathcal{P}_N\psi_{\mathrm{c},m}(t) \qquad \textit{and} \qquad \xi_{\mathrm{c},m}^N = \mathcal{P}_N\psi_{\mathrm{c},m}(t) - \psi_{\mathrm{c},m}^N(t) \tag{3.89}$$

so that $e_{\mathrm{c},m}^N(t) = \eta_{\mathrm{c},m}^N(t) + \xi_{\mathrm{c},m}^N(t)$. Here $\eta_{\mathrm{c},m}^N$ is the m-th collided consistency error *and $\xi_{\mathrm{c},m}^N$ is the* m-th collided stability error. *The error e_M^N is simply called the* hybrid error.

This next lemma gives a one-step analysis of the growth of the error in the uncollided and collided components from t_{m-1} to t_m.

Lemma 3.3.2. *Let $1 \leq m \leq M$, then if $t \in [t_{m-1}, t_m)$, the m-th uncollided and collided errors satisfy, respectively,*

$$\|e^N_{\mathrm{u},m}\|_{L^2(X\times\mathbb{S}^2)}(t) \leq e^{-\sigma(t-t_{m-1})/\varepsilon^2}\|e^N_{m-1}\|_{L^2(X\times\mathbb{S}^2)}(t^-_{m-1}), \tag{3.90a}$$

$$\begin{aligned}\|e^N_{\mathrm{c},m}\|_{L^2(X\times\mathbb{S}^2)}(t) \leq &\left(1 - e^{-\sigma(t-t_{m-1})/\varepsilon^2}\right)\|e^N_{m-1}\|_{L^2(X\times\mathbb{S}^2)}(t^-_{m-1}) \\ &+ \|\eta^N_{\mathrm{c},m}\|_{L^2(X\times\mathbb{S}^2)}(t) + \frac{1}{\varepsilon}\|\Omega\cdot\nabla_x\eta^N_{\mathrm{c},m}\|_{L^1([t_{m-1},t];L^2(X\times\mathbb{S}^2))}.\end{aligned} \tag{3.90b}$$

Proof. Subtracting (3.14a) from (3.12a), yields the following evolution equation for $e^N_{\mathrm{u},m}$,

$$\varepsilon\partial_t e^N_{\mathrm{u},m} + \Omega\cdot\nabla_x e^N_{\mathrm{u},m} + \frac{\sigma}{\varepsilon}e^N_{\mathrm{u},m} = 0, \qquad e^N_{\mathrm{u},m}\big|_{t=t_{m-1}} = e^N_{m-1}(t^-_{m-1}), \tag{3.91}$$

where $e^N_0(t^-_0) = 0$. Thus applying (3.84a) from Lemma 3.3.1, with $r = 0$ and a zero source term yields (3.90a). To prove (3.90b), we subtract from (3.14b) the projection applied to (3.12b). This gives the following P_N equation for $\xi^N_{\mathrm{c},m}$

$$\varepsilon\partial_t\xi^N_{\mathrm{c},m} + \mathcal{P}_N(\Omega\cdot\nabla_x\xi^N_{\mathrm{c},m}) + \frac{\sigma}{\varepsilon}\xi^N_{\mathrm{c},m} = \frac{\sigma}{\varepsilon}(\overline{\xi^N_{\mathrm{c},m}} + \overline{e^N_{\mathrm{u},m}}) - \mathcal{P}_N(\Omega\cdot\nabla_x\eta^N_{\mathrm{c},m}), \tag{3.92a}$$

$$\xi^N_{\mathrm{c},m}\big|_{t=t_{m-1}} = 0. \tag{3.92b}$$

We apply Lemma 3.2.6 to (3.92a) with zero initial data and source $\varepsilon^{-2}\sigma\overline{e^N_{\mathrm{u},m}} - \varepsilon^{-1}\mathcal{P}_N(\Omega\cdot\nabla_x\eta^N_{\mathrm{c},m})$. Combined with bound (3.90a), the estimate on $\xi^N_{\mathrm{c},m}$ becomes

$$\begin{aligned}\|\xi^N_{\mathrm{c},m}\|_{L^2(X\times\mathbb{S}^2)}(t) &\leq \frac{\sigma}{\varepsilon^2}\int_{t_{m-1}}^t \|e^N_{\mathrm{u},m}\|_{L^2(X\times\mathbb{S}^2)}(\tau)\,d\tau + \frac{1}{\varepsilon}\int_{t_{m-1}}^t \|\Omega\cdot\nabla_x\eta^N_{\mathrm{c},m}\|_{L^2(X\times\mathbb{S}^2)}(\tau)\,d\tau \\ &\leq \frac{\sigma}{\varepsilon^2}\|e^N_{m-1}\|(t^-_{m-1})\int_{t_{m-1}}^t e^{-\sigma(\tau-t_{m-1})/\varepsilon^2}\,d\tau \\ &\quad + \frac{1}{\varepsilon}\int_{t_{m-1}}^t \|\Omega\cdot\nabla_x\eta^N_{\mathrm{c},m}\|_{L^2(X\times\mathbb{S}^2)}(\tau)\,d\tau \\ &= \left(1 - e^{-\sigma(t-t_{m-1})/\varepsilon^2}\right)\|e^N_{m-1}\|_{L^2(X\times\mathbb{S}^2)}(t^-_{m-1}) \\ &\quad + \frac{1}{\varepsilon}\int_{t_{m-1}}^t \|\Omega\cdot\nabla_x\eta^N_{\mathrm{c},m}\|_{L^2(X\times\mathbb{S}^2)}(\tau)\,d\tau.\end{aligned} \tag{3.93}$$

Adding $\|\eta^N_{\mathrm{c},m}\|_{L^2(X\times\mathbb{S}^2)}(t)$ to the both sides recovers (3.90b). □

Now we can state an error for all time that only depends on the approximation properties of the spherical harmonic discretization on the solution.

Theorem 3.3.1. *The hybrid error e_M^N satisfies*

$$\|e_M^N\|_{L^2(X\times\mathbb{S}^2)}(t_M^-) \leq \sum_{m=1}^{M} \left(\|\eta_{\mathrm{c},m}^N\|_{L^2(X\times\mathbb{S}^2)}(t_m^-) + \frac{1}{\varepsilon}\|\Omega\cdot\nabla_x\eta_{\mathrm{c},m}^N\|_{L^1([t_{m-1},t_m];L^2(X\times\mathbb{S}^2))} \right). \quad (3.94)$$

Proof. Adding the inequalities in Lemma 3.3.2 and taking the limit $t \to t_{m+1}^-$ gives

$$\begin{aligned}\|e_M^N\|_{L^2(X\times\mathbb{S}^2)}(t_M^-) \leq \|e_{M-1}^N\|_{L^2(X\times\mathbb{S}^2)}(t_{M-1}^-) + \|\eta_{\mathrm{c},M}^N\|_{L^2(X\times\mathbb{S}^2)}(t_M^-) \\ + \frac{1}{\varepsilon}\|\Omega\cdot\nabla_x\eta_{\mathrm{c},M}^N\|_{L^1([t_{M-1},t_M];L^2(X\times\mathbb{S}^2))}\end{aligned} \quad (3.95)$$

Exhausting this recursion until $e_0^N(t_0^-) = 0$ yields the result. □

Lemma 3.3.3. *Let $s \geq 0$, $N \geq \max\{0, s-1\}$, and $1 \leq m \leq M$. Then the m-th projection error $\eta_{\mathrm{c},m}^N$ satisfies,*

$$\|\eta_{\mathrm{c},m}^N\|_{L^2(X\times\mathbb{S}^2)}(t_m^-) \leq \frac{\Delta t}{\varepsilon(N+1)^s} \sup_{\tau\in[t_{m-1},t_m]} |\psi_{\mathrm{c},m}|_{H^{1,s}}(\tau). \quad (3.96)$$

Proof. An application of Lemma 3.2.10 with $\alpha = t_{m-1}$, $\beta = t_m$, $\eta_{\mathrm{c},m}^N(t_{m-1}) = 0$, an isotropic source $q = \frac{\sigma}{\varepsilon^2}\overline{\psi_{\mathrm{u},m}}$, along with the fact that $\mathcal{A}_\alpha[f](t) \leq \varepsilon^{-1}(t-\alpha)\sup_{\tau\in[\alpha,t]} f(\tau)$ (cf. (3.40) with $k=1$)) and the spectral estimate in Lemma 3.2.2, gives

$$\begin{aligned}\|\eta_{\mathrm{c},m}^N\|_{L^2(X\times\mathbb{S}^2)}(t_m^-) &\leq \frac{\Delta t}{\varepsilon} \sup_{\tau\in[t_{m-1},t_m]} \|\widetilde{\mathcal{P}}_N(\Omega\cdot\nabla_x\psi_{\mathrm{c},m}))\|_{L^2(X\times\mathbb{S}^2)}(\tau)\,d\tau \\ &\leq \frac{\Delta t}{\varepsilon(N+1)^s} \sup_{\tau\in[t_{m-1},t_m]} |\Omega\cdot\nabla_x\psi_{\mathrm{c},m}|_{H^{0,s}}(\tau)\,d\tau \leq \frac{\Delta t}{\varepsilon(N+1)^s} \sup_{\tau\in[t_{m-1},t_m]} |\psi_{\mathrm{c},m}|_{H^{1,s}}(\tau).\end{aligned} \quad (3.97)$$

□

3.3.3 Estimating hybrid error in terms of the data

Finally, we will apply the approximation properties and stability estimates to the estimate in Theorem 3.3.1 to obtain an estimate that depends only on the regularity of the data.

Theorem 3.3.2. *Let $s \geq 1$ and $N \geq s-1$. If $q \in L^1([0,T]; H^{s+1,0})$ and $g \in H^{s+1,0}$, then*

$$\|e_M^N\|_{L^2(X\times\mathbb{S}^2)}(T^-) \leq \frac{2C_s}{(N+1)^s}\left(|g|_{H^{s+1,0}} + T\,|q|_{L^\infty([0,T];H^{s+1,0})}\right) \tag{3.98}$$

$$\times \min\left(\frac{\varepsilon^{s-1}s!T}{\sigma^s}, \frac{\Delta t^s T}{\varepsilon^{s+1}}\min\left(1, \frac{\Delta t\sigma}{\varepsilon^2}\right)\right). \tag{3.99}$$

Proof. Using $\|\Omega\cdot\nabla_x\psi_{\mathrm{c},m}\|_{L^2(X\times\mathbb{S}^2)} \leq \|\nabla_x\psi_{\mathrm{c},m}\|_{L^2(X\times\mathbb{S}^2)}$, and applying Lemma 3.2.2 and Lemma 3.3.3 to Theorem 3.3.1 yields, for $N \geq s-1$,

$$\begin{aligned}\|e_M^N\|_{L^2(X\times\mathbb{S}^2)}(t_M^-) &\leq \frac{1}{(N+1)^s}\sum_{m=1}^M\left(\frac{\Delta t}{\varepsilon}\sup_{\tau\in[t_{m-1},t_m]}|\psi_{\mathrm{c},m}|_{H^{1,s}}(\tau) + \frac{1}{\varepsilon}|\psi_{\mathrm{c},m}|_{L^1([t_{m-1},t_m];H^{1,s})}\right)\\ &\leq \frac{2\Delta t}{\varepsilon(N+1)^s}\sum_{m=1}^M\sup_{\tau\in[t_{m-1},t_m]}|\psi_{\mathrm{c},m}|_{H^{1,s}}(\tau)\end{aligned}$$

Applying Corollary 3.2.1 with $q = \frac{\sigma}{\varepsilon^2}\overline{\psi_{\mathrm{u},m}}$ and $r = 1$,

$$\|e_M^N(t_M^-)\|_{L^2(X\times\mathbb{S}^2)} \leq 2C_s s!\frac{\sigma}{\varepsilon^3}\frac{\Delta t}{(N+1)^s}\sum_{m=1}^M\sup_{\tau\in[t_{m-1},t_m]}\mathcal{A}^s_{t_{m-1}}\left[|\psi_{\mathrm{u},m}|_{L^1([t_{m-1},\bullet];H^{s+1,0})}\right](\tau). \tag{3.100}$$

For the summand above, it follows from (3.84b), the monotonicity of $\mathcal{A}_\alpha$, and Lemma 3.2.3 that

$$\begin{aligned}\sup_{\tau\in[t_{m-1},t_m]}\mathcal{A}^s_{t_{m-1}}\left[|\psi_{\mathrm{u},m}|_{L^1([t_{m-1},\bullet];H^{s+1,0})}\right](\tau) &\leq \sup_{\tau\in[t_{m-1},t_m]}|\psi_{\mathrm{u},m}|_{L^1([t_{m-1},t_m];H^{s+1,0})}\mathcal{A}^s_{t_{m-1}}[\mathbb{1}](\tau)\\ &\leq \frac{\varepsilon^2}{\sigma}\left(|g|_{H^{s+1,0}} + |q|_{L^1([0,T];H^{s+1,0})}\right)\mathcal{A}^s_{t_{m-1}}[\mathbb{1}](t_m)\\ &\leq \frac{\varepsilon^2}{\sigma}\left(|g|_{H^{s+1,0}} + |q|_{L^1([0,T];H^{s+1,0})}\right)\min\left(\frac{\varepsilon^s}{\sigma^s}, \frac{1}{s!}\left(\frac{\Delta t}{\varepsilon}\right)^s\right).\end{aligned} \tag{3.101}$$

Pluggin the above bound into (3.100) yields (since $T = M\Delta t$)

$$\begin{aligned}\|e_M^N\|_{L^2(X\times\mathbb{S}^2)}(t_M^-) &\leq 2C_s s!\frac{\Delta t M}{\varepsilon(N+1)^s}\left(|g|_{H^{s+1,0}} + |q|_{L^1([0,T];H^{s+1,0})}\right)\min\left(\frac{\varepsilon^s}{\sigma^s}, \frac{\Delta t^s}{s!\varepsilon^s}\right)\\ &= \frac{2C_s}{(N+1)^s}\left(|g|_{H^{s+1,0}} + |q|_{L^1([0,T];H^{s+1,0})}\right)\min\left(\frac{\varepsilon^{s-1}s!T}{\sigma^s}, \frac{\Delta t^s T}{\varepsilon^{s+1}}\right).\end{aligned} \tag{3.102}$$

On the other hand, it follows from (3.84c) that

$$\begin{aligned}
&\sup_{\tau\in[t_{m-1},t_m]} \mathcal{A}^s_{t_{m-1}}\left[|\psi_{\mathrm{u},m}|_{L^1([t_{m-1},\bullet];H^{s+1,0})}\right](\tau)\\
&\qquad\le \varepsilon(|g|_{H^{s+1,0}} + T\,|q|_{L^\infty([0,T];H^{s+1,0})}) \sup_{\tau\in[t_{m-1},t_m]} \mathcal{A}^{s+1}_{t_{m-1}}[\mathbb{1}](\tau) \qquad (3.103)\\
&\qquad\le \varepsilon(|g|_{H^{s+1,0}} + T\,|q|_{L^\infty([0,T];H^{s+1,0})}) \min\left(\frac{\varepsilon^{s+1}}{\sigma^{s+1}}, \frac{1}{(s+1)!}\left(\frac{\Delta t}{\varepsilon}\right)^{s+1}\right).
\end{aligned}$$

Plugging this into (3.100) yields,

$$\|e^N_M\|_{L^2(X\times\mathbb{S}^2)}(t^-_M) \le \frac{2C_s}{(N+1)^s}\left(|g|_{H^{s+1,0}} + T\,|q|_{L^\infty([0,T);H^{s+1,0})}\right)\min\left(\frac{\varepsilon^{s-1}s!T}{\sigma^s}, \frac{\Delta t^{s+1}\sigma T}{\varepsilon^{s+3}}\right). \qquad (3.104)$$

Taking a minimum of the right hand sides of (3.102) and (3.104) yields the result. □

3.4 Return to the original transport model

In this section we will show error estimates for the model (3.1). The analogous discretizations for the models are the following. For the non-splitting P_N discretization we seek $\Psi^{\varepsilon,N} \in C([0,T); X\times\mathbb{P}_N(\mathbb{S}^2))$, satisfying

$$\varepsilon\partial_t\Psi^{\varepsilon,N} + \mathcal{P}_N(\Omega\cdot\nabla_x\Psi^{\varepsilon,N}) + \frac{\sigma_\mathrm{t}}{\varepsilon}\Psi^{\varepsilon,N} = \left(\frac{\sigma_\mathrm{t}}{\varepsilon} - \varepsilon\sigma_\mathrm{a}\right)\overline{\Psi^{\varepsilon,N}} + \varepsilon\mathcal{P}_N Q, \qquad (3.105a)$$

$$\Psi^{\varepsilon,N}|_{t=0} = \mathcal{P}_N g \qquad (3.105b)$$

and for the hybrid, we seek $\Psi^{\varepsilon,N}_m = \Psi^{\varepsilon,N}_{\mathrm{u},m} + \Psi^{\varepsilon,N}_{\mathrm{c},m}$ where for each $m\in\{1,2,\dots,M\}$

$$\left(\Psi^{\varepsilon,N}_{\mathrm{u},m}, \Psi^{\varepsilon,N}_{\mathrm{c},m}\right) \in C([t_{m-1},t_m); X\times L^2(\mathbb{S}^2)) \times C([t_{m-1},t_m); X\times\mathbb{P}_N(\mathbb{S}^2)) \qquad (3.106)$$

satifies

$$\varepsilon\partial_t\Psi^{\varepsilon,N}_{\mathrm{u},m} + \Omega\cdot\nabla_x\Psi^{\varepsilon,N}_{\mathrm{u},m} + \frac{\sigma_\mathrm{t}}{\varepsilon}\Psi^{\varepsilon,N}_{\mathrm{u},m} = \varepsilon Q, \qquad (3.107a)$$

$$\varepsilon\partial_t\Psi^{\varepsilon,N}_{\mathrm{c},m} + \mathcal{P}_N\left(\Omega\cdot\nabla_x\Psi^{\varepsilon,N}_{\mathrm{c},m}\right) + \frac{\sigma_\mathrm{t}}{\varepsilon}\Psi^{\varepsilon,N}_{\mathrm{c},m} = \left(\frac{\sigma_\mathrm{t}}{\varepsilon} - \varepsilon\sigma_\mathrm{a}\right)\overline{\Psi^{\varepsilon,N}_{\mathrm{u},m}} + \overline{\Psi^{\varepsilon,N}_{\mathrm{c},m}}, \qquad (3.107b)$$

$$(3.107c)$$

$$\Psi^{\varepsilon,N}_{\mathrm{c},m}\big|_{t=t_{m-1}} = 0, \quad \Psi^{\varepsilon,N}_{\mathrm{u},m}\big|_{t=t_{m-1}} = \begin{cases} g, & m=1,\\ \Psi^{\varepsilon,N}_{\mathrm{u},m-1}(t^-_{m-1}) + \Psi^{\varepsilon,N}_{\mathrm{c},m-1}(t^-_{m-1}) & m>1. \end{cases} \qquad (3.107d)$$

We define the correspondent P_N error and m-th hybrid error respectively as,

$$e^{\varepsilon,N} = \Psi^{\varepsilon} - \Psi^{\varepsilon,N} \quad \text{and} \quad e_m^{\varepsilon,N} = \Psi_m^{\varepsilon} - \Psi_m^{\varepsilon,N}$$

Since $\psi^N = e^{\sigma_{\mathrm{a}} t}\Psi^{\varepsilon,N}$ and $\psi_m^N = e^{\sigma_{\mathrm{a}} t}\Psi_m^{\varepsilon,N}$, applying Theorems 3.2.1 and 3.3.2 gives the following estimate for the P_N error for the original transport model (3.1).

Theorem 3.4.1. *Let $s \geq 1$ and $N \geq s-1$, $Q \in L^\infty([0,T]; H^{i,j})$ and $g \in H^{i,j}$ for each i, j such that $0 \leq j \leq s$, $i + j \leq s+1$. Then*

$$\begin{aligned}
\|e^{\varepsilon,N}\|_{L^2(X\times\mathbb{S}^2)}(T) \leq{}& \frac{e^{-(\sigma_{\mathrm{t}}+\varepsilon^2\sigma_{\mathrm{a}})T/\varepsilon^2}}{(N+1)^s}|g|_{H^{0,s}} + \frac{1}{(N+1)^s}|Q|_{L^\infty([0,T];H^{0,s})}\min\left(\frac{\varepsilon^2}{\sigma_{\mathrm{t}}}, T\right) \\
&+2C_s\frac{1}{(N+1)^s}\Bigg(\left[e^{-\sigma_{\mathrm{a}}T}|g|_{H^{s+1,0}} + T|Q|_{L^\infty([0,T];H^{s+1,0})}\right]\min\left(\frac{\varepsilon^{s-1}s!T}{\sigma_{\mathrm{t}}^s}, \left(\frac{T}{\varepsilon}\right)^{s+1}\right) \\
&\qquad + e^{-(\sigma_{\mathrm{t}}+\varepsilon^2\sigma_{\mathrm{a}})T}\sum_{i=0}^{s-1}|g|_{H^{1+i,s-i}}\binom{s}{i}\frac{T^{i+1}}{\varepsilon^{i+1}} \\
&\qquad + \sum_{i=0}^{s-1}|Q|_{L^\infty([0,T];H^{1+i,s-i})}\binom{s}{i}\min\left(\frac{\varepsilon^{i+1}T}{\sigma_{\mathrm{t}}^{i+1}}, \frac{1}{(i+1)!}\frac{T^{i+2}}{\varepsilon^{i+1}}\right)\Bigg)
\end{aligned} \tag{3.108}$$

Meanwhile for the hybrid approximation.

Theorem 3.4.2. *Let $s \geq 1$ and $N \geq s-1$, $Q \in L^1([0,T]; H^{s+1,0})$ and $g \in H^{s+1,0}$, we have*

$$\begin{aligned}
\|e_M^{\varepsilon,N}\|_{L^2(X\times\mathbb{S}^2)}(T^-) \leq{}& \frac{2C_s}{(N+1)^s}\left(e^{-\sigma_{\mathrm{a}}T}\,|g|_{H^{s+1,0}} + T\,|Q|_{L^\infty([0,T];H^{s+1,0})}\right) \\
&\times \min\left(\frac{\varepsilon^{s-1}s!T}{\sigma_{\mathrm{t}}^s}, \frac{\Delta t^s T}{\varepsilon^{s+1}}\min\left(1, \frac{\Delta t\sigma_{\mathrm{t}}}{\varepsilon^2}\right)\right).
\end{aligned} \tag{3.109}$$

CHAPTER 4

SUMMARY AND CONCLUSION

In this thesis, we studied two numerical schemes for kinetic equations. In Chapter 2, we proved theoretically and demonstrated computationally the effectiveness of the SIAC filter to the DG solutions of the nonlinear VM system. We proved the superconvergence of order $(2k + \frac{1}{2})$ in the negative norm of the DG solutions. This is nontrivial for nonlinear systems, and is achieved by identifying a suitable dual problem. The numerical experiments verify the performance of the filter in reducing spurious oscillations in the numerical errors. For low order k, the resolution of the numerical solution is greatly enhanced, which is highly desirable for long time kinetic simulations. In the future, we plan to prove superconvergence for the divided difference of the numerical solution to fully justify the enhanced resolution of the post-processed solution. Another interesting project will be to apply this post-processing technique to different kinetic equations that use the DG method.

In Chapter 3, we derived multiscale error estimates for the P_N approximation of the RTE and for a hybrid approximation for the RTE that is built using the P_N approximation. By construction, the hybrid is is more expensive; we use these error estimates to understand the benefits of the additional expense for different parameter regimes. At each time step in the hybrid approximation, the collided equation is equipped with isotropic initial conditions and zero initial condition. In scattering dominating regimes, this property is key to improved estimates over the monolithic P_N approach. Meanwhile, in purely absorbing regimes, the hybrid captures the RTE solution exactly. In the future, we intend to revisit the current analysis for more general problems on non-periodic domains, with non-constant cross-sections and inflow boundary conditions. In addition, we intend to explicitly examine the effects of angular discretization errors in the treatment of the uncollided equation, which for the purposes of the current chapter was assumed to be solved exactly.

BIBLIOGRAPHY

[1] Andrés Galindo-Olarte, Juntao Huang, Jennifer K Ryan, and Yingda Cheng. Superconvergence and accuracy enhancement of discontinuous galerkin solutions for vlasov-maxwell equations. *arXiv preprint arXiv:2210.07908*, 2022.

[2] Cory D Hauck and Ryan G McClarren. A collision-based hybrid method for time-dependent, linear, kinetic transport equations. *Multiscale Modeling & Simulation*, 11(4):1197–1227, 2013.

[3] Andrés Galindo-Olarte, Victor P DeCaria, and Cory D Hauck. Numerical analysis of a hybrid method for radiation transport. *arXiv:2306.04714*, 2023.

[4] Lorenzo Pareschi. Kinetic equations: computation. *arXiv:1311.7230v1*, 2013.

[5] Giacomo Dimarco and Lorenzo Pareschi. Numerical methods for kinetic equations. *Acta Numerica*, 23:369–520, 2014.

[6] F. Califano, F. Pegoraro, S. V. Bulanov, and A. Mangeney. Kinetic saturation of the Weibel instability in a collisionless plasma. *Phys. Rev. E*, 57(6):7048–7059, 1998.

[7] A. Mangeney, F. Califano, C. Cavazzoni, and P. Travnicek. A numerical scheme for the integration of the Vlasov-Maxwell system of equations. *J. Comp. Phys.*, 179(2):495–538, 2002.

[8] F. Califano, F. Pegoraro, and S. V. Bulanov. Impact of kinetic processes on the macroscopic nonlinear evolution of the electromagnetic-beam-plasma instability. *Phys. Rev. Lett.*, 84:3602–3605, 2000.

[9] F. Califano, N. Attico, F. Pegoraro, G. Bertin, and S. V. Bulanov. Fast formation of magnetic islands in a plasma in the presence of counterstreaming electrons. *Phys. Rev. Lett.*, 86(23):5293–5296, 2001.

[10] N.J. Sircombe and T.D. Arber. VALIS: A split-conservative scheme for the relativistic 2d Vlasov-Maxwell system. *J. Comp. Phys.*, 228(13):4773 – 4788, 2009.

[11] Nicolas Besse, Guillaume Latu, Alain Ghizzo, Eric Sonnendrüker, and Pierre Bertrand. A wavelet-MRA-based adaptive semi-Lagrangian method for the relativistic Vlasov-Maxwell system. *J. Comp. Phys.*, 227(16):7889 – 7916, 2008.

[12] Akihiro Suzuki and Toshikazu Shigeyama. A conservative scheme for the relativistic Vlasov-Maxwell system. *J. Comp. Phys.*, 229(5):1643 – 1660, 2010.

[13] F. Huot, A. Ghizzo, P. Bertrand, E. Sonnendrüker, and O. Coulaud. Instability of the time splitting scheme for the one-dimensional and relativistic Vlasov-Maxwell system. *J. Comp. Phys.*, 185(2):512 – 531, 2003.

[14] Bernardo Cockburn and Chi-Wang Shu. Runge-Kutta discontinuous Galerkin methods for convection-dominated problems. *J. Sci. Comput.*, 16:173–261, 2001.

[15] R. E. Heath, I. M. Gamba, P. J. Morrison, and C. Michler. A discontinuous Galerkin method for the Vlasov-Poisson system. *J. Comp. Phys.*, 231:1140–1174, 2012.

[16] R. E. Heath. Numerical analysis of the discontinuous Galerkin method applied to plasma physics. 2007. Ph. D. dissertation, the University of Texas at Austin.

[17] Y. Cheng, I. M. Gamba, and P. J. Morrison. Study of conservation and recurrence of Runge-Kutta discontinuous Galerkin schemes for Vlasov-Poisson systems. *J. Sci. Comp. accepted*, 2012. preprint arXiv:1209.6413v2 [math.NA].

[18] Y. Cheng and I. M. Gamba. Numerical study of Vlasov-Poisson equations for infinite homogeneous stellar systems. *Comm. Nonlin. Sci. Num. Sim.*, 17, 2012.

[19] Yingda Cheng, Irene M Gamba, Fengyan Li, and Philip J Morrison. Discontinuous Galerkin methods for the Vlasov-Maxwell equations. *SIAM Journal on Numerical Analysis*, 52(2):1017–1049, 2014.

[20] Yingda Cheng, Andrew J Christlieb, and Xinghui Zhong. Energy-conserving discontinuous Galerkin methods for the Vlasov–Ampere system. *Journal of Computational Physics*, 256:630–655, 2014.

[21] He Yang and Fengyan Li. Discontinuous galerkin methods for relativistic Vlasov–Maxwell system. *Journal of Scientific Computing*, 73(2):1216–1248, 2017.

[22] James H Bramble and Alfred H Schatz. Higher order local accuracy by averaging in the finite element method. *Mathematics of Computation*, 31(137):94–111, 1977.

[23] Bernardo Cockburn, Mitchell Luskin, Chi-Wang Shu, and Endre Süli. Enhanced accuracy by post-processing for finite element methods for hyperbolic equations. *Mathematics of Computation*, 72(242):577–606, 2003.

[24] Liangyue Ji, Yan Xu, and Jennifer K Ryan. Negative-order norm estimates for nonlinear hyperbolic conservation laws. *Journal of Scientific Computing*, 54(2):531–548, 2013.

[25] Xiong Meng and Jennifer K Ryan. Discontinuous Galerkin methods for nonlinear scalar hyperbolic conservation laws: divided difference estimates and accuracy enhancement. *Numerische mathematik*, 136(1):27–73, 2017.

[26] Xiong Meng and Jennifer K Ryan. Divided difference estimates and accuracy enhancement of discontinuous Galerkin methods for nonlinear symmetric systems of hyperbolic conservation laws. *IMA Journal of Numerical Analysis*, 38(1):125–155, 2018.

[27] Michael Steffan, Sean Curtis, Robert M Kirby, and Jennifer Ryan. Investigation of smoothness enhancing accuracy-conserving filters for improving streamline integration through discontinuous fields. *IEEE Transactions on Visualization and Computer Graphics*, 14(3):680–692, 2008.

[28] Liangyue Ji, Paulien Van Slingerland, Jennifer K Ryan, and Kees Vuik. Superconvergent error estimates for position-dependent smoothness-increasing accuracy-conserving (SIAC) post-processing of discontinuous Galerkin solutions. *Mathematics of computation*, pages 2239–2262, 2014.

[29] Gerald C Pomraning. *The equations of radiation hydrodynamics*. Courier Corporation, 2005.

[30] Elmer Eugene Lewis and Warren F Miller. *Computational methods of neutron transport*. John Wiley and Sons, Inc., New York, NY, 1984.

[31] K.M. Case and P.F. Zweifel. *Linear Transport Theory*. Addison-Wesley series in nuclear engineering. Addison-Wesley Publishing Company, 1967.

[32] Edward W Larsen and Joseph B Keller. Asymptotic solution of neutron transport problems for small mean free paths. *Journal of Mathematical Physics*, 15(1):75–81, 1974.

[33] Alain Bensoussan, Jacques L Lions, and George C Papanicolaou. Boundary layers and homogenization of transport processes. *Publications of the Research Institute for Mathematical Sciences*, 15(1):53–157, 1979.

[34] Mohammed Lemou and Luc Mieussens. A new asymptotic preserving scheme based on micro-macro formulation for linear kinetic equations in the diffusion limit. *SIAM Journal on Scientific Computing*, 31(1):334–368, 2008.

[35] Luis Chacon, Guangye Chen, Dana A Knoll, C Newman, H Park, William Taitano, Jeff A Willert, and Geoffrey Womeldorff. Multiscale high-order/low-order (holo) algorithms and applications. *Journal of Computational Physics*, 330:21–45, 2017.

[36] Marvin L Adams and Edward W Larsen. Fast iterative methods for discrete-ordinates particle transport calculations. *Progress in nuclear energy*, 40(1):3–159, 2002.

[37] James S Warsa, Todd A Wareing, and Jim E Morel. Krylov iterative methods and the degraded effectiveness of diffusion synthetic acceleration for multidimensional sn calculations in problems with material discontinuities. *Nuclear science and engineering*, 147(3):218–248, 2004.

[38] Raymond E Alcouffe. A first collision source method for coupling monte carlo and discrete ordinates for localized source problems. In *Monte-Carlo Methods and Applications in Neutronics, Photonics and Statistical Physics: Proceedings of the Joint Los Alamos National Laboratory-Commissariat à l'Energie Atomique Meeting Held at Cadarache Castle, Provence, France April 22–26, 1985*, pages 352–366. Springer, 2006.

[39] Michael M Crockatt, Andrew J Christlieb, C Kristopher Garrett, and Cory D Hauck. An arbitrary-order, fully implicit, hybrid kinetic solver for linear radiative transport using integral deferred correction. *Journal of Computational Physics*, 346:212–241, 2017.

[40] Michael M Crockatt, Andrew J Christlieb, C Kristopher Garrett, and Cory D Hauck. Hybrid methods for radiation transport using diagonally implicit runge–kutta and space–time discontinuous galerkin time integration. *Journal of Computational Physics*, 376:455–477, 2019.

[41] Michael M Crockatt, Andrew J Christlieb, and Cory D Hauck. Improvements to a class of hybrid methods for radiation transport: Nyström reconstruction and defect correction methods. *Journal of Computational Physics*, 422:109765, 2020.

[42] Ben Whewell, Ryan G McClarren, Cory D Hauck, and Minwoo Shin. Multigroup Neutron Transport Using a Collision-Based Hybrid Method. *Nuclear science and engineering*, 2023.

[43] Vincent Heningburg and Cory D Hauck. Hybrid solver for the radiative transport equation using finite volume and discontinuous galerkin. *arXiv preprint arXiv:2002.02517*, 2020.

[44] Martin Frank, Cory Hauck, and Kerstin Küpper. Convergence of filtered spherical harmonic equations for radiation transport. *Communications in Mathematical Sciences*, 14(5):1443–1465, 2016.

[45] Zheng Chen and Cory Hauck. Multiscale convergence properties for spectral approximations of a model kinetic equation. *Mathematics of Computation*, 88(319):2257–22 93, 2019.

[46] Sean Curtis, Robert M Kirby, Jennifer K Ryan, and Chi-Wang Shu. Postprocessing for the discontinuous Galerkin method over nonuniform meshes. *SIAM Journal on Scientific Computing*, 30(1):272–289, 2008.

[47] Philippe G Ciarlet. The finite element method for elliptic problems. *Bull. Amer. Math. Soc*, 1:800–802, 1979.

[48] Guri Ivanovich Marchuk. Construction of adjoint operators in non-linear problems of mathematical physics. *Sbornik: Mathematics*, 189(10):1505, 1998.

[49] Michel Crouzeix and Vidar Thomée. The stability in L_p and W_p^1 of the L_2-projection onto finite element function spaces. *Mathematics of Computation*, 48(178):521–532, 1987.

[50] Blanca Ayuso de Dios, José Antonio Carrillo de la Plata, and C-W Shu. Discontinuous Galerkin methods for the one-dimensional Vlasov-Poisson system. 2009.

[51] Susanne C Brenner, L Ridgway Scott, and L Ridgway Scott. *The mathematical theory of finite element methods*, volume 3. Springer, 2008.

[52] Sigal Gottlieb and Chi-Wang Shu. Total variation diminishing Runge-Kutta schemes. *Mathematics of computation*, 67(221):73–85, 1998.

[53] Robert Dautray and Jacques-Louis Lions. *Mathematical analysis and numerical methods for science and technology: volume 6 evolution problems II*, volume 6. Springer Science & Business Media, 1999.

[54] Cory D Hauck and Robert B Lowrie. Temporal regularization of the P_N equations. *Multiscale Modeling & Simulation*, 7(4):1497–1524, 2009.

[55] Kendall Atkinson and Weimin Han. *Spherical harmonics and approximations on the unit sphere: an introduction*, volume 2044. Springer Science & Business Media, 2012.

[56] Feng Dai and Yuan Xu. *Approximation Theory and Harmonic Analysis on Spheres and Balls*. Springer, 01 2013.

APPENDIX

Proof of Lemma 2.3.3

By using equation (2.21a), the divergence free properties of $\mathbf{A_1}, \mathbf{A_2}$ and the boundary conditions, we have the following

$$\frac{1}{2}\frac{d}{dt}\|\varphi\|^2 = -\int_\Omega (\mathbf{A_3}\cdot\mathbf{F})\varphi\, d\mathbf{x}d\mathbf{v} \leq C(\|\varphi\|^2 + \|\mathbf{F}\|^2),$$

where C depends on $\|\mathbf{A_3}\|_{L^\infty((0,T);L^\infty(\Omega))}$. On the other hand using equations (2.21b) and (2.21c), Gauss theorem on the physical space integrals and integration by parts on the velocity space variables,

$$\begin{aligned}\frac{1}{2}\frac{d}{dt}\|\mathbf{F}\|^2 + \frac{1}{2}\frac{d}{dt}\|\mathbf{D}\|^2 &= \int_{\Omega_x}(\nabla_\mathbf{x}\times\mathbf{D}\cdot\mathbf{F} - \nabla_\mathbf{x}\times\mathbf{F}\cdot\mathbf{D})\, d\mathbf{x} - \int_{\Omega_v}\varphi\nabla_\mathbf{v} g\cdot\mathbf{F}\, d\mathbf{x}d\mathbf{v}\\ &\quad + \int_{\Omega_v}\varphi(\mathbf{v}\times\nabla_\mathbf{v} g)\mathbf{D}\, d\mathbf{x}d\mathbf{v}\\ &= -\int_\Omega \varphi\nabla_\mathbf{v} g\cdot\mathbf{F}\, d\mathbf{x}d\mathbf{v} + \int_\Omega \varphi(\mathbf{v}\times\nabla_\mathbf{v} g)\mathbf{D}\, d\mathbf{x}d\mathbf{v}\\ &\leq C\left(\|\mathbf{F}\|^2 + \|\mathbf{D}\|^2 + \|\varphi\|^2\right),\end{aligned}$$

where C depends on $\|g\|_{L^\infty((0,T);W^{1,\infty}(\Omega))}$.

Now we add the tow inequalities above, to obtain

$$\frac{1}{2}\frac{d}{dt}\|\varphi\|^2 + \frac{1}{2}\frac{d}{dt}\|\mathbf{F}\|^2 + \frac{1}{2}\frac{d}{dt}\|\mathbf{D}\|^2 \leq C\left(\|\mathbf{F}\|^2 + \|\mathbf{D}\|^2 + \|\varphi\|^2\right), \tag{.1}$$

where C depends on $\|\mathbf{A_3}\|_{L^\infty((0,T);L^\infty(\Omega))}$ and $\|g\|_{L^\infty((0,T);W^{1,\infty}(\Omega))}$. An application of Gronwall's inequality allow us to conclude. Now since we are considering the full Sobolev norm, we still need to estimate the L^2 norms of the higher order derivatives $\partial^\beta_\mathbf{x}\partial^\gamma_\mathbf{v}$, to do so we apply $\partial^\beta_\mathbf{x}\partial^\gamma_\mathbf{v}$ to the system (2.21) and then we repeat the same steps that we took above.

Fundamental ODE result

If $|\cdot|$ denotes either one of the semi-norms and norms defined throughout the chapter, one of the usual procedures is finding a bound for $|\varphi|$, by analyzing a differential

inequality, that looks like

$$\frac{1}{2}\frac{d}{dt}|\varphi|^2 + \kappa|\varphi|^2 \leq \chi(t)|\varphi|,$$

where $\kappa \geq 0$ is a constant and $\chi(t) \geq 0$ for all times t. Formally one would divide by $|\varphi|$ and integrate in time. However this computation is not rigorous if there exists some time t^*, for which $|\varphi|(t^*) = 0$. The following lemma makes the formal calculation rigorous.

Lemma .0.1. *Assume χ is a non-negative continuous function on $[\alpha, \beta]$. Assume $\phi \in C^1([\alpha, \beta])$, $\phi \geq 0$, and satisfies the following differential inequality*

$$\frac{1}{2}(\phi(t)^2)' + \kappa\phi(t)^2 \leq \chi(t)\phi(t), \quad \phi(\alpha) = \phi_\alpha \geq 0. \tag{.2}$$

Then for all $t \in [\alpha, \beta]$,

$$\phi(t) + \kappa\int_\alpha^t \phi(\tau)\,d\tau \leq \phi_\alpha + \int_\alpha^t \chi(\tau)\,d\tau. \tag{.3}$$

Furthermore

$$\phi(t) \leq e^{-\kappa(t-\alpha)}\phi_\alpha + \int_\alpha^t e^{-\kappa(t-\tau)}\chi(\tau)\,d\tau. \tag{.4}$$

Proof. We prove first (.3). Since ϕ and χ are non-negative functions, it follows that for any arbitrary $\delta > 0$, the following differential inequality holds

$$\phi(t)\phi'(t) + \kappa\phi(t)^2 \leq \chi(t)(\phi(t) + \delta), \tag{.5}$$

dividing both sides of the inequality by $\phi + \delta$, and integrating in time, we arrive at

$$\begin{aligned}\phi(t) + \kappa\int_\alpha^t \phi(\tau)\,d\tau &\leq \phi_\alpha + \int_\alpha^t \chi(\tau)\,d\tau + \delta\ln\left|\frac{\phi(t)+\delta}{\phi_\alpha+\delta}\right| + \kappa\delta\int_\alpha^t \frac{\phi(\tau)}{\phi(\tau)+\delta}\,d\tau,\\ &\leq \phi_\alpha + \int_\alpha^t \chi(\tau)\,d\tau + \delta\ln\left|\frac{\phi(t)+\delta}{\phi_\alpha+\delta}\right| + \kappa\delta(t-\alpha),\end{aligned}$$

the conclusion follows taking $\delta \to 0^+$.

We next prove (.4). When $\kappa = 0$ the result follows immediately from (.3):

$$\phi(t) \leq \phi_\alpha + \int_\alpha^t \chi(\tau)\,d\tau. \tag{.6}$$

For the general case we multiply (.2) by $e^{2\kappa t}$, obtaining

$$\frac{1}{2}\left[\Phi(t)^2\right]' \leq e^{\kappa t}\chi(t)\Phi(t), \tag{.7}$$

where $\Phi(t) = e^{\kappa t}\phi(t)$. Applying (.6) to Φ and undoing the transformation yields (.4). $\square$