8-2019

Numerical Methods for Radiative Transport Equations

Vincent Heningburg
University of Tennessee

To the Graduate Council:

I am submitting herewith a dissertation written by Vincent Heningburg entitled "Numerical Methods for Radiative Transport Equations." I have examined the final electronic copy of this dissertation for form and content and recommend that it be accepted in partial fulfillment of the requirements for the degree of Doctor of Philosophy, with a major in Mathematics.

Cory Hauck, Major Professor

We have read this dissertation and recommend its acceptance:

Xiaobing Feng, Ohannes Karakashian, Anthony Mezzacappa

Accepted for the Council:

Dixie L. Thompson

Vice Provost and Dean of the Graduate School

(Original signatures are on file with official student records.)

Numerical Methods for Radiative Transport Equations

A Dissertation Presented for the

Doctor of Philosophy

Degree

The University of Tennessee, Knoxville

Vincent Edmund Heningburg

August 2019

To everyone I consider family, you know who you are.

Acknowledgments

I would like to express my sincere gratitude to my advisor Dr. Cory Hauck. Your guidance and patience has been immeasurable, and I am quite confident that I could not have done this without you.

I would also like to thank my committee, Profs. Feng, Karakashian, and Mezzacappa. You all have helped me considerably in my education, and I cherish the conversations I've had with each of you throughout these many years.

I would like to thank Dr. Yulong Xing, my former co-advisor, for helping to start my interest in numerical methods for PDE's.

I would also like to thank all of my teachers at the University of Tennessee for their dedication to their craft, and I thank the Bredesen Center for giving me the chance to study at this great university.

I thank Kristopher Garrett for his mentorship at Los Almos National Laboratory, which became the groundwork for Chapter 4. I would also like to thank him for the calculation of an approximation to the analytic solution I use in my experiments in the line source benchmark.

I would like to thank Oak Ridge National Laboratory and its staff for providing the necessary resources I used in the process of writing this dissertation.

I thank Richard Barnard, Zheng Chen, Micheal Crockatt, Joseph Daws, Anton Dereventsov, Nicholas Dexter, Eirik Endeve, Zachary Grant, Armenak Petrosyan, Qiwei Sheng, Jeremy Trageser, Hoang Tran, and all my other colleagues at ORNL for their support and advice. I would like to give special thanks to Miroslav Stoyanov for all his assistance with the main supercomputer I used in my research, and Ming Tse Paul Laiu for the considerable help with my countless questions and for the reviews of my papers.

I would like to thank my undergraduate advisor Prof. Luis Cueva-Parra, and my former classmates Mark Allen Guest II and Pratik Patel from Auburn University at Montgomery for introducing me to the joy of Mathematics and challenging me to be better.

I would like to thank all my friends at the University of Tennessee for their support throughout the years. I wish to give special thanks to my adopted family members Kacy and Stefan Schnake, and Mustafa Elmas for all the love and support you have shown me thus far. This journey would have been much less enjoyable without you.

Finally, I would like to thank my sisters Alyssa, Avory, Octavia, and Katherine, my brothers Chris and Raunte, and my parents Joanne, Darryl, and Delphia for everything you have done for me throughout my life.

Abstract

In this dissertation, we present and analyze a discrete ordinates (S_N) discretization of a filtered radiative transport equation (RTE). Under certain conditions, S_N discretizations of the standard RTE create numeric artifacts, known as "ray-effects"; the goal of using a filter is to remove such artifacts. We analyze convergence of the filtered discrete ordinates solution to the solution of the non-filtered RTE, taking into account the effect of the filter as well as the usual quadrature and truncation errors that arise in discrete ordinates methods.

We also present a hybrid spatial discretization for the radiative transport equation that combines a second-order discontinuous Galerkin (DG) method and a second-order finite volume (FV) method. The strategy relies on a simple operator splitting that has been used previously to combine different angular discretizations. Unlike standard FV methods with upwind fluxes, the hybrid approach is able to accurately simulate problems in scattering dominated regimes. However, it requires less memory and yields a faster time to solution than a standard DG approach. In addition, the underlying splitting allows naturally for hybridization in both space and angle.

We demonstrate, via the simulation of two benchmark problems, the effectiveness of the filtering approach in reducing ray effects. In addition, we also examine efficiency of both methods, in particular the balance between improved accuracy and additional cost of including the filter, and the ability of the spatial hybrid to leverage its efficiency to produce more accurate results.

Table of Contents

List of Tables

List of Figures

Nomenclature

$(\cdot)_N$	Quadrature rule of order N
$\mathbf{S}^N$	External source vector w.r.t quadrature of order N
$\boldsymbol{\psi}_0^N$	Initial condition vector w.r.t quadrature of order N
$\boldsymbol{\psi}^\Theta$	Solution satisfying the filtered discrete ordinates equations
$\mathcal{A}$	Averaging operator over $\mathbb{S}^2$
$(\cdot,\cdot)_N$	Bilinear operator for continuous functions w.r.t quadrature of order N
$\mathcal{E}^N$	Discrete evaluation operator for continuous functions w.r.t quadrature of order N
$\mathcal{F}^D$	Filtering operator of degree D
$\mathcal{I}^{N,M}$	Discrete expansion operator of degree M w.r.t bilinear operator $(\cdot,\cdot)_N$
$\mathcal{L}$	Streaming operator
$\mathcal{L}_\epsilon$	Scaled streaming operator
$\lvert\cdot\rvert_N$	Semi-norm for continuous functions associated with $(\cdot,\cdot)_N$
$\mathcal{P}^M$	Orthogonal projection from $L^2(\mathbb{S}^2)$ to $\mathbb{P}^M(\mathbb{S}^2)$ w.r.t the standard inner product
$\mathcal{Q}$	Scattering source operator
$\mathcal{Q}^D$	Scattering source operator truncated to degree D
$\mathcal{W}_{h,0}^{\mathrm{u}}$	Product space $(\mathcal{X}_{h,0})^{N_{\mathrm{u}}^*}$

$\mathcal{W}_{h,1}^{\mathrm{c}}$ Product space $(\mathcal{X}_{h,1})^{N_{\mathrm{c}}^*}$

$\mathcal{X}_{h,i}$ Finite dimensional subspace of $L^2(X)$ of piece-wise defined polynomials, $i \in \{0,1\}$

Ω_i^N i-th discrete angle in $\mathbb{S}^2$ for quadrature of order N

$\langle \cdot, \cdot \rangle_N$ Discrete inner product w.r.t quadrature of order N

w_i^N i-th weight for quadrature of order N

ϵ Scaling parameter

Γ Phase space $X \times \mathbb{S}^2$

Υ^N Gradient matrix w.r.t quadrature of order N

$\hat{u}_\ell$ ℓ-th moment for $u \in L^2[-1,1]$

$\hat{v}_{\ell,k}^N$ ℓ,k-th discrete moment of vector $\mathbf{v}$

$\hat{v}_{\ell,k}$ ℓ,k-th moment for $v \in L^2(\mathbb{S}^2)$

$(\cdot,\cdot)$ Inner product on $L^2(\mathbb{S}^2)$

$\|\cdot\|_{L_N^2(X)}$ Discrete L^2 norm with respect to discrete inner product $\langle \cdot, \cdot \rangle_N$

∇_r Gradient in space

N_{c}^* Number of points and weights in quadrature of order N_{c}

N_{u}^* Number of points and weights in quadrature of order N_{u}

$\mathrm{FS}_{N,P}^D$ Solution to the filtered discrete ordinates of order N with anisotropy degree D and filter degree P

$\mathrm{FS}_{N,P}$ Isotropic solution to the filtered discrete ordinates of order N with filter degree P

P_N Spherical harmonics method of order N

S_N Discrete ordinates method of order N

$\mathrm{S}_{N_\mathrm{u}}\mathrm{S}_{N_\mathrm{c}}$ Solution method of spatial hybrid using quadrature orders of N_u and N_u for the uncollided and collided equations respectively

Ω Angle in unit sphere $\mathbb{S}^2$.

$\partial\Gamma$ Boundary of phase space

$\partial\Gamma^-$ Inflow boundary of phase space

$\partial\Gamma^-$ Outflow boundary of phase space

∂X Boundary of X

$\partial X_i^\pm$ Outflow/inflow boundary w.r.t Ω_i^N

ψ_i^Θ i-th component of $\boldsymbol{\psi}^\Theta$

$\mathcal{Q}_{\mathrm{s},\epsilon}$ Scaled scattering operator

$\mathcal{Q}_{\mathrm{s},\epsilon}$ Scaled scattering source operator

$Q_{\mathrm{s},\epsilon}^{\mathrm{m,n}}$ Discrete scattering source operator from finite dimensional space of order N_n to finite dimensional space of order N_n, with $\mathrm{m},\mathrm{n} \in \{\mathrm{u},\mathrm{c}\}$

σ_a Absorption cross section

σ_f Constant filter strength

σ_s Scattering cross section

$\sigma_{\mathrm{s},\epsilon}$ Scaled scattering cross-section

σ_t Total cross section

ψ Solution satisfying the radiative transport equation

ψ_0 Initial condition on Γ

ψ_b Boundary condition on $(0,T)\times\partial\Gamma^-$

$\mathbb{S}^2$ Unit sphere in $\mathbb{R}^3$

Θ	Multi-index for indices D, P, N
$\tilde{v}_{\ell,k}^{N}$	Discrete moment for continuous function v w.r.t bilinear operator $(\cdot,\cdot)_N$
A^{n}	Discrete averaging operator w.r.t quadrature of order N_{n}, $\mathrm{n} \in \{\mathrm{u}, \mathrm{c}\}$
$F^{P,N}$	Filtering matrix of degree D w.r.t quadrature of order N
f_h^{c}	Solution satisfying the discretized collided equations
f_h^{u}	Solution satisfying the discretized uncollided equations
$f_{i,j}^{P,N}$	Non-weighted i,j component of matrix $F^{P,N}$
$g_{i,j}^{D,N}$	Non-weighted i,j component of matrix $Q^{D,N}$
L^N	Streaming matrix w.r.t quadrature of order N
$L^{D,N}$	Streaming matrix of degree D w.r.t quadrature of order N
L_ϵ^{c}	Discrete streaming operator for the collided equations
L_ϵ^{u}	Discrete streaming operator for the uncollided equations
$m_{\ell,k}$	Real-valued normalized spherical harmonic function of degree ℓ and order k
$n(r)$	Normal vector at point r
N^*	Number of points or weights for quadrature of order N
N_{c}	Order of quadrature for collided equations
N_{u}	Order of quadrature for uncollided equations
P_ℓ	Legendre polynomial of degree ℓ
Q^N	Scattering source matrix w.r.t quadrature of order N
$Q^{D,N}$	Scattering source matrix of degree D w.r.t quadrature of order N
r	Point in space $(x, y, z) \in \mathbb{R}^3$

S	Known external source
t	Time
X	Spatial domain

Chapter 1

Introduction

1.1 Review of works

Radiative transport equations (RTEs) are kinetic equation models that are used to describe the movement of particles—including neutrons [15, 53], photons [63, 66], neutrinos [60–62, 74] and charged particles [79]—through a surrounding material medium. As they pass through, these particles interact with the material via scattering and emission/absorption processes.

Depending on the particles, several numerical methods have been used to simulate the transport equation. In many situations, numerical solutions are sought for the time-independent transport equation. However for certain problems like pulsed neutron experiments, photon transport in stellar atmospheres, and nuclear reactors, it becomes necessary to follow the time-dependent behavior of particle transport problems [53]. These particles have a wide range of different time scales. In particular, for nuclear reactors the time scale for neutrons produced by fission is on the order of 10^{-4} to 10^{-8} seconds while the characteristic half lives of delayed neutrons range from a few to several seconds [44]. Because of these wide disparities in characteristic times, many different approximation methods have been developed that allow for longer time steps. These include different forms of the kinetic equations that ignore certain time scales. For example, many multi-group kinetic equations are reduced to a prompt neutron transport equation to avoid the slower time scales of particles [53]. This form is also applicable to non-fissionable materials and photon transport. From this point many established time integration methods are used

to discretize the time variable. These include explicit methods (forward Euler, explicit Runge-Kutta, Adams-Bashforth) that have easy implementations of high order accurate methods at short time steps, and implicit methods (backward Euler, diagonolly implicit Runge-Kutta, Adams-Moulton, BDF) that allow for stable implementations at much longer time steps. When slower characteristic behaviors must be considered, then certain semi-implicit methods (TIMEX) are used to preserve some of the stability benefits of implicit methods without costly iteration methods for the inversion of the transport operator [36]. More details of certain time differencing methods are given in Chapter 2.

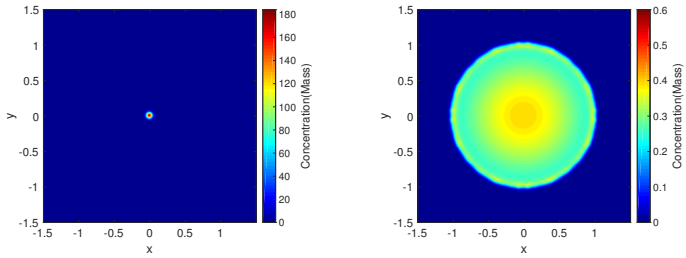
Many angular discretization methods exist for transport equations. In one dimensional geometries, spectral methods such as the P_N method approximate the solution with a truncated series of Legendre polynomials [25]. In the double P_N or DP_N, different expansions are used dependent on the direction of travel. It can be shown that under certain conditions these methods are equivalent to the discrete ordinates method which only requires the solution to be valid on a finite set of directions [53]. This is not the case for multi-dimensional geometries though. Discussion of the discrete ordinates method is left for later. Integral transport methods are another type of discretization [5, 21, 43, 70]. These are based on integrating out the angular dependence from the transport equation. These methods are almost exclusively used in neutron transport, but they are rarely used for photons as well [53]. The most widely used technique to solve the resultant integral transport equations is the method of collision probabilities [5, 21, 43, 70]. Even-parity transport methods are yet another way to discretize the radiation transport equation [42, 52, 78], although they are not as widely used as discrete ordinates or integral transport methods. They are derived by splitting the boundary into two pieces, one with a vacuum boundary condition and one with a reflective condition, and then dividing the solution into even and odd angular-parity components [53]. Other techniques include finite-element methods [55, 56, 68] and variational methods [9, 41]. Additionally, the Monte Carlo method [26, 73, 73] is a very popular technique for solving the transport equation for a variety of particles, in which a finite set of particles are simulated through the use of a pseudo random number generator.

Spatial discretization methods are highly dependent on the geometry of the problem being simulated, and in some cases, the angular discretization chosen. In one-dimensional,

or slab geometry, methods like finite difference and diamond difference are very popular, requiring very little computer memory when compared to other methods. However these methods have low accuracy in scattering dominated regimes with under resolved meshes. These methods have also been shown to be susceptible to negative fluxes. Remedies of this issue include the step method and the negative-flux-fixup used in conjunction with the diamond difference formulation [14]. Additionally, linear discontinuous methods [35, 68] are very popular as they are more accurate than aforementioned methods and are less susceptible to negative fluxes. Many of these methods have extensions in non-Cartesian geometries, where the representation of the direction a particle travels changes over a straight line. For multi-dimensional Cartesian geometries using discrete ordinates, we try to limit our choices to procedures that allow for the formation of a solution by sweeping the spatial grid in the direction of neutron travel.

1.2 Motivation

One of the methods discussed to discretize solutions to RTEs is the discrete ordinates (S_N) method [7, 31, 45, 48, 50, 64, 77]. This is a collocation method that approximates the radiative transport equation on a predefined set of directions and uses an associated quadrature rule to approximate integrals. The S_N method benefits from being easy to implement with good accuracy and flexibility, leading to algorithms of reasonable computational efficiency [53, 64]. There are, however, limitations. In problems with scattering dominated regions, the convergence of iterative solvers may become unacceptably slow [53]. Also, solutions obtained from the S_N method are not invariant under coordinate rotations. This lack of rotational invariance leads to "ray-effects": numerical artifacts aligned with the ordinate directions, often appearing in problems with regions of little scattering or localized sources [10, 49, 58]. These artifacts are demonstrated through a benchmark problem called the line source problem [23], where an initial pulse of particles distributed isotropically along an infinite line in space moves through a purely scattering material medium as time evolves. The density of particles for the initial condition and the analytic solution at $t = 1$ are shown in Figure 1.1. When the line source problem is solved using discrete ordinates,

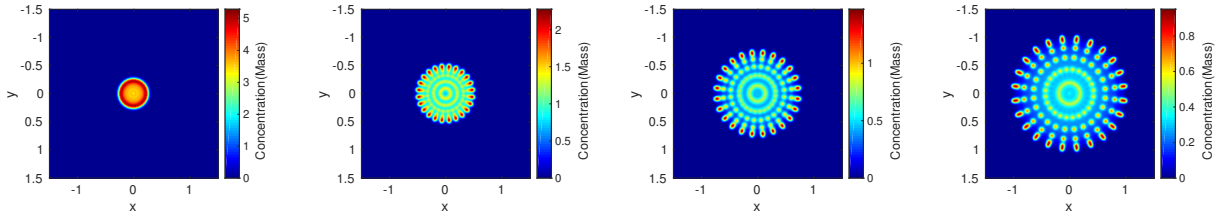


(a) Initial Condition, $t = 0$ (b) Semi-analytic solution, $t = 1$

Figure 1.1: Initial condition and semi-analytic solution for the line source benchmark.

then ray-effects are formed throughout the solution process, as demonstrated in Figure 1.2. Several attempts to remedy ray effects have been considered, including increasing the number of ordinates [49, 54], selecting particular quadrature sets [49, 54, 75], introducing trial functions which produce a direction-to-direction coupling in the representation of the streaming operator [40], modifying the differential operator to become similar to that of spherical harmonic operators [49], and use of bilinear and piece-wise constant finite element angular discretization[10].

This dissertation proposes to use a filtering technique applied to the discrete ordinates method to reduce the occurrence of ray-effects. The filtering technique is based on a modification of the solution algorithm to the spherical harmonics (P_N) method [58], which is known to have solutions with non-physical oscillations in regions of very little scattering. The method was further developed in [22, 67], where a modified hyperbolic system for the expansion coefficients is derived from a modified RTE with an additional term in the form of an anisotropic scattering operator. The benefit of this filtered RTE is that one can apply to it any type of angular discretization. We choose to modify the technique for the application of discrete ordinates simulations since discrete ordinates is the standard method to discretize

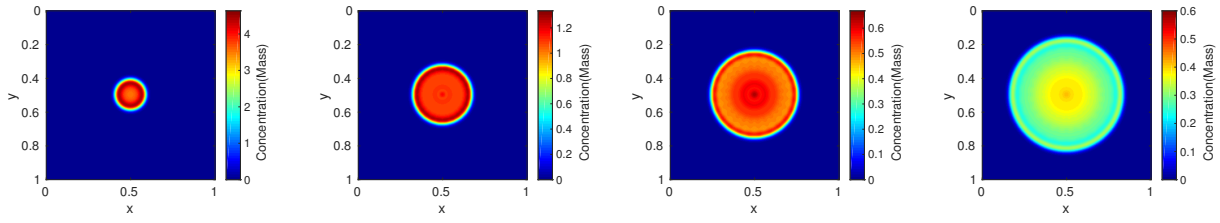


(a) $t = 0.25$ (b) $t = 0.5$ (c) $t = 0.75$ (d) $t = 1$

Figure 1.2: Discrete ordinates numerical solutions for the line source benchmark.

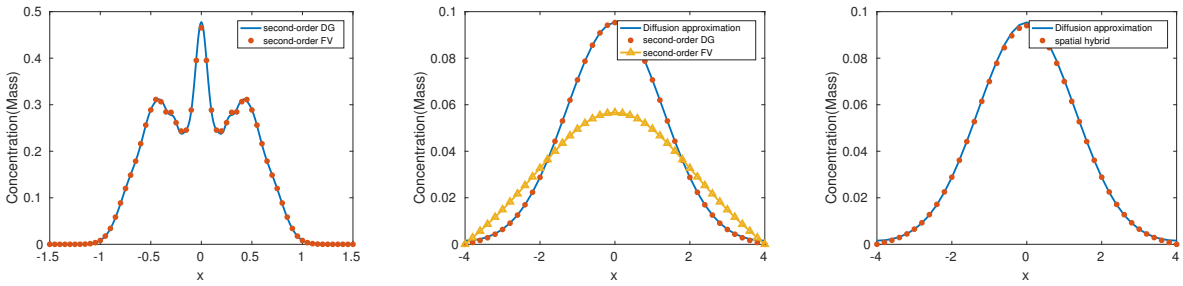
the angular component of transport equations. Figure 1.3 shows an example of a filtered discrete ordinates solution for the line source benchmark.

Another challenge in simulating the solution of the transport equation is capturing the diffusion limit numerically, where in scattering dominated regimes, the solution of the transport equation is accurately approximated by the solution of a diffusion equation. To accomplish this many numerical methods choose the discontinuous Galerkin (DG) method for spatial discretization. It was originally formulated in [68] for the purpose of solving neutron transport problems. Additionally, it offers better accuracy for problems with scattering dominated regimes over traditional discretizations like finite difference or finite volume (FV) methods allowing DG to capture the diffusion limit without resolving the mesh size to the minimal mean free path [2, 28, 47]. This however comes at the cost of increased computational time and memory usage. This dissertation proposes a spatial hybrid discretization strategy to reduce memory expenditure and accurately capture the diffusion limit in scattering dominated regimes (i.e. multiple mean free paths per cell). The strategy is based on a splitting scheme that separates the equation into two components: one for the first collision or collided component and one for streaming particles, which we will refer to as the uncollided component [3]. We note that the splitting scheme has been used to combine different angular discretizations [18, 59], but that it allows for the flexibility to discretize any of the variables of the phase space in each equation separately. Therefore, we choose to simulate the collided equation with discontinuous Galerkin and the uncollided equation with finite volume. This has the effect of reducing the number of degrees of freedom in the overall scheme and leads to a less computationally complex code. Figure 1.4 gives several comparisons between different spatial discretizations to the line source benchmark.



(a) $t = 0.25$ (b) $t = 0.5$ (c) $t = 0.75$ (d) $t = 1$

Figure 1.3: Filtered discrete ordinates numerical solutions for the line source benchmark.



(a) Non-scattering dominated (b) Scattering dominated (c) Scattering dominated

Figure 1.4: Comparison of different spatial discretizations in both non-scattering dominated, (a), and scattering dominated regimes, (b) and (c).

Figure 1.4 (a), shows numerical solutions using the DG and FV methods in a non-scattering dominated regime. In Figure 1.4 (b), the DG and FV methods are compared against each other in a scattering dominated regime. The figure also shows what the solution to the diffusion equation is. We see that the FV method is unable to capture this limit. In Figure 1.4 (c), the proposed spatial hybrid is compared against the solution of the diffusion equation, demonstrating that the spatial hybrid method is able to capture the diffusion limit.

1.3 Scope of dissertation

The remainder of this work is organized as follows. In Chapter 2, we discuss the general background information for the topics of this dissertation. This includes the general form of the radiative transport equation and brief explanations of its notation. We include a few reductions of the equation that will serve as the focus for the development and applications of several numerical strategies and analysis. These reductions include the unit-speed, mono-energy form and a reduction to two-dimensional space. We also discuss the conditions in which the solution to a scaled form of the transport equation can be approximated by the solution of a diffusion equation. This discussion also includes a selection of numerical spatial discretizations and under which conditions will the methods preserve the diffusion limit. Finally, a few angular discretizations are presented. In the case of discrete ordinates, the discussion serves as a setting for the remainder of this work as it will be the main angular

discretization used throughout. For spherical harmonics and filtered spherical harmonics, the discussion gives further insight into the development of the method used in Chapter 3.

In Chapter 3, we apply the discrete ordinates method to a filtered modified form of the radiative transport equation proposed by Radice, Abdikamalov, Rezzolla, and Ott [67], and further developed by Frank, Hauck, and Kupper [22]. The work of these authors, expanded from the work of Hauck and McClarren [58], involved using a filtering process on a spectral method based discretization of the radiative transport equation, whose solution is represented as moments of the angular flux distribution. It was therefore necessary to achieve something similar but with a notion of discrete moments instead, leading to the filtered discrete ordinates equations. After discretizing the time and space variables, the resulting equations can be reformulated into a set of equations in the discrete moments, and then an appropriate iterative solver can be employed to solve the system. My research shows that the filtered discrete ordinates equations are stable and consistent to the solution of the radiative transport equation in the L^2 norm. My main result from the chapter is in the form of analysis that shows that the convergence order is highly dependent on the accuracy of the chosen quadrature rule, as well as the regularity of the solution and order of the filter. In the event that a quadrature rule with sufficient accuracy is chosen, then the minimum of the regularity and order of the filter would be an upper bound for the convergence of the numerical solution. Through some benchmark problems assuming 2-D in space simplification, we show that under certain circumstances the filtered discrete ordinates method is effective at reducing the occurrences of ray-effects. In most situations, the addition of the filter operator resulted in smaller errors in the L^2 norm. In the case where the scattering source term involved anisotropic scattering, where the direction a particle travels after a collision is dependent on the angle it was traveling before, the method proved to be efficient, as it did not add a significant amount of computational time to get a numerical solution with accuracy the same or better than a non-filtered simulation.

In Chapter 4, we formulate a spatial hybrid discretization strategy to reduce memory expenditure and accurately capture the diffusion limit in scattering dominated regimes (i.e. multiple mean free paths) using both DG and FV spatial discretizations. This has the effect of reducing the number of degrees of freedom in the overall scheme and leads to a less

computationally complex code. For the proposed spatial hybrid scheme, we show that it preserves the asymptotic diffusion limit in scattering dominated regimes. We also show that it can be combined with existing angular hybrid methods. Finally, we show that this method is more efficient, in terms of memory usage and computational time, than the standard DG discretization of the transport equation as well as the angular hybrid method using DG in both collided and uncollided equations while giving similar results.

In Chapter 5, we draw final conclusions and also propose any future work that had been considered, but was unable to be expanded upon fully.

1.4 Summary of main contributions

In Chapter 3, a set of filtered discrete ordinates equations are presented as a means to reduce the occurrence of ray-effects found in standard discrete ordinates solutions to radiation transport problems. We have investigated convergence properties of the filtered discrete ordinates equations (3.2) and have demonstrated how they can be discretized in time implicitly using a standard Krylov framework.

Following the analysis of the filtered P_N equations in [22], we show that the convergence rates depend on both the regularity of the solution to the transport equation and the order of the filter. According to our analysis, since the numerical solution is based on point-wise approximations, we expect slightly worse convergence rates than the ones in [22] when the order of the filter is not the limiting factor. Convergence rates also depend implicitly on the underlying precision of the quadrature rule. Numerical results also show that the choice of the expansion index P in the filter operator is important for both obtaining good qualitative results.

We have investigated the efficiency of filtered discrete ordinates when compared with standard S_N methods. Although the filter is effective at reducing errors, including the occurrence of ray-effects, the cost associated with producing a filtered result shows that the method may be ill-advised for problems with isotropic scattering. Indeed, using a standard method with more ordinates may produce a solution with sufficient accuracy in less time.

However, in problems involving anisotropic scattering, the filtered results achieve significantly better accuracy in a more efficient manner.

In Chapter 4, we present a hybrid spatial discretization of the radiation transport equation (RTE) based on the formulation introduced in [59]. Following the approach in [28], we show that, like standard DG, the hybrid spatial discretization converges, in the limit of infinite scattering, to a consistent discretization solution of the diffusion limit (2.52). We also demonstrate the hybrid approach is more efficient, in terms of memory usage and time to solution than the standard DG. The formulation in [59] allows for hybridization in both space and angle, and we show how a combination of the two can improve the efficiency of simulations for two benchmark problems.

Chapter 2

Background

2.1 Radiative transport equation

2.1.1 Formulation

According to [53], a general transport equation describing particles interacting with a material medium through scattering and emission/absorption processes at time t, position $r = (x, y, z)$, speed c, angle $\Omega = (\Omega_x, \Omega_y, \Omega_z)$, and energy level E is described in the following way. Suppose V is a measurable subset of the phase space with differential $drd\Omega dE$, and $f(t, r, \Omega, E)$ is a function such that

$$\int_V f(t, r, \Omega, E)\ drd\Omega dE \tag{2.1}$$

describes the total number of particles at time t with $(r, \Omega, E) \in V$. Then the angular flux $\psi(t, r, \Omega, E)$ is defined as

$$\psi(t, r, \Omega, E) = cf(t, r, \Omega, E), \tag{2.2}$$

and satisfies the following:

$$\frac{1}{c}\frac{\partial}{\partial t}\psi(t, r, \Omega, E) + \Omega \cdot \nabla_r \psi(t, r, \Omega, E) + \sigma_{\mathrm{t}}(r, E)\psi(t, r, \Omega, E) = S(t, r, \Omega, E). \tag{2.3}$$

Here the total cross section $\sigma_{\mathrm{t}}(r, E)$ is the probable number of collisions per unit path length per particle. Additionally, $S(t, r, \Omega, E)\ dtdrd\Omega$ is the emission concentration defined as the number of particles emitted from collisions and sources with direction $d\Omega$ about Ω with energies dE about E and in the incremental volume dr during the time increment dt. The three contributions:

$$S = S_{\mathrm{ex}} + S_{\mathrm{s}} + S_{\mathrm{f}} \tag{2.4}$$

are attributed to external sources, scattered particles, and fission neutrons respectively. By external sources, we refer to a known distribution of source particles independent of the angular flux ψ, whereas both S_{s} and S_{f} are functions of ψ.

For this thesis, we consider the case $S_{\mathrm{f}} = 0$, implying that fission does not occur within the system. This case is referred to as a nonmultiplying system. For the scattering contribution, we introduce the differential scattering cross section $\Sigma_{\mathrm{s}}(r, \Omega \cdot \Omega', E \to E')$, where $\Sigma_{\mathrm{s}}(r, \Omega \cdot \Omega', E \to E')\ d\Omega dE$ is the probability per unit path length that particles at position r with energy E' traveling in direction Ω' scatter into dE about E and into the cone of directions $d\Omega$ about Ω. Since $\psi(t, r, \Omega, E)\ dtdrd\Omega dE$ is the total number of path lengths traveled by particles in the differential phase space during dt, we integrate over E' and Ω' and define the scattering source contribution as

$$S_{\mathrm{s}}(t, r, \Omega, E) = \int_{E'} \int_{\Omega'} \Sigma_{\mathrm{s}}(r, \Omega \cdot \Omega', E \to E')\psi(t, r, \Omega', E')\ d\Omega' dE'. \tag{2.5}$$

For the remainder of this thesis, we consider a reduction to a form concerning the systems of mono-energetic particles moving with unit speed. This implies that $c = 1$ and ψ, q and σ_{t} are independent of E. We also assume that the differential scattering cross section is separable over its independent variables r and $\Omega \cdot \Omega'$ into the product of two functions:

$$\Sigma_{\mathrm{s}}(r, \Omega \cdot \Omega') = \sigma_{\mathrm{s}}(r) g(\Omega \cdot \Omega'), \tag{2.6}$$

where the scattering cross section σ_{s} is the probable number of collisions that result in a change of direction per unit path length per particle, and the phase function g is dependent

on only the angle between two directions and gives the probability of a particle traveling in direction Ω' to change to direction Ω after a collision.

Let $X \subset \mathbb{R}^3$ be an open domain with Lipschitz continuous boundary ∂X, let $\mathbb{S}^2$ be the unit sphere in $\mathbb{R}^3$, and let $\Gamma := X \times \mathbb{S}^2$. Define the following sets:

$$\partial\Gamma := \partial X \times \mathbb{S}^2, \quad \partial\Gamma^- := \{(r,\Omega) \in \partial\Gamma \colon \Omega \cdot n(r) < 0\}, \quad \partial\Gamma^+ := \{(r,\Omega) \in \partial\Gamma \colon \Omega \cdot n(r) > 0\},$$

where $n(r)$ is the unit normal vector at r. We consider a linear radiative transport equation (RTE) [16] with complete initial-boundary value problem taking the form

$$(\partial_t + \mathcal{L})\psi(t,r,\Omega) = \sigma_{\mathrm{s}}(r)(\mathcal{Q}\psi)(t,r,\Omega) + S(t,r,\Omega), \quad (t,r,\Omega) \in (0,T) \times \Gamma, \tag{2.7a}$$

$$\psi(t,r,\Omega) = \psi_0(r,\Omega), \quad (t,r,\Omega) \in \{0\} \times \Gamma, \tag{2.7b}$$

$$\psi(t,r,\Omega) = \psi_{\mathrm{b}}(t,r,\Omega), \quad (t,r,\Omega) \in (0,T) \times \partial\Gamma^-, \tag{2.7c}$$

where

$$(\mathcal{L}\psi)(t,r,\Omega) = \Omega \cdot \nabla_r \psi(t,r,\Omega) + \sigma_{\mathrm{t}}(r)\psi(t,r,\Omega), \tag{2.8}$$

and ψ_0 and ψ_{b} are the initial condition and boundary data respectively. Here we have relabeled

$$S(t,r,\Omega) := S_{\mathrm{ex}}(t,r,\Omega) \quad \text{and} \quad \sigma_{\mathrm{s}}(r)(\mathcal{Q}\psi)(t,r,\Omega) := S_{\mathrm{s}}(t,r,\Omega) \tag{2.9}$$

where the source S is a known function and the total cross section σ_{t} and the scattering cross section σ_{s} are non-negative functions. Additionally we define the absorption cross section

$$\sigma_{\mathrm{a}}(r) = \sigma_{\mathrm{t}}(r) - \sigma_{\mathrm{s}}(r) \tag{2.10}$$

to also be non-negative. We let $\mathcal{Q}$ be a scattering source operator and define it as

$$(\mathcal{Q}v)(\Omega) = \int_{\mathbb{S}^2} g(\Omega \cdot \Omega')v(\Omega')\, d\Omega', \quad \forall\, v \in L^2(\mathbb{S}^2), \tag{2.11}$$

where the phase function $g\colon [-1,1] \to \mathbb{R}^+$ is a non-negative L^1 function, normalized so that

$$\int_{\mathbb{S}^2} g(\Omega \cdot \Omega')\, d\Omega' = 2\pi \int_{-1}^{1} g(\mu)\, d\mu = 1, \quad \forall\, \Omega \in \mathbb{S}^2. \tag{2.12}$$

The existence and uniqueness of the solution of (2.7) are discussed in [19]. When $\psi_{\mathrm{b}} = 0$, the main result is the following.

Theorem 2.1. *[19, XXI. 2.3]. Suppose the data for (2.7) satisfies*

$$\sigma_{\mathrm{s}}, \sigma_{\mathrm{a}} \in L^\infty((0,T) \times X), \quad S \in L^p((0,T) \times \Gamma), \quad \psi_0 \in L^p(\Gamma), \quad p \in [1,\infty).$$

Then (2.7) has a unique weak solution $\psi \in C^0([0,T]; L^p(\Gamma))$. If in addition,

$$\Omega \cdot \nabla_x \psi_0 \in L^p(\Gamma), \quad \psi_0|_{\partial\Gamma^-} = 0, \quad \text{and} \quad S \in C^1([0,T]; L^p(\Gamma)),$$

then ψ is a strong solution of (2.7) with

$$\psi \in C^1([0,T]; L^p(\Gamma)), \quad \Omega \cdot \nabla_x \psi \in C([0,T]; L^p(\Gamma)), \quad \psi(t)|_{\partial\Gamma^-} = 0, \ \forall\, t \in [0,T].$$

Extending this result to the case that $\psi_{\mathrm{b}} \neq 0$ requires some additional technical assumptions, which are discussed in more detail in [19].

2.1.2 Properties of $\mathcal{Q}$

Many of the properties of the scattering source operator $\mathcal{Q}$ are described using spherical harmonics and Legendre polynomials. Let $m_{\ell,k}$ be the real-valued spherical harmonic of degree ℓ and order k[7, 65], normalized such that

$$\int_{\mathbb{S}^2} m_{\ell,k}(\Omega) m_{\ell',k'}(\Omega)\, d\Omega = \delta_{\ell,\ell'}\delta_{k,k'}, \tag{2.13}$$

where δ is the Kronecker delta function. According to [7, Eqn. 2.36], these functions have the following bound

$$\|m_{\ell,k}\|_{L^\infty(\mathbb{S}^2)} \leq \sqrt{\frac{2\ell+1}{4\pi}}, \quad \forall\ \ell \geq 0,\ -\ell \leq k \leq \ell. \tag{2.14}$$

For any non-negative integer M, the set $\{m_{\ell,k} : 0 \leq \ell \leq M, |k| \leq \ell\}$ forms a basis of $\mathbb{P}^M(\mathbb{S}^2)$, the space of polynomials on $\mathbb{S}^2$ with degree M or less. Additionally, let P_ℓ be the Legendre polynomial of degree ℓ [1, Ch. 8], normalized such that

$$\int_{-1}^{1} P_\ell(\mu) P_{\ell'}(\mu)\ d\mu = \frac{2\delta_{\ell,\ell'}}{2\ell+1}, \quad \forall\ \ell, \ell' \geq 0. \tag{2.15}$$

This normalization of the Legendre polynomials ensures that

$$\|P_\ell\|_{L^\infty([-1,1])} \leq 1, \quad \forall\ \ell \geq 0. \tag{2.16}$$

Using these polynomials, the moments of a generic function $u \in L^2[-1,1]$ and $v \in L^2(\mathbb{S}^2)$, are given by

$$\hat{u}_\ell = 2\pi \int_{-1}^{1} u(\mu) P_\ell(\mu)\ d\mu, \quad \text{and} \quad \hat{v}_{\ell,k} = \int_{\mathbb{S}^2} v(\Omega) m_{\ell,k}(\Omega)\ d\Omega. \tag{2.17}$$

When $u = g$, (2.12) and (2.16) imply, respectively, that

$$\hat{g}_0 = 1 \quad \text{and} \quad |\hat{g}_\ell| \leq 1, \quad \forall\ \ell \geq 0. \tag{2.18}$$

The operator $\mathcal{Q}$ satisfies the following properties:

- $\mathcal{Q}$ is bounded in $L^2(\mathbb{S}^2)$:

$$\|\mathcal{Q}v\|_{L^2(\mathbb{S}^2)} \leq \|v\|_{L^2(\mathbb{S}^2)}, \quad \forall\ v \in L^2(\mathbb{S}^2). \tag{2.19}$$

  This can be shown using Jensen's inequality with respect to the measure $g(\Omega \cdot \Omega')\ d\Omega'$.

- $\mathcal{Q}$ is self adjoint with respect to the usual inner product on $L^2(\mathbb{S}^2)$:

$$\int_{\mathbb{S}^2} v\mathcal{Q}u \; d\Omega = \int_{\mathbb{S}^2} u\mathcal{Q}v \; d\Omega, \quad \forall \; u, v \in L^2(\mathbb{S}^2). \tag{2.20}$$

- $\mathcal{Q}$ is conservative:

$$\int_{\mathbb{S}^2} \mathcal{Q}v \; d\Omega = \int_{\mathbb{S}^2} v \; d\Omega, \quad \forall \; v \in L^2(\mathbb{S}^2). \tag{2.21}$$

- $\mathcal{Q}$ has a sparse L^2 expansion [53, Sec. 1-4]:

$$\mathcal{Q}v = \sum_{\ell=0}^{\infty} \sum_{k=-\ell}^{\ell} \hat{g}_\ell \hat{v}_{\ell,k} m_{\ell,k}(\Omega), \quad \forall \; v \in L^2(\mathbb{S}^2). \tag{2.22}$$

- The spherical harmonics are eigenfunctions of $\mathcal{Q}$:

$$\mathcal{Q}m_{\ell,k} = \hat{g}_\ell m_{\ell,k}, \quad \forall \; \ell \geq 0, \; -\ell \leq k \leq \ell. \tag{2.23}$$

  This is a direct consequence of (2.22) with normalization (2.13).

For numerical purposes, the expansion in (2.22) is often truncated at some finite degree $D \in \mathbb{N}$. This gives rise to a truncated scattering source operator $\mathcal{Q}^D$, given by

$$(\mathcal{Q}^D v)(\Omega) := \sum_{\ell=0}^{D} \sum_{k=-\ell}^{\ell} \hat{g}_\ell \hat{v}_{\ell,k} m_{\ell,k}(\Omega), \quad \forall \; v \in L^2(\mathbb{S}^2). \tag{2.24}$$

2.1.3 2-D reduction

In this thesis many experiments reference a form of (2.7) where $X \subset \mathbb{R}^2$. However, Ω is still represented as a vector in 3-D space, $\Omega = (\Omega_x, \Omega_y, \Omega_z)$. Specifically, Ω can be represented as a function of two parameters: an azimuthal angle $\varphi \in [0, 2\pi]$ and a polar angle $\theta \in [0, \pi]$. If we define $\mu = \cos(\theta)$, then

$$\Omega = \left(\sqrt{1-\mu^2}\cos(\varphi), \sqrt{1-\mu^2}\sin(\varphi), \mu \right). \tag{2.25}$$

To formulate the 2-D reduction of (2.7), we let ψ be independent of the third spatial component z. This leads the transport operator $\mathcal{L}$ in (2.8) to become.

$$(\mathcal{L}\psi)(t,r,\Omega) = (\Omega_x\partial_x + \Omega_y\partial_y)\psi(t,r,\Omega) + \sigma_{\mathrm{t}}(r)\psi(t,r\Omega), \tag{2.26}$$

In this 2-D reduction, the solution ψ is even in the polar parameter μ, (i.e. $\psi(t,r,\varphi,\mu) = \psi(t,r,\varphi,-\mu)$). For the the scattering source term $\mathcal{Q}\psi$ from (2.22), this leads to several moments being zero. This will occur when the matching spherical harmonic is odd in μ, or $\hat{\psi}_{\ell,k} = 0$ when $\ell + k \equiv 1 \pmod 2$.

2.2 Angular discretization

In this section several angular discretizations of (2.7) are presented. The discrete ordinates method will serve as one of the main focuses of this thesis, as the methods developed in this work all use this discretization for the angular component of (2.7). The other discretizations, spherical harmonics and filtered spherical harmonics, serve as motivation for one of the methods developed in this thesis. They are presented here for completion.

2.2.1 Discrete ordinates (S_N)

We consider a family of quadrature rules indexed by a positive integer N. A quadrature rule of order N is defined by a set of N^* discrete angles $\{\Omega_i^N\}_{i=1}^{N^*} \subset \mathbb{S}^2$ and weights $\{w_i^N\}_{i=1}^{N^*}$. Here $N^* = N^*(N)$ is a positive integer that is monotonically increasing as a function of N; the exact form of this relationship depends on the quadrature choice. At a minimum, we require that

$$w_i^N > 0, \quad \forall\ i \geq 1 \qquad \text{and} \qquad \sum_{i=1}^{N^*} w_i^N = 4\pi. \tag{2.27}$$

The discrete ordinates (S_N) equations [48, 50, 53] approximate the solution of the RTE at a fixed set of points in $\mathbb{S}^2$ and approximate any integrals by some quadrature using the same

set of points. Then the vector-valued function

$$\boldsymbol{\psi}^N(t,r) = [\psi_1^N(t,r), \psi_2^N(t,r), \ldots, \psi_{N^*}^N(t,r)]^T \tag{2.28}$$

satisfies the S_N approximation of (2.7)

$$(\partial_t + L^N)\boldsymbol{\psi}^N(t,r) = \sigma_s(r)Q^N\boldsymbol{\psi}^N(t,r) + \mathbf{S}^N(t,r), \quad (t,r) \in (0,T)\times X \tag{2.29a}$$

$$\boldsymbol{\psi}^N(t,r) = \boldsymbol{\psi}_0^N(r), \quad (t,r) \in \{0\}\times X \tag{2.29b}$$

$$\psi_i^N(t,r) = \psi_b(t,r,\Omega_i^N), \quad (t,r) \in (0,T)\times \partial X_i^-, 1 \le i \le N^* \tag{2.29c}$$

where

$$(L^N\boldsymbol{\psi}^N)(t,r) = \Upsilon^N\boldsymbol{\psi}^N(t,r) + \sigma_a(r)\boldsymbol{\psi}^N(t,r) + \sigma_s(r)R^N\boldsymbol{\psi}^N(t,r) \tag{2.30}$$

and the entries of the $N^*\times N^*$ matrices Υ^N, Q^N, R^N, and g^N are

$$\Upsilon_{i,j}^N = \delta_{i,j}\Omega_i^N\cdot\nabla_r, \quad Q_{i,j}^N = w_j^N g_{i,j}^N, \quad R_{i,j}^N = \delta_{i,j}\sum_{k=1}^{N^*} Q_{i,k}^N, \quad g_{i,j}^N = g(\Omega_i^N\cdot\Omega_j^N), \quad 1\le i,j\le K. \tag{2.31}$$

Additionally

$$(\mathbf{S}^N)_i = S(t,r,\Omega_i^N), \quad (\boldsymbol{\psi}_0^N)_i = \psi_0(r,\Omega_i^N), \quad \partial X_i^\pm := \{x\in\partial X:\ \pm\Omega_i^N\cdot n(r) > 0\} \tag{2.32}$$

are, respectively, the discretized source, initial condition, and outflow/inflow boundaries with respect to Ω_i^N. The matrix operator Q^N inherits many of the properties of $\mathcal{Q}$ seen in Subsection 2.1.1. Define the weighted inner product and the associated norm as follows:

$$\langle\mathbf{u},\mathbf{v}\rangle_N = \sum_{i=1}^{N^*} w_i^N u_i v_i, \quad \|\mathbf{u}\|_{\ell_N^2} = \sqrt{\langle\mathbf{u},\mathbf{u}\rangle_N}, \quad \forall\ \mathbf{u},\mathbf{v}\in\mathbb{R}^{N^*}. \tag{2.33}$$

Then since g^N is symmetric, and Q^N is a self adjoint operator with respect to the inner product $\langle \cdot, \cdot \rangle_N$, i.e.,

$$\left\langle Q^N \mathbf{v}, \mathbf{u} \right\rangle_N = \left\langle \mathbf{v}, Q^N \mathbf{u} \right\rangle_N, \quad \forall\ \mathbf{u}, \mathbf{v} \in \mathbb{R}^{N^*}. \tag{2.34}$$

The presence of R^N in (2.29a) is to ensure a discrete form of the conservation property (2.21):

$$\left\langle \mathbf{e}^N, Q^N \mathbf{v} \right\rangle_N = \left\langle \mathbf{e}^N, R^N \mathbf{v} \right\rangle_N, \quad \forall\ \mathbf{v} \in \mathbb{R}^{N^*}, \tag{2.35}$$

where $\mathbf{e}^N = (1, \ldots, 1)^T \in \mathbb{R}^{N^*}$. Ideally R^N is the identity matrix, but this is not the case when the quadrature approximation of the integral of g is not exact.

2.2.2 Spherical harmonics (P_D)

The spherical harmonics method is another way to discretize the transport equation in angle. It is formed by approximating the solution of (2.7a) with a finite linear combination of the angular moments, seen in (2.17), and the spherical harmonic basis functions. For any $D \in \mathbb{N}$,

$$\psi(t, r, \Omega) \approx \psi_{\mathrm{P}_D}(t, r, \Omega) := \sum_{\ell=0}^{D} \sum_{k=-\ell}^{\ell} \hat{\psi}_{\ell,k}(t, r) m_{\ell,k}(\Omega). \tag{2.36}$$

Although traditionally referred to as the P_N equations, we will instead use the index D when referencing the expansion of the approximate solution so as to remain consistent with the rest of the thesis. As a spectral method, the P_D equations have been shown that if the solution to (2.7) is smooth enough, then the appoximation ψ_{P_D} will have formal spectral convergence and will preserve rotational invariance in the solution ψ [29, 34, 53]. The method then takes the inner product on $L^2(\mathbb{S})$ of (2.7) with the orthonormal basis of the space

$$\mathbb{P}_D = \left\{ \sum_{\ell=0}^{D} \sum_{k=-\ell}^{\ell} c_{\ell,k} m_{\ell,k} : c_{\ell,k} \in \mathbb{R} \text{ for } 0 \leq \ell \leq D, |k| \leq \ell \right\}. \tag{2.37}$$

Let $\mathbf{m}_\ell$ be the vector of the $2\ell + 1$ spherical harmonics in $\mathbb{P}_D$ with degree ℓ and let $\mathbf{m}$ be the vector of all $\mathbf{m}_\ell$ where $\ell \leq D$. Then we can write the expansion of the approximate solution

(2.36) as

$$\psi_{\mathrm{P}_D} = \mathbf{m}^T \mathbf{u}_{\mathrm{P}_D}, \tag{2.38}$$

where $\mathbf{u}_{\mathrm{P}_D}$ solves the following:

$$\big((\partial_t + \mathcal{L})\mathbf{m}^T \mathbf{u}_{\mathrm{P}_D}(t,r), \mathbf{m}\big) = \sigma_{\mathrm{s}}(r) G^D \mathbf{u}_{\mathrm{P}_D}(t,r) + \mathbf{S}(t,r), \qquad (t,r) \in (0,T) \times X, \tag{2.39a}$$

$$\mathbf{u}_{\mathrm{P}_D}(0,r) = (\psi_0(r,\cdot), \mathbf{m}), \qquad r \in X, \tag{2.39b}$$

where $(\cdot,\cdot)$ is the inner product on $L^2(\mathbb{S}^2)$, $\mathbf{S}(t,r) = (S(t,r,\cdot),\mathbf{m})$, and G^D is a diagonal matrix with components $G^D_{(\ell,k),(\ell,k)} = \hat{g}_\ell$. This is a consequence of the finite expansion of the solution ψ_{P_D} and the fact that the basis functions $\mathbf{m}$ are eigen functions of $\mathcal{Q}$ (2.23). Using a recursion relation of the spherical harmonics

$$\Omega_j \mathbf{m}_\ell = \mathbf{a}^{(j)}_{\ell+1}\mathbf{m}_{\ell+1} + \left(\mathbf{a}^{(j)}_{\ell-1}\right)^T \mathbf{m}_{\ell-1}, \quad \mathbf{a}^{(j)}_\ell \in \mathbb{R}^{(2\ell-1)\times(2\ell+1)}, \quad j \in \{x,y,z\}, \tag{2.40}$$

we can form (2.39) explicitly as

$$\partial_t \mathbf{u}_{\mathrm{P}_D} + \Upsilon \cdot \nabla_r \mathbf{u}_{\mathrm{P}_D} + \sigma_{\mathrm{t}} \mathbf{u}_{\mathrm{P}_D} = \sigma_{\mathrm{s}} G^D \mathbf{u}_{\mathrm{P}_D} + \mathbf{S}, \tag{2.41}$$

where $\Upsilon = (\mathbf{m}\Omega, \mathbf{m})$, and the inner product between Υ and the gradient is

$$\Upsilon \cdot \nabla_r = \Upsilon_x \partial_x + \Upsilon_y \partial_y + \Upsilon_z \partial_z \tag{2.42}$$

where Υ_x, Υ_y and Υ_z have the following form:

$$\Upsilon_j = \begin{bmatrix} 0 & \mathbf{a}^{(j)}_1 & & \\ \left(\mathbf{a}^{(j)}_1\right)^T & 0 & \mathbf{a}^{(j)}_2 & \\ & \ddots & \ddots & \ddots \\ & & \left(\mathbf{a}^{(j)}_{D+1}\right)^T & 0 \end{bmatrix}, \quad j \in \{x,y,z\}. \tag{2.43}$$

Exact expression for $\mathbf{a}^{(j)}_\ell$ are discussed in [22, App. A].

2.2.3 Filtered Spherical harmonics (FP_D)

One of the issues with the spherical harmonic equations is that it is known to produce solutions with non-physical oscillations when the problem includes regions of very little scattering and absorption. This is a result from Gibbs phenomena, where a non-smooth function is approximated by a smooth basis [58]. One of the methods used to alleviate this issue is to use a filtering method. The filtered transport equation suggested in [22] is a modification of (2.7) that depends on a non-negative integer D and takes the form

$$\partial_t \psi^D(t,r,\Omega) + (\mathcal{L}\psi^D)(t,r,\Omega) = \sigma_{\mathrm{s}}\mathcal{Q}\psi^D(t,r,\Omega) + \sigma_{\mathrm{f}}\mathcal{F}^D\psi^D(t,r,\Omega) + S(t,r,\Omega). \tag{2.44}$$

Here the constant $\sigma_{\mathrm{f}} > 0$ is a filter strength,

$$(\mathcal{F}^D v)(\Omega) = \sum_{\ell=0}^{D}\sum_{k=-\ell}^{\ell} f_\ell^D m_{\ell,k}(\Omega)\hat{v}_{\ell,k}, \quad \forall\ v \in L^2(\mathbb{S}^2), \quad \text{and} \quad f_\ell^D = \log\left(f\left(\frac{\ell}{D+1}\right)\right), \tag{2.45}$$

where the filter function f is defined as follows:

Definition 2.2. *A filter of order $\alpha \in \mathbb{N}$ is a real-valued function $f \in C^\alpha(\mathbb{R}^+)$ that satisfies*

$$\textit{(i)}\ f(0) = 1, \quad \textit{(ii)}\ f^{(a)}(0) = 0, \quad \forall\ a \le \alpha - 1,\ a \in \mathbb{N}, \quad \textit{(iii)}\ \ f^{(\alpha)}(0) \neq 0,$$
$$\textit{(iv)}\ f(\eta) \in (0,1] \textit{ for any } \eta \in [0,1), \quad \textit{(v)}\ f \textit{ is monotonically decreasing.}$$

In the literature there are somewhat different definitions of a filter [12, 34, 58, 67, 76]. Typically, all definitions require conditions (i) and (ii). Condition (iii) is added to ensure the order of the filter is a unique property, and conditions (iv) and (v) are added to ensure stability. In [58], a filtering process was introduced as an update to the time integration scheme. After each time step, each term in the spherical harmonic expansion (2.36) is multiplied by an order-dependent coefficient. This usually results in filtering the higher order moments of the expansion more so than the low order ones. With a chosen filter

strength σ_{f}, this gave the new expansion

$$\psi_{\mathrm{FP}_D} := \sum_{\ell=0}^{D} \sum_{k=-\ell}^{\ell} \left(f\left(\frac{\ell}{D+1}\right)\right)^{\sigma_{\mathrm{f}}} \hat{\psi}_{\ell,k} m_{\ell,k}. \tag{2.46}$$

This formulation has the benefit of remaining rotationally invariant as it treats all moments of degree ℓ the same. Later in [22, 67], it was shown that if the filter process was changed to depend on the time step, then the new expansion would be consistent with a system of modified equations:

$$\partial_t \mathbf{u}_{\mathrm{FP}_D} + \Upsilon \cdot \nabla_r \mathbf{u}_{\mathrm{FP}_D} + \sigma_{\mathrm{t}} \mathbf{u}_{\mathrm{FP}_D} = \sigma_{\mathrm{s}} G^D \mathbf{u}_{\mathrm{FP}_D} + \sigma_{\mathrm{f}} G_{\mathrm{f}}^D \mathbf{u}_{\mathrm{FP}_D} + \mathbf{S}, \tag{2.47}$$

where G_{f}^D is a diagonal matrix with entries $(G_f^D)_{(\ell,k),(\ell,k)} = f_\ell^D$.

In [22], the analysis shows that the method is stable and consistent, and depending on the regularity of the solution of (2.7a) and the order of the filter, the analysis also shows what convergence order one could expect. In Chapter 3, we will show how this technique can be leveraged to a discrete ordinates formulation of (2.7a), and through similar analysis we will show under what conditions the method is stable and consistent. Additionally, convergence orders are determined when the order of the quadrature, N, the expansion index D of the scattering operator, and the expansion index P of the filter operator are all allowed to be independent of each other.

2.3 Diffusion limit

A well known approximation to the solution of (2.7) in scattering dominated regions is the diffusion limit. Consider an isotropic scaled version of (2.7) for dimensionless parameter ϵ:

$$(\epsilon\partial_t + \mathcal{L}_\epsilon)\psi(t,r,\Omega) = (\mathcal{Q}_{\mathrm{s},\epsilon}\psi)(t,r) + \epsilon S(t,r,\Omega), \quad (t,r,\Omega) \in (0,T)\times\Gamma, \tag{2.48a}$$

$$\psi(t,r,\Omega) = \psi_0(x,\Omega), \quad (t,r,\Omega) \in \{0\}\times\Gamma, \tag{2.48b}$$

$$\psi(t,r,\Omega) = \psi_{\mathrm{b}}(t,r,\Omega), \quad (t,r,\Omega) \in (0,T)\times\partial\Gamma^-, \tag{2.48c}$$

where

$$(\mathcal{L}_\epsilon \psi)(t, r, \Omega) = \Omega \cdot \nabla_r \psi(t, r, \Omega) + \frac{\sigma_{\mathrm{t}}(r)}{\epsilon} \psi(t, r, \Omega), \quad \sigma_{\mathrm{s},\epsilon} = \frac{\sigma_{\mathrm{t}}(r)}{\epsilon} - \epsilon \sigma_{\mathrm{a}}(r), \tag{2.49}$$

and

$$(\mathcal{Q}_{\mathrm{s},\epsilon})(r) = \left(\frac{\sigma_{\mathrm{t}}(r)}{\epsilon} - \epsilon \sigma_{\mathrm{a}}(r) \right) \mathcal{A}, \quad \text{with} \quad \mathcal{A} v = \frac{1}{4\pi} \int_{\mathbb{S}^2} v \, d\Omega, \quad \forall \, v \in L^1(\mathbb{S}^2). \tag{2.50}$$

In scattering dominated regions, (2.48a) can be approximated by the solution of a diffusion equation [30, 46]; that is, if

$$\inf_{r \in X} \sigma_{\mathrm{s}}(r) > 0 \quad \text{and} \quad \inf_{r \in X} \sigma_{\mathrm{a}}(r) > 0, \tag{2.51}$$

then for any compactly embedded subset $X_0 \Subset X$, there is an ϵ small enough that $\psi(t, r, \Omega) = \phi(t, r) + O(\epsilon)$ for all $r \in X_0$, where ϕ satisfies the following diffusion equation

$$\partial_t \phi(t, r) - \nabla_r \left(\frac{1}{3\sigma_{\mathrm{t}}(r)} \nabla_r \phi(t, r) \right) + \sigma_{\mathrm{a}}(r) \phi(t, r) = (\mathcal{A} q)(t, r). \tag{2.52}$$

Here X_0 must be bounded away from ∂X due to the boundary layers of width $O(\epsilon)$ which can appear in the solution of (2.48a) but not the solution of (2.52) [47].

For this thesis, we are interested in numerical methods that can capture the (interior) diffusion limit. In other words, we seek discretizations of (2.48a) that in the limit $\epsilon \to 0$, become a stable and consistent discretization of (2.52). In addition, we would like to capture the steady-state limit, which naturally lends itself toward an implicit time integration scheme. DG methods with sufficiently rich trial spaces [2, 28, 47] and finite volume methods with modified fluxes [27, 33, 38, 39] are two common spatial discretizations strategy for capturing the diffusion limit. Modifying the flux precludes the use of mesh sweeping techniques that rely on upwind information. Such techniques are a common tool for the iterative solution of steady-state and implicitly integrated problems. Hence DG methods with upwind fluxes are often preferable, even though they require more unknowns than a finite volume approach having the same formal order of accuracy.

Several angular discretizations have been shown to work well with upwind DG. The initial analysis of the diffusion limit can be found in [47]. There the discrete ordinates method is employed in a one-dimensional slab geometry. This analysis was later extended to the multi-dimensional setting in [2] for a variety of the different geometries. In [28], a finite element discretization is used for the angular variables, and the authors re-establish the results from [2] using functional analytic tools. Specifically, it is shown that the upwind DG approximation can capture the (interior) diffusion limit as long as the trial space is rich enough to support global linear functions. In [57], a spherical harmonic (P_N) angular discretization is combined with an upwind DG spatial discretization and a semi-implicit time integration scheme in order to achieve the diffusion limit. Due to its popularity and easy implementation, we use discrete ordinates for angular discretization in the paper.

2.4 Temporal and spatial discretizations

Part of the difficulty of solving transport equations numerically is making sure all the various discretizations that are employed work well together and serve to achieve certain objectives. The type of methods used in transport equations tend to be very problem dependent. One of my objectives is to develop a more robust strategy capable of solving a wider range of problems. For temporal discretizations, the methods can be described as either explicit, semi-implicit, or implicit. Explicit methods do very well in regimes that are not very diffusive or for problems that have a short time scale. They also lead to methods that are very easily parallelizable. However, certain problems require small CFL conditions which can become impractical at longer time scales. For semi-implicit methods, the multi-scale nature of certain problems can require an explicit discretization with an impractical time step restriction in the diffusion limit [59]. With a fully implicit method, the collisional, steady-state, and diffusive time scales can be resolved without restriction to the time step. Throughout this thesis, many of the methods are formulated using the backward Euler method. However, to increase the accuracy of the numerical solutions showcased in this work, the second-order diagonally implicit Runge Kutta method is used [4]. Further details of the form of the stages of the method using discrete ordinates can be found in Appendix A.

Two spatial discretizations used to get the diffusion limit are DG [2, 28, 47] or modified FV [38]. High-order finite volume approximations have been shown to give accurate results in streaming regimes with minimal computational cost; however, modifying the flux to achieve the diffusion limit prevents the use of sweeping based on upwind information. Another consideration is the upwind DG approximation using at least linear polynomial spaces when restricted to a spatial cell. This has been shown [2, 28, 47] to converge to the correct diffusion limit away from any boundary layers and maintains the ability to sweep through the spatial computational cells.

2.5 Benchmark Problems

In this section we discuss the details for two benchmark problems that are used throughout this thesis.

2.5.1 Line source benchmark

The line source is a benchmark problem that was first formulated in [23] as means to verify various time-dependent particle transport methods and assess any strengths or weaknesses. A robust review of various angular discretization techniques for the line source problem can be found in [24]. The problem describes an initial pulse of particles distributed isotropically along an infinite line in space moving through a purely scattering material medium as time evolves. In the reduced two-dimensional geometry, the initial pulse is expressed as a delta function at the origin of the two dimensional domain.

In order to simulate the line source, we approximate the initial condition with a Gaussian distribution with small standard deviation β:

$$\psi(0, x, y, \Omega) = \frac{1}{8\beta^2\pi^2} e^{\frac{-(x^2+y^2)}{2\beta^2}}. \tag{2.53}$$

This problem is simulated to final time $t = 1$ with an absorption cross section $\sigma_{\mathrm{a}} = 0$, scattering cross section $\sigma_{\mathrm{s}} = 1$, and source $S = 0$. All filtered numerical solutions use a filter

strength $\sigma_{\mathrm{f}} = 1.0$ and a 4th-order exponential filter function

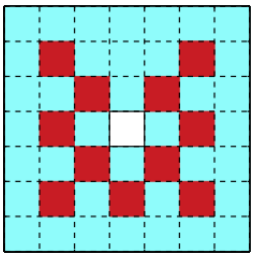
$$f(\eta) = \exp(\ln(\epsilon) \times \eta^4) \tag{2.54}$$

where $\epsilon = 2^{-52}$ is the machine epsilon using double precision. The semi-analytic solution to the line source with Gaussian initial condition (2.53) is described in [24] and is computed by convolution with a Green's function.

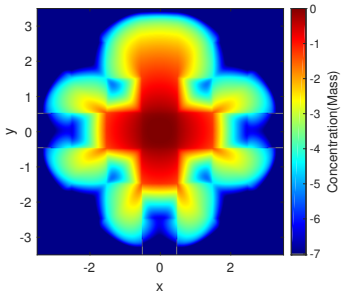
2.5.2 Lattice benchmark

The lattice test was first proposed in [11] as a cartoon loosely based on a nuclear reactor core assembly. The problem is a checkerboard of highly scattering and highly absorbing regions with vacuum boundaries as shown in Figure 2.1. The computational domain is a 7×7 square divided into smaller squares with side length one. The middle square is an isotropic source, surrounded by a checkerboard of purely scattering and purely absorbing squares as shown in Figure 2.1(a). The value at the locations given in Figure 2.1(a) are presented in Table 2.1.

The reference solution, unless stated otherwise, is simulated on a 504×504 spatial grid using the code StaRMAP [71], which uses a spherical harmonics approximation in angle, a staggered grid in space, and operator splitting in time. The reference solution was run with a high angular resolution using polynomials up to degree 129.



(a) problem layout



(b) P_{129} reference solution

Figure 2.1: Lattice problem layout and reference solution.

Graph (a) lists the locations for material cross-sections and source. Graph (b) shows cell-averaged particle concentrations at $t = 2.8$ on a logarithmic scale for the P_{129} reference.

Table 2.1: Material properties of Figure 2.1a.

Location	σ_a	σ_s	S
red squares	10	0	0
blue squares	0	1	0
white square	0	1	1

Chapter 3

Filtered discrete ordinates equations for radiative transport

In this chapter the filtered discrete ordinates equations are formulated by modifying (2.29a) with an additional filter operator that severs as a new scattering operator for the higher discrete moments. The stability and convergence of this new equation is analyzed, and several numerical examples are given to demonstrate the effectiveness of the filter, in terms of qualitative improvement and efficiency.

3.1 Formulation

Let $\Theta = \{D, P, N\}$ be a shorthand notation for given choices of non-negative integers D and P and positive integer N. Here D is the truncation index in the expansion of the collision operator $\mathcal{Q}^D$, P is the expansion index of the filter operator, and N is the index of the S_N quadrature for a particular solution. Then let

$$\boldsymbol{\psi}^{\Theta} = [\psi_1^{\Theta}, \ldots, \psi_{N^*}^{\Theta}]^T \tag{3.1}$$

be the solution to the filtered discrete ordinates equations

$$(\partial_t + L^{D,N})\boldsymbol{\psi}^{\Theta}(t,r) = (\sigma_{\mathrm{s}}(t,r)Q^{D,N} + \sigma_{\mathrm{f}}F^{P,N})\boldsymbol{\psi}^{\Theta} + \mathbf{S}^N(t,r),\ (t,r) \in (0,T)\times X \tag{3.2a}$$

$$\boldsymbol{\psi}^{\Theta}(t,r) = \boldsymbol{\psi}_0^N(r), \qquad (t,r) \in \{0\}\times X \tag{3.2b}$$

$$\psi_i^{\Theta}(t,r) = \psi_{\mathrm{b}}(t,r,\Omega_i^N), \qquad (t,r) \in (0,T)\times \partial X_i^-, \tag{3.2c}$$

for $1 \le i \le N^*$ where

$$(L^{D,N}\boldsymbol{\psi}^{\Theta})(t,r) = \Upsilon^N\boldsymbol{\psi}^{\Theta}(t,r) + \sigma_{\mathrm{a}}(t,r)\boldsymbol{\psi}^{\Theta}(t,r) + \sigma_{\mathrm{s}}(t,r)R^{D,N}\boldsymbol{\psi}^{\Theta}(t,r). \tag{3.3}$$

The matrices $Q^{D,N}, R^{D,N}, F^{P,N} \in \mathbb{R}^{N^*\times N^*}$ approximate $\mathcal{Q}$, the identity operator, and $\mathcal{F}^P$, respectively , and are given by

$$Q_{i,j}^{D,N} = w_j^N g_{i,j}^{D,N}, \quad R_{i,j}^{D,N} = \delta_{i,j}\sum_{k=1}^{N^*} Q_{i,k}^{D,N}, \quad F_{i,j}^{P,N} = w_j^N f_{i,j}^{P,N}, \quad 1 \le i,j \le N^*, \tag{3.4}$$

where

$$g_{i,j}^{D,N} := \sum_{\ell=0}^{D}\sum_{k=-\ell}^{\ell} \hat{g}_\ell m_{\ell,k}(\Omega_i^N)m_{\ell,k}(\Omega_j^N), \quad \text{and} \quad f_{i,j}^{P,N} := \sum_{\ell=0}^{P}\sum_{k=-\ell}^{\ell} f_\ell^P m_{\ell,k}(\Omega_i^N)m_{\ell,k}(\Omega_j^N). \tag{3.5}$$

Let $\mathbf{v} := [v_1, v_2, \ldots, v_{N^*}]^T$. Then in terms of the discrete moments

$$\hat{v}_{\ell,k}^N = \sum_{i=1}^{N^*} w_i^N m_{\ell,k}(\Omega_i^N)v_i, \tag{3.6}$$

$$(Q^{D,N}\mathbf{v})_i = \sum_{\ell=0}^{D}\sum_{k=-\ell}^{\ell} \hat{g}_\ell m_{\ell,k}(\Omega_i^N)\hat{v}_{\ell,k}^N, \quad (F^{P,N}\mathbf{v})_i = \sum_{\ell=0}^{D}\sum_{k=-\ell}^{\ell} f_\ell^P m_{\ell,k}(\Omega_i^N)\hat{v}_{\ell,k}^N. \tag{3.7}$$

These formulas are discrete analogs of (2.24) and (2.45).

3.2 Stability and convergence analysis

3.2.1 Stability

In this subsection, we show L^2 stability of the filtered discrete ordinates equations (3.2). The proof is a straight-forward energy estimate, but it requires properties of $Q^{D,N}$, $R^{D,N}$, and $F^{P,N}$ that are established in the following two propositions.

Proposition 3.1. *Suppose that the components of $g^{D,N}$, defined in (3.5), are non-negative. Then*

$$\left\langle Q^{D,N}\mathbf{v},\mathbf{v}\right\rangle_N \leq \left\langle R^{D,N}\mathbf{v},\mathbf{v}\right\rangle_N, \qquad \forall\ \mathbf{v}\in\mathbb{R}^{N^*}. \tag{3.8}$$

Proof. The proof is a simple application of the Cauchy-Schwarz inequality:

$$\begin{aligned}\left\langle Q^{D,N}\mathbf{v},\mathbf{v}\right\rangle_N &= \sum_{i=1}^{N^*}\sum_{j=1}^{N^*} w_i^N w_j^N g_{i,j}^{D,N} v_i v_j \\ &\leq \left(\sum_{i=1}^{N^*}\sum_{j=1}^{N^*} w_i^N w_j^N g_{i,j}^{D,N} v_i^2\right)^{1/2}\left(\sum_{i=1}^{N^*}\sum_{j=1}^{N^*} w_i^N w_j^N g_{i,j}^{D,N} v_j^2\right)^{1/2} \\ &= \sum_{i=1}^{N^*} w_i^N v_i^2 R_{i,i}^{D,N} = \left\langle R^{D,N}\mathbf{v},\mathbf{v}\right\rangle_N.\end{aligned} \tag{3.9}$$

□

Remark 1. *Given any $D>0$, it should be possible to find a set of points $\{\Omega_i^N\}_{i=1}^{N^*}\subset\mathbb{S}^2$ such that the non-negative condition is satisfied for all $1\leq i,j\leq N^*$.*

Proposition 3.2. *For any set of points $\{\Omega_i^N\}_{i=1}^{N^*}\subset\mathbb{S}^2$ and the set of associated weights $\{w_i^N\}_{i=1}^{N^*}$,*

$$\left\langle F^{P,N}\mathbf{v},\mathbf{v}\right\rangle_N \leq 0, \qquad \forall\ \mathbf{v}\in\mathbb{R}^{N^*}. \tag{3.10}$$

Proof. Using the formula for $F^{P,N}\mathbf{v}$ in (3.7) gives

$$\left\langle F^{P,N}\mathbf{v},\mathbf{v}\right\rangle_N = \sum_{i=1}^{N^*} w_i^N v_i \sum_{\ell=0}^{P}\sum_{k=-\ell}^{\ell} f_\ell^P m_{\ell,k}(\Omega_i^N)\hat{v}_{\ell,k}^N = \sum_{\ell=0}^{P}\sum_{k=-\ell}^{\ell} f_\ell^P\left(\hat{v}_{\ell,k}^N\right)^2 \leq 0. \tag{3.11}$$

□

For the following lemma we require a bit more notation. Let $u_{\mathrm{b}}\colon \partial\Gamma \to \mathbb{R}$, $\mathbf{u}\colon X \to \mathbb{R}^{N^*}$, and $\mathbf{u}_{\mathrm{b}}^N\colon \partial X \to \mathbb{R}^{N^*}$ be measurable functions where $\mathbf{u}_{\mathrm{b}}^N := [u_{\mathrm{b}}(r,\Omega_1^N), u_{\mathrm{b}}(r,\Omega_2^N), \dots, u_{\mathrm{b}}(r,\Omega_{N^*}^N)]^T$. We define the following norms:

$$\|\mathbf{u}\|_{L^2_N(X)} = \sqrt{\int_X \langle \mathbf{u},\mathbf{u}\rangle_N \, dr}, \quad \|\mathbf{u}_{\mathrm{b}}^N\|^2_{L^2_N(\partial X^{\pm})} = \sum_{i=1}^{N^*} w_i^N \int_{\partial X_i^{\pm}} |\Omega_i^N \cdot n(r)|\, |u_{\mathrm{b}}(r,\Omega_i^N)|^2 dr. \tag{3.12}$$

Lemma 3.3 (L^2 Stability)**.** *Suppose the data in (3.2) satisfies the following:*

- $\sigma_{\mathrm{s}}(t,r) \geq 0, \quad 0 < a \leq \sigma_{\mathrm{a}}(t,r), \ \forall\ t \in [0,T], \ x \in X; \quad \mathbf{S}^N \in L^\infty([0,T]; L^2_N(X));$
- $\|\boldsymbol{\psi}_{\mathrm{b}}^N\|^2_{L^2_N(\partial X^{\pm})} \in L^\infty[0,T], \quad$ *where* $\boldsymbol{\psi}_{\mathrm{b}}^N(t,r) := [\psi_{\mathrm{b}}(t,r,\Omega_1^N), \psi_{\mathrm{b}}(t,r,\Omega_2^N), \dots, \psi_{\mathrm{b}}(t,r,\Omega_{N^*}^N)]^T,$

and the components of $g^{D,N}$, defined in (3.5), are non-negative. Then for $t \in [0,T]$, the following bound holds[1]

$$\begin{aligned}\|\boldsymbol{\psi}^\Theta(t)\|^2_{L^2_N(X)} \leq e^{-at}\|\boldsymbol{\psi}^\Theta(0)\|^2_{L^2_N(X)} + \frac{1}{a^2}\left(1-e^{-at}\right)\|\mathbf{S}^N\|^2_{L^\infty([0,t];L^2_N(X))} \\ + \frac{1}{a}\left(1-e^{-at}\right)\|\boldsymbol{\psi}_b^N\|^2_{L^\infty([0,t];L^2_N(\partial X^-))}.\end{aligned} \tag{3.13}$$

Proof. We begin by taking the inner product of (3.2a) with $\boldsymbol{\psi}^\Theta$, integrating over X, applying the chain rule, and integrating by parts where appropriate to get the following:

$$\begin{aligned}\frac{1}{2}\partial_t\|\boldsymbol{\psi}^\Theta\|^2_{L^2_N(X)} + \frac{1}{2}\|\boldsymbol{\psi}_{\mathrm{b}}^N\|^2_{L^2_N(\partial X^+)} + \int_X \left\langle \sigma_{\mathrm{a}}\boldsymbol{\psi}^\Theta, \boldsymbol{\psi}^\Theta\right\rangle_N dr - \sigma_{\mathrm{f}}\int_X \left\langle F^{P,N}\boldsymbol{\psi}^\Theta, \boldsymbol{\psi}^\Theta\right\rangle_N dr \\ + \int_X \left\langle \sigma_{\mathrm{s}}(R^{D,N} - Q^{D,N})\boldsymbol{\psi}^\Theta, \boldsymbol{\psi}^\Theta\right\rangle_N dr = \int_X \left\langle \mathbf{S}^N, \boldsymbol{\psi}^\Theta\right\rangle_N dr + \frac{1}{2}\|\boldsymbol{\psi}_{\mathrm{b}}^N\|^2_{L^2_N(\partial X^-)}.\end{aligned} \tag{3.14}$$

Propositions 3.1 and 3.2 imply that

$$\int_X \left\langle \sigma_{\mathrm{s}}(R^{D,N} - Q^{D,N})\boldsymbol{\psi}^\Theta, \boldsymbol{\psi}^\Theta\right\rangle_N dr \geq 0 \quad \text{and} \quad \int_X \left\langle -\sigma_{\mathrm{f}}F^{P,N}\boldsymbol{\psi}^\Theta, \boldsymbol{\psi}^\Theta\right\rangle_N dr \geq 0. \tag{3.15}$$

[1] Given a generic time-dependent function $u\colon t \mapsto u(t) \in B$, with B a normed vector space, we abuse notation slightly by writing $\|u(t)\|_B = \|u\|_B(t)$.

Removing appropriate terms in (3.14) and applying the lower bound of σ_{a}, we get

$$\frac{1}{2}\partial_t\|\boldsymbol{\psi}^{\Theta}\|^2_{L^2_N(X)} + a\|\boldsymbol{\psi}^{\Theta}\|^2_{L^2_N(X)} \leq \int_X \left\langle \mathbf{S}, \boldsymbol{\psi}^{\Theta} \right\rangle_N \, dr + \frac{1}{2}\|\boldsymbol{\psi}^N_{\mathrm{b}}\|^2_{L^2_N(\partial X^-)}. \tag{3.16}$$

By Young's inequality,

$$\int_X \left\langle \mathbf{S}, \boldsymbol{\psi}^{\Theta} \right\rangle_N \, dr \leq \frac{a}{2}\|\boldsymbol{\psi}^{\Theta}\|^2_{L^2_N(X)} + \frac{1}{2a}\|\mathbf{S}^N\|^2_{L^2_N(X)}. \tag{3.17}$$

Therefore,

$$\frac{1}{2}\partial_t\|\boldsymbol{\psi}^{\Theta}\|^2_{L^2_N(X)} + \frac{a}{2}\|\boldsymbol{\psi}^{\Theta}\|^2_{L^2_N(X)} \leq \frac{1}{2a}\|\mathbf{S}^N\|^2_{L^2_N(X)} + \frac{1}{2}\|\boldsymbol{\psi}^N_{\mathrm{b}}\|^2_{L^2_N(\partial X^-)}. \tag{3.18}$$

Multiplying by $2e^{at}$ and integrating both sides from 0 to t gives

$$\begin{aligned} e^{at}\|\boldsymbol{\psi}^{\Theta}(t)\|^2_{L^2_N(X)} &\leq \|\boldsymbol{\psi}^{\Theta}(0)\|^2_{L^2_N(X)} + \int_0^t e^{as}\left[\frac{1}{a}\|\mathbf{S}^N(s)\|^2_{L^2_N(X)} + \|\boldsymbol{\psi}^N_{\mathrm{b}}(s)\|^2_{L^2_N(\partial X^-)}\right] ds \\ &\leq \|\boldsymbol{\psi}^{\Theta}(0)\|^2_{L^2_N(X)} + \frac{1}{a}\left[e^{at}-1\right]\left(\frac{1}{a}\|\mathbf{S}^N\|^2_{L^\infty([0,t];L^2_N(X))} + \|\boldsymbol{\psi}^N_{\mathrm{b}}\|^2_{L^\infty([0,t];L^2_N(\partial X^-))}\right). \end{aligned} \tag{3.19}$$

Multiplying both sides by e^{-at} gives the desired result. □

3.2.2 Preliminaries and notation

Define the evaluation operator $\mathcal{E}^N \colon C^0(\mathbb{S}^2) \to \mathbb{R}^{N^*}$ by

$$\mathcal{E}^N v = [v(\Omega^N_1), v(\Omega^N_2), \ldots, v(\Omega^N_{N^*})]^T, \quad \forall\ v \in C^0(\mathbb{S}^2). \tag{3.20}$$

For any $u, v \in C^0(\mathbb{S}^2)$ define the bilinear form and associated semi-norm:

$$(u,v)_N = \left\langle \mathcal{E}^N u, \mathcal{E}^N v \right\rangle_N \quad \text{and} \quad |v|_N = \sqrt{(v,v)_N}\,. \tag{3.21}$$

Definition 3.4. *Let N be a positive integer and M be a non-negative integer. The operator $\mathcal{I}^{N,M}\colon C^0(\mathbb{S}^2) \to \mathbb{P}^M$ is given by*

$$\mathcal{I}^{N,M} v(\Omega) = \sum_{\ell=0}^{M} \sum_{k=-\ell}^{\ell} \tilde{v}_{\ell,k}^N m_{\ell,k}(\Omega), \ \forall\ v \in C^0(\mathbb{S}^2), \quad \text{where } \tilde{v}_{\ell,k}^N = (v, m_{\ell,k})_N\,. \tag{3.22}$$

Definition 3.5. *Let N be a positive integer, and let $(v)_N := (v,1)_N$ be a quadrature rule with abscissas $\{\Omega_i^N\}_{i=1}^{N^*}$ and weights $\{w_i^N\}_{i=1}^{N^*}$. We say the rule $(\cdot)_N$ has* **precision** *p if*

$$(v)_N = \int_{\mathbb{S}^2} v(\Omega)\ d\Omega, \quad \forall\ v \in \mathbb{P}^p(\mathbb{S}^2), \tag{3.23}$$

and there exists some $v \in \mathbb{P}^{p+1}(\mathbb{S}^2)$ such that

$$(v)_N \neq \int_{\mathbb{S}^2} v(\Omega)\ d\Omega. \tag{3.24}$$

The following lemma generalizes the result in [29, Lem. 8.5] to the case when $N \neq M$. The proofs are similar.

Lemma 3.6. *Let N be a positive integer and M be a non-negative integer. If the quadrature rule $(\cdot)_N$ has precision $p \geq 2M$, then for any $v \in C^0(\mathbb{S}^2)$ and $u \in \mathbb{P}^M(\mathbb{S}^2)$,*

$$\text{(i) } \mathcal{I}^{N,M} u = u, \quad \text{(ii) } (\mathcal{I}^{N,M} v, u) = \left(\mathcal{I}^{N,M} v, u\right)_N = (v,u)_N\,, \quad \text{(iii) } |v - \mathcal{I}^{N,M} v|_N \leq |v-u|_N, \tag{3.25}$$

where $(\cdot,\cdot)$ is the usual inner product in $L^2(\mathbb{S}^2)$.

Proof. For any $u \in \mathbb{P}^M(\mathbb{S}^2)$, the precision of the quadrature rule implies that $\hat{u}_{\ell,k} = \tilde{u}_{\ell,k}^N$ for all ℓ, k such that $|k| \leq \ell \leq M$. Hence

$$\mathcal{I}^{N,M} u = \sum_{\ell=0}^{M} \sum_{k=-\ell}^{\ell} \tilde{u}_{\ell,k}^N m_{\ell,k} = \sum_{\ell=0}^{M} \sum_{k=-\ell}^{\ell} \hat{u}_{\ell,k} m_{\ell,k} = u, \tag{3.26}$$

which proves (i).

The first equality in (ii) is a direct consequence of the assumption on the quadrature precision. For the second equality in (ii), a direct calculation gives, for any $u \in \mathbb{P}^M(\mathbb{S}^2)$ and $v \in C^0(\mathbb{S}^2)$,

$$\left(\mathcal{I}^{N,M} v, u\right)_N = \left(\sum_{\ell=0}^{M} \sum_{k=-\ell}^{\ell} \tilde{v}_{\ell,k} m_{\ell,k}, u \right)_N = \sum_{\ell=0}^{M} \sum_{k=-\ell}^{\ell} \tilde{v}_{\ell,k} \left(m_{\ell,k}, u\right)_N = \sum_{\ell=0}^{M} \sum_{k=-\ell}^{\ell} \tilde{v}_{\ell,k} \hat{u}_{\ell,k}, \tag{3.27a}$$

$$\left(v, u\right)_N = \left(v, \sum_{\ell=0}^{M} \sum_{k=-\ell}^{\ell} \hat{u}_{\ell,k} m_{\ell,k} \right)_N = \sum_{\ell=0}^{M} \sum_{k=-\ell}^{\ell} \hat{u}_{\ell,k} \left(v, m_{\ell,k}\right)_N = \sum_{\ell=0}^{M} \sum_{k=-\ell}^{\ell} \tilde{v}_{\ell,k} \hat{u}_{\ell,k}, \tag{3.27b}$$

where in the first line above, we have again used the fact that $\hat{u}_{\ell,k} = \tilde{u}^N_{\ell,k}$ when $|k| \leq \ell \leq M$.

To show (iii), we use the fact that $\left(v - \mathcal{I}^{N,M} v, u - \mathcal{I}^{N,M} v\right)_N = 0$, which follows from (ii). When combined with the Cauchy-Schwarz inequality, this gives

$$|v - \mathcal{I}^{N,M} v|_N^2 = \left(v - \mathcal{I}^{N,M} v, v - u\right)_N + \left(v - \mathcal{I}^{N,M} v, u - \mathcal{I}^{N,M} v\right)_N \leq |v - \mathcal{I}^{N,M} v|_N \, |v - u|_N. \tag{3.28}$$

If $v = \mathcal{I}^{N,M} v$, (iii) holds trivially. Otherwise, dividing both sides by $|v - \mathcal{I}^{N,M} v|_N$ gives the desired result. □

Let $\mathcal{P}^M \colon L^2(\mathbb{S}^2) \to \mathbb{P}^M(\mathbb{S}^2)$ be the orthogonal projection with respect to the standard inner product, and define $H^r(\mathbb{S}^2)$ to be the completion of $C^\infty(\mathbb{S}^2)$ with respect to the norm [7, Def. 3.23]

$$\|v\|^2_{H^r(\mathbb{S}^2)} = \sum_{\ell=0}^{\infty} \sum_{k=-\ell}^{\ell} (\ell + 0.5)^{2r} |\hat{v}_{\ell,k}|^2, \quad \forall \; r \geq 0. \tag{3.29}$$

Since the unit sphere is a compact 2-D manifold, by Sobolev Embedding Theorem [8, Thm. 2.10 and Thm. 2.20], for any $r > 1$, $H^r(\mathbb{S}^2) \subset C^0(\mathbb{S}^2)$. As a consequence, $\mathcal{I}^{N,M}$ is well-defined for functions in $H^r(\mathbb{S}^2)$ when $r > 1$. Moreover there exists a constant $C = C(r)$ such that

$$\|v\|_{L^\infty(\mathbb{S}^2)} \leq C \|v\|_{H^r(\mathbb{S}^2)}, \quad \forall \; v \in H^r(\mathbb{S}^2). \tag{3.30}$$

Lemma 3.7. *Let N be a positive integer, and suppose $(\cdot)_N$ is a quadrature rule with precision $p \in \mathbb{N}$. Then for any non-negative integer $M \leq p/2$ and for any $v \in H^r(\mathbb{S}^2)$, $r > 1$, there exists a constant C, independent of M, such that*

$$\|\mathcal{P}^M v - \mathcal{I}^{N,M} v\|_{L^2(\mathbb{S}^2)} \leq C M^{1+\delta-r} \|v\|_{H^r(\mathbb{S}^2)}, \tag{3.31}$$

with δ an arbitrary positive constant.

Proof. Let $v \in H^r(\mathbb{S}^2)$ for some $r > 1$. The Sobolev Embedding Theorem implies that v is continuous; hence so are $v - \mathcal{I}^{N,M} v$ and $v - \mathcal{P}^M v$. If $p \geq 2M$, then

$$\|\mathcal{P}^M v - \mathcal{I}^{N,M} v\|_{L^2(\mathbb{S}^2)} = |\mathcal{P}^M v - \mathcal{I}^{N,M} v|_N \leq |v - \mathcal{I}^{N,M} v|_N + |v - \mathcal{P}^M v|_N \leq 2|v - \mathcal{P}^M v|_N, \tag{3.32}$$

where the last inequality follows by applying Lemma 3.6 (iii) with $u = \mathcal{P}^M v$. A direct calculation using the definition of the semi-norm $|\cdot|$ shows that

$$|v - \mathcal{P}^M v|_N \leq \sqrt{4\pi} \|v - \mathcal{P}^M v\|_{L^\infty(\mathbb{S}^2)}. \tag{3.33}$$

Substituting (3.33) into (3.32) and then applying (3.30) gives

$$\|\mathcal{P}^M v - \mathcal{I}^{N,M} v\|_{L^2(\mathbb{S}^2)} \leq C \|v - \mathcal{P}^M v\|_{H^{1+\delta}(\mathbb{S}^2)}, \tag{3.34}$$

where the constant C does not depend on M and δ is an arbitrary positive constant. Lastly, according to [29, Lem. 8.1 and Thm. 8.2],

$$\|v - \mathcal{P}^M v\|_{H^{1+\delta}(\mathbb{S}^2)} \leq C M^{1+\delta-r} \|v\|_{H^r(\mathbb{S}^2)}, \quad \forall v \in H^r(\mathbb{S}^2). \tag{3.35}$$

Together (3.34) and (3.35) give the desired result. □

Assumption 1. *Hereafter, we make the additional assumption that the filter f satisfies*

$$\textit{(vi)}\ f(\eta) \geq C(1-\eta)^k, \quad \eta \in [\eta_0, 1] \tag{3.36}$$

for some $k \geq 0$, some constant C, and some $\eta_0 \in (0, 1)$.

3.2.3 Convergence of the filtered discrete ordinates equations

We now show L^2 convergence of the solution of the filtered discrete ordinates equations (3.2) to the solution ψ of the RTE (2.7). Let

$$\mathbf{e}^{\Theta} = \boldsymbol{\psi} - \boldsymbol{\psi}^{\Theta}, \quad \text{where } \boldsymbol{\psi} = \mathcal{E}^N \psi, \tag{3.37}$$

be the error on the quadrature points $\{\Omega_i^N\}_{i=1}^{N^*}$. Our main result is the following:

Theorem 3.8. *Assume that $\psi \in C^0([0,T]; L^2(X; H^r(\mathbb{S}^2)))$ for $r > 3/2$ and that g satisfies the following additional regularity condition:*

$$\sum_{\ell=0}^{\infty} (\ell + 0.5)^{2s} \hat{g}_\ell^2 < \infty, \tag{3.38}$$

where $s \geq 1$, and the components of $g^{D,N}$, defined in (3.5), are non-negative. Additionally, suppose the quadrature rule $(\cdot)_N$ associated to the discrete ordinates equation (2.29) has precision $p \geq \max\{2D, 2P\}$ and the filter function f used to define $\mathcal{F}^P$ in (2.45) is of order α. Then for any $t \in [0,T]$, there exists a constant C, which depend on g and ψ, such that

$$\begin{aligned} \|\mathbf{e}^{\Theta}(t)\|_{L_N^2(X)} \leq C \Big[& \|\mathcal{E}^N \mathcal{Q}^D \psi - Q^{D,N} \boldsymbol{\psi}\|_{L^\infty([0,t]; L_N^2(X))} \\ & + \|\mathcal{E}^N \mathcal{Q} \psi - \mathcal{E}^N \mathcal{Q}^D \psi\|_{L^\infty([0,t]; L_N^2(X))} + \|F^{P,N} \boldsymbol{\psi}\|_{L^\infty([0,t]; L_N^2(X))} \Big], \end{aligned} \tag{3.39}$$

where the terms above satisfy the following bounds:

$$\|\mathcal{E}^N \mathcal{Q}^D \psi - Q^{D,N} \boldsymbol{\psi}\|_{L^\infty([0,t]; L_N^2(X))} \leq C(D+1)^{-\left(r - \frac{3+\delta}{2}\right)}, \tag{3.40a}$$

$$\|\mathcal{E}^N \mathcal{Q} \psi - \mathcal{E}^N \mathcal{Q}^D \psi\|_{L^\infty([0,t]; L_N^2(X))} \leq C(D+1)^{-(r+s-(1+\delta))}, \tag{3.40b}$$

$$\|F^{P,N} \boldsymbol{\psi}\|_{L^\infty([0,t]; L_N^2(X))} \leq \begin{cases} C(P+1)^{-(r-3/2-\delta)}, & \alpha > r - \delta - 3/2, \\ C(P+1)^{-\alpha+\epsilon}, & \alpha \leq r - \delta - 3/2, \end{cases} \tag{3.40c}$$

for any $\delta, \epsilon > 0$.

Remark 2. *Lemma 3.11 provides a more general estimate for the quadrature error in (3.40a) when $p \geq 2D$. The estimate in (3.40a) is obtained by letting $p = 2D$ in Lemma 3.11. The*

error will decrease if p is allowed to be greater than $2D$. Additionally, the estimates in Theorem 3.8 depend implicitly on N, as the precision of the quadrature will need to increase if D and P increase. This implies that the theorem is not valid for any quadrature whose precision does not increase when N increases.

According to Theorem 3.8, the discrete L^2 error created by approximating the solution of the RTE using filtered discrete ordinates has three contributions: a quadrature error (3.40a), a truncation error (3.40b) due to the polynomial approximation of the phase function, and an error due to the filter (3.40c). The truncation error decays much more rapidly than the other two terms due to the regularity of the phase function. Indeed for an even integer s, the condition on (3.38) implies that $g \in H^s[-1,1]$. What Theorem 3.8 does not do is to explain why the filtered equations in practice have better accuracy than the non-filtered ones. This analysis will be left to future work.

Proof of Theorem 3.8. Applying $\mathcal{E}^N$ to (2.7) and subtracting the equations for $\boldsymbol{\psi}^\Theta$ in (3.2), we see that $\mathbf{e}^\Theta$ satisfies

$$(\partial_t + L^{D,N})\mathbf{e}^\Theta(t,r) = (\sigma_{\mathrm{s}}(t,r)Q^{D,N} + \sigma_{\mathrm{f}}F^{P,N})\mathbf{e}^\Theta(t,r) + \mathbf{S}_{\mathbf{e}}(t,r),\ (t,r) \in (0,T)\times X \tag{3.41a}$$

$$\mathbf{e}^\Theta(t,r) = 0, \qquad (t,r) \in \{0\}\times X \tag{3.41b}$$

$$(\mathbf{e}^\Theta)_i(t,r) = 0, \qquad (t,r) \in (0,T)\times \partial X_i^-, \tag{3.41c}$$

for $1 \le i \le N^*$ where

$$\mathbf{S}_{\mathbf{e}}(t,r) = \sigma_{\mathrm{s}}(t,r)(R^{D,N} - I^N)\boldsymbol{\psi}(t,r) + \sigma_{\mathrm{s}}(t,r)(\mathcal{E}^N\mathcal{Q}\psi(t,r) - Q^{D,N}\boldsymbol{\psi}(t,r)) - \sigma_{\mathrm{f}}F^{P,N}\boldsymbol{\psi}(t,r). \tag{3.42}$$

The inequality in (3.39) is then achieved by applying Lemma 3.3 to (3.41) and treating $\mathbf{S}_{\mathbf{e}}$ as a new source, which is in $L^\infty([0,T],L^2_N(X))$ if $\psi \in L^\infty([0,T];L^2(X;H^r(\mathbb{S}^2)))$. Application of the triangle inequality and Lemma 3.9 give the right-hand side of (3.39). Estimates of each term are shown in the lemmas below. □

Lemma 3.9. *Suppose that the quadrature rule $(\cdot)_N$ has precision $p \geq D$. Then,*

$$R^{D,N} = I^N. \tag{3.43}$$

Proof. The proof is a calculation using the definitions of $Q^{D,N}$ and $R^{D,N}$ in (3.4):

$$(R^{D,N})_{i,i} = \sum_{j=1}^{N^*} Q_{i,j}^{D,N} = \sum_{\ell=0}^{D} \sum_{k=-\ell}^{\ell} \hat{g}_\ell (m_{\ell,k})_N m_{\ell,k}(\Omega_i^N). \tag{3.44}$$

As a result of the quadrature precision,

$$(m_{\ell,k})_N = \int_{\mathbb{S}^2} m_{\ell,k}(\Omega)\, d\Omega = m_{0,0}^{-1} \delta_{\ell,0}. \tag{3.45}$$

Since $\hat{g}_0 = 1$, plugging (3.45) into (3.44) gives the desired result. □

Lemma 3.10 (Estimate of truncation error)**.** *Suppose that for some $r > 1$ and $s \geq 0$, $v \in H^r(\mathbb{S}^2)$ and*

$$\sum_{\ell=0}^{\infty} (\ell + 0.5)^{2s}\, \hat{g}_\ell^2 < \infty. \tag{3.46}$$

Then there exists a constant C depending on r and s such that

$$\|\mathcal{E}^N \mathcal{Q} v - \mathcal{E}^N \mathcal{Q}^D v\|_{\ell_N^2} \leq C(D+1)^{-(r+s-(1+\delta))} \|v\|_{H^r(\mathbb{S}^2)} \tag{3.47}$$

for any $\delta > 0$.

Proof. For any $u \in C^0(\mathbb{S}^2)$, it is straightforward to show that

$$\|\mathcal{E}^N u\|_{\ell_N^2} \leq \sqrt{4\pi} \|u\|_{L^\infty(\mathbb{S}^2)}. \tag{3.48}$$

Let $r > 1$. If $v \in H^r(\mathbb{S}^2)$, then $\mathcal{Q}v - \mathcal{Q}^D v \in H^r(\mathbb{S}^2) \subset C^0(\mathbb{S}^2)$ so that point-wise evaluation makes sense and $\mathcal{E}^N \mathcal{Q} v - \mathcal{E}^N \mathcal{Q}^D v$ is well-defined. Moreoever, (3.48) implies that

$$\|\mathcal{E}^N \mathcal{Q} v - \mathcal{E}^N \mathcal{Q}^D v\|_{\ell_N^2} \leq C \|\mathcal{Q} v - \mathcal{Q}^D v\|_{L^\infty(\mathbb{S}^2)} \leq C \|\mathcal{Q} v - \mathcal{Q}^D v\|_{H^{1+\delta}(\mathbb{S}^2)} \tag{3.49}$$

for any $\delta > 0$. A direct calculation using the norm definition in (3.29) gives

$$\|\mathcal{Q}v - \mathcal{Q}^D v\|^2_{H^{1+\delta}(\mathbb{S}^2)} = \sum_{\ell=D+1}^{\infty} \sum_{k=-\ell}^{\ell} (\ell + 0.5)^{2(1+\delta)} \hat{g}_\ell^2 \hat{v}_{\ell,k}^2. \tag{3.50}$$

From assumption (3.46), there exists a constant C such that $\hat{g}_\ell^2 \leq C\,(\ell + 0.5)^{-2s}$. Applying this bound to (3.50) gives

$$\begin{aligned} \|\mathcal{Q}v - \mathcal{Q}^D v\|^2_{H^{1+\delta}(\mathbb{S}^2)} &\leq C \sum_{\ell=D+1}^{\infty} \sum_{k=-\ell}^{\ell} (\ell + 0.5)^{-2(s-(1+\delta))} \hat{v}_{\ell,k}^2 \qquad (3.51) \\ &\leq C(D+1)^{-2(r+s-(1+\delta))} \sum_{\ell=D+1}^{\infty} \sum_{k=-\ell}^{\ell} (\ell + 0.5)^{2r} \hat{v}_{\ell,k}^2 \leq C(D+1)^{-2(r+s-(1+\delta))} \|v\|^2_{H^r(\mathbb{S}^2)}. \end{aligned}$$

Taking the square root of both sides gives the desired result. □

Lemma 3.11 (Estimate of quadrature error)**.** *Suppose for some $r > \frac{3}{2}$, $v \in H^r(\mathbb{S}^2)$, and*

$$\sum_{\ell=0}^{\infty} (\ell + 0.5)^2 \hat{g}_\ell^2 < \infty. \tag{3.52}$$

Let $D \geq 0$ and let the quadrature rule $(\cdot)_N$ have precision $p \geq 2D$. Then there exists a constant C depending on r and g such that

$$\|\mathcal{E}^N \mathcal{Q}^D v - Q^{D,N} \mathcal{E}^N v\|_{\ell_N^2} \leq C(p - D + 1)^{-(r - \frac{3+\delta}{2})} \|v\|_{H^r(\mathbb{S}^2)} \tag{3.53}$$

for any $\delta > 0$.

Proof. Let $D' = p - D + 1$. A direct calculation gives

$$\|\mathcal{E}^N \mathcal{Q}^D v - Q^{D,N} \mathcal{E}^N v\|^2_{\ell_N^2} = \sum_{\ell=0}^{D} \sum_{k=-\ell}^{\ell} \hat{g}_\ell^2 \left| \hat{v}_{\ell,k} - \hat{v}_{\ell,k}^N \right|^2, \tag{3.54}$$

where as a result of the quadrature precision,

$$\left| \hat{v}_{\ell,k} - \hat{v}_{\ell,k}^N \right|^2 = \left| \sum_{\ell'=D'}^{\infty} \sum_{k'=-\ell}^{\ell'} \hat{v}_{\ell',k'} \left(m_{\ell,k}, m_{\ell',k'} \right)_N \right|^2, \tag{3.55}$$

Using the Cauchy-Schwarz inequality, (3.48), and the bound in (2.14) gives

$$\begin{aligned}(m_{\ell,k}, m_{\ell',k'})_N \leq \|\mathcal{E}^N m_{\ell,k}\|_{\ell_N^2} \|\mathcal{E}^N m_{\ell',k'}\|_{\ell_N^2} &\leq 4\pi \|m_{\ell,k}\|_{L^\infty(\mathbb{S}^2)} \|m_{\ell',k'}\|_{L^\infty(\mathbb{S}^2)} \\ &\leq \sqrt{2\ell+1}\sqrt{2\ell'+1}.\end{aligned} \tag{3.56}$$

Substituting (3.56) into (3.55) gives

$$\left|\hat{v}_{\ell,k} - \hat{v}_{\ell,k}^N\right|^2 \leq (2\ell+1) \left| \sum_{\ell'=D'}^{\infty} \sum_{k'=-\ell'}^{\ell'} \hat{v}_{\ell',k'} \sqrt{2\ell'+1} \right|^2. \tag{3.57}$$

We factor the square root above as

$$\sqrt{2\ell'+1} = (2\ell'+1)^{r_1} (2\ell'+1)^{r_2}, \tag{3.58}$$

where $r_1 + r_2 = 0.5$. Then application of the Cauchy-Schwarz inequality gives

$$\begin{aligned}\sum_{\ell'=D'}^{\infty} \sum_{k'=-\ell'}^{\ell'} \hat{v}_{\ell',k'} \sqrt{2\ell'+1} &= \sum_{\ell'=D'}^{\infty} \sum_{k'=-\ell'}^{\ell'} \hat{v}_{\ell',k'} (2\ell'+1)^{r_1} (2\ell'+1)^{r_2} \\ &\leq \left(\sum_{\ell'=D'}^{\infty} \sum_{k'=-\ell'}^{\ell'} \hat{v}_{\ell',k'}^2 (2\ell'+1)^{2r_1} \right)^{1/2} \left(\sum_{\ell'=D'}^{\infty} (2\ell'+1)^{2r_2+1} \right)^{1/2}.\end{aligned} \tag{3.59}$$

In order to bound the second term above, we set $r_2 = -1 - \delta/2$ for some arbitrarily small $\delta > 0$. Substituting (3.59) into (3.57) with this value of r_2 gives

$$\left|\hat{v}_{\ell,k} - \hat{v}_{\ell,k}^N\right|^2 \leq C(2\ell+1) \sum_{\ell'=D'}^{\infty} \sum_{k'=-\ell'}^{\ell'} \hat{v}_{\ell',k'}^2 (2\ell'+1)^{3+\delta}. \tag{3.60}$$

Substituting (3.60) into (3.54) gives

$$\begin{aligned}\|\mathcal{E}^N \mathcal{Q}^D v - Q^{D,N} \mathcal{E}^N v\|_{\ell_N^2}^2 &\leq C \left(\sum_{\ell=0}^{D} \sum_{k=-\ell}^{\ell} (2\ell+1) \hat{g}_\ell^2 \right) \left(\sum_{\ell'=D'}^{\infty} \sum_{k'=-\ell'}^{\ell'} \hat{v}_{\ell',k'}^2 (2\ell'+1)^{3+\delta} \right) \\ &\leq C(D')^{-2\left(r - \frac{3+\delta}{2}\right)} \left(\sum_{\ell=0}^{D} (2\ell+1)^2 \hat{g}_\ell^2 \right) \left(\sum_{\ell'=D'}^{\infty} \sum_{k'=-\ell'}^{\ell'} \hat{v}_{\ell',k'}^2 (\ell'+0.5)^{2r} \right).\end{aligned} \tag{3.61}$$

The second sum above can be bounded by the $H^r(\mathbb{S}^2)$ norm of v (cf. (3.29)). Thus with the assumption in (3.52), (3.61) becomes

$$\begin{aligned}\|\mathcal{E}^N\mathcal{Q}^D v - Q^{D,N}\mathcal{E}^N v\|_{\ell^2_N}^2 &\le C(D')^{-2\left(r-\frac{3+\delta}{2}\right)}\sum_{\ell=0}^{D}(2\ell+1)^2\hat{g}_\ell^2\|v\|_{H^r(\mathbb{S}^2)}^2 \\ &\le C(D')^{-2\left(r-\frac{3+\delta}{2}\right)}\|v\|_{H^r(\mathbb{S}^2)}^2.\end{aligned} \tag{3.62}$$

Taking the square root on both sides gives the desired result. □

Lemma 3.12 (Estimate of filter error)**.** *Suppose the filter function f is of order α and satisfies Assumption 1. Let $v \in H^r(\mathbb{S}^2)$ for some $r > 1$, and suppose the quadrature rule $(\cdot)_N$ has precision $p \ge 2P$. Then*

$$\|F^{P,N}\mathcal{E}^N v\|_{\ell^2_N} \le \begin{cases} C(P+1)^{-r+\delta+3/2}\|v\|_{H^r(\mathbb{S}^2)}, & \alpha > r-\delta-3/2, \\ C(P+1)^{-\alpha+\epsilon}\|v\|_{H^r(\mathbb{S}^2)}, & \alpha \le r-\delta-3/2, \end{cases} \tag{3.63}$$

for any $\epsilon > 0$ and $\delta > 0$.

Proof. We show that

$$\|F^{P,N}\mathcal{E}^N v\|_{\ell^2_N}^2 \le C\sum_{\ell=1}^{P}(f_\ell^P)^2\ell^{-2(r-\delta-1)}\|v\|_{H^r(\mathbb{S}^2)}^2 \tag{3.64}$$

for some constant C and any $\delta > 0$. The result then follows from the proof of the filter error estimate for [22, Thm. 3.3], which shows that for all $q > 0$,

$$\sum_{\ell=1}^{P}(f_\ell^P)^2\ell^{-2q} \le C(P+1)^{1-\theta} \tag{3.65}$$

where $\theta = \min\{2q, 2\alpha+1-2\epsilon\}$ and $\epsilon > 0$ is arbitrary. The bound in (3.65) relies on Assumption 1; see [22] for details.

Since $v \in H^r(\mathbb{S}^2)$ with $r > 1$, $v \in C^0(\mathbb{S}^2)$. Using the definition of $F^{P,N}$ in (3.4) and (3.5), a direct calculation gives

$$\begin{aligned}\|F^{P,N}\mathcal{E}^N v\|_{\ell^2_N}^2 &= \sum_{i=1}^{N^*} w_i^N \left(\sum_{\ell=0}^{P}\sum_{k=-\ell}^{\ell} f_\ell^P m_{\ell,k}(\Omega_i^N)\tilde{v}_{\ell,k}^N\right)\left(\sum_{\ell'=0}^{P}\sum_{k'=-\ell'}^{\ell'} f_{\ell'}^P m_{\ell',k'}(\Omega_i^N)\tilde{v}_{\ell',k'}^N\right) \\ &= \sum_{\ell=0}^{P}\sum_{k=-\ell}^{\ell}\sum_{\ell'=0}^{P}\sum_{k'=-\ell'}^{\ell'} f_\ell^P f_{\ell'}^P \tilde{v}_{\ell,k}^N \tilde{v}_{\ell',k'}^N \left(m_{\ell,k}, m_{\ell',k'}\right)_N \\ &= \sum_{\ell=0}^{P}\sum_{k=-\ell}^{\ell} (f_\ell^P)^2 (\tilde{v}_{\ell,k}^N)^2.\end{aligned} \tag{3.66}$$

where the last line follows from the quadrature assumption, i.e., $\left(m_{\ell,k}, m_{\ell',k'}\right)_N = \delta_{\ell,\ell'}\delta_{k,k'}$ for all terms in the above sum. This term can be further decomposed as follows:

$$\begin{aligned}\sum_{\ell=0}^{P}\sum_{k=-\ell}^{\ell} (f_\ell^P)^2 (\tilde{v}_{\ell,k}^N)^2 &= \sum_{\ell=1}^{P} (f_\ell^P)^2 \|(\mathcal{I}^{N,\ell} - \mathcal{I}^{N,\ell-1})v\|_{L^2(\mathbb{S}^2)}^2 \\ &\le C\sum_{\ell=1}^{P} (f_\ell^P)^2 \left[\|(\mathcal{P}^\ell - \mathcal{P}^{\ell-1})v\|_{L^2(\mathbb{S}^2)}^2 + \|(\mathcal{I}^{N,\ell} - \mathcal{P}^\ell)v\|_{L^2(\mathbb{S}^2)}^2\right],\end{aligned} \tag{3.67}$$

where the $\ell = 0$ case is dropped since $f_0^P = 0$. Consequently, the error is exactly zero when $P = 0$. The last inequality above follows from two applications of the triangle inequality and the fact that

$$\|(\mathcal{I}^{N,\ell-1} - \mathcal{P}^{\ell-1})v\|_{L^2(\mathbb{S}^2)}^2 \le \|(\mathcal{I}^{N,\ell} - \mathcal{P}^\ell)v\|_{L^2(\mathbb{S}^2)}^2. \tag{3.68}$$

The first term in (3.67) satifies the bound

$$\|(\mathcal{P}^\ell - \mathcal{P}^{\ell-1})v\|_{L^2(\mathbb{S}^2)}^2 \le \|(\mathcal{I} - \mathcal{P}^{\ell-1})v\|_{L^2(\mathbb{S}^2)}^2 \le C\ell^{-2r}\|v\|_{H^r(\mathbb{S}^2)}^2, \tag{3.69}$$

where $\mathcal{I}$ is the identity operator and the second inequality is a standard result. (See, for example, [7, Eqn. 3.105].) For the second term in (3.67), Lemma 3.7 implies

$$\|(\mathcal{I}^{N,\ell} - \mathcal{P}^\ell)v\|_{L^2(\mathbb{S}^2)}^2 \le C\ell^{-2(r-\delta-1)}\|v\|_{H^r(\mathbb{S}^2)}^2 \tag{3.70}$$

for any positive δ. Substituting (3.69) and (3.70) into (3.67) leads to (3.64). Letting $q = r - \delta - 1$ concludes the proof. □

3.3 Numerical results

In the following numerical sections, we use norms based on the integral of the solutions in angle as proxies for the norm in Theorem 3.8. In Subsection 3.3.1, for a given uniform mesh $\mathcal{T}_h$ of domain X we use an approximation to the continuous L^2 norm for every $K \in \mathcal{T}_h$ to measure the error between the particle concentrations of the true solution $\phi(t, r)$ and the numerical solution $\Phi^{\Theta,h}(t, r)$ at a time t. Therefore

$$||\phi(t,\cdot) - \Phi^{\Theta,h}(t,\cdot)||_{L^2(X)} = \sqrt{\sum_{K\in\mathcal{T}_h} ||\phi(t,\cdot) - \Phi^{\Theta,h}(t,\cdot)||^2_{L^2(K)}}, \tag{3.71}$$

where $||\cdot||_{L^2(K)}$ is approximated with a high resolution quadrature.

In Subsection 3.3.2 the discrete L^2 and L^∞ norms used to measure the error between the numerical and true solutions at a time t are calculated for the cell-averaged particle concentration as follows. Given a uniform mesh $\mathcal{T}_h$ of domain X with cells $\mathcal{K}_{i,j}$ having size $|\mathcal{K}_{i,j}| = \Delta x \Delta y$ for all $1 \le i \le N_1$, $1 \le j \le N_2$, let $\mathbf{\Phi}^{\Theta,h}_{i,j}(t)$ be the numerically computed cell-averaged particle concentration on cell $\mathcal{K}_{i,j}$ at time t, whose calculation is dependent on the method used, and let

$$\boldsymbol{\phi}_{i,j}(t) = \frac{1}{\Delta x \Delta y} \int_{\mathcal{K}_{i,j}} \int_{\mathbb{S}^2} \psi(t, x, y, \Omega)\ d\Omega\ dx\ dy \tag{3.72}$$

be the exact cell-averaged particle concentration on cell $\mathcal{K}_{i,j}$. Then the discrete L^2 norm we use for a vector $\mathbf{\Phi}^{\Theta,h}(t)$ of numerically computed cell-averaged particle concentrations and a vector $\boldsymbol{\phi}(t)$ of cell-averaged particle concentrations of the exact solution is

$$\|\mathbf{\Phi}^{\Theta,h}(t) - \boldsymbol{\phi}(t)\|_{L^2(X)} = \left(\Delta x \Delta y \sum_{i=1}^{N_1}\sum_{j=1}^{N_2} |\mathbf{\Phi}^{\Theta,h}_{i,j}(t) - \boldsymbol{\phi}_{i,j}(t)|^2\right)^{1/2} \tag{3.73}$$

and the L^∞ norm is

$$\|\boldsymbol{\Phi}^{\Theta,h}(t) - \boldsymbol{\phi}(t)\|_{L^\infty(X)} = \max_{i,j} |\boldsymbol{\Phi}^{\Theta,h}_{i,j}(t) - \boldsymbol{\phi}_{i,j}(t)|. \tag{3.74}$$

Three different quadrature families are employed in the following numerical results: a product quadrature [6] with Gauss points and weights for the polar component and equally spaced points and equal weights for the azimuthal component, the Lebedev quadrature [51], and a standard two-dimensional triangular S_N quadrature [13, 64]. For product and Lebedev quadrature indexed by N, the precision is $2N-1$. For triangular quadrature indexed by $N \geq 2$, the precision is 3. The number of quadrature points associated with each quadrature family is given in Table 3.1, both in the full three-dimensional setting and in the reduced two-dimensional geoemetry in which we test the method. Solutions to problems using any of these quadratures with a particular set indexed by N are labeled as an S_N^D solution, where D is the truncation index in the expansion of the collision operator (see (2.24)). Likewise, if said solution is also filtered using a filter operator with expansion up to P as seen in (2.45), then those solutions are labeled as an $FS_{N,P}^D$ solution. If $D = 0$ in either case, then the index is suppressed, i.e. $S_N := S_N^0$, and $FS_{N,P} := FS_{N,P}^0$. This is the case when isotropic scattering is considered.

3.3.1 Filter order effect on convergence

In this section, we investigate the effect that the filter has on convergence order for smooth solutions. We will use the approximation to the continuous L^2 norm described in (3.71) for the calculations in this section. In absence of other errors, we expect the numerical solutions to converge with the order of the filter. In order to isolate the error due to the filter, we

Table 3.1: Scaling of quadrature points.

	Lebedev	Product	Triangular
3-D	$4N^2/3$	$2N^2$	$N(N+2)$
2-D	$2N^2/3$	N^2	$N(N+2)/2$

Number of quadrature points associated with an order N quadrature rule of each quadrature family. Numbers for the Lebedev family are asymptotic approximates.

consider isotropic scattering:

$$g(\mu_0) = \frac{1}{4\pi}, \quad \mu_0 \in [-1, 1], \tag{3.75}$$

and a smooth manufactured solution:

$$v(x, y, \Omega) = 900x^2y^2(1-x)^2(1-y)^2m_{1,1}^2(\Omega), \qquad m_{1,1}(\Omega) = \sqrt{\frac{3}{4\pi}}\Omega_x. \tag{3.76}$$

The constant phase function ensures that the truncation error in (3.40b) is zero, while a smooth solution ensures that regularity is not the limiting factor for the convergence order in (3.40c).

Let $X = (0, 1) \times (0, 1) \subset \mathbb{R}^2$. Then v satisfies the following boundary-value problem

$$\mathcal{L}\psi(x, y, \Omega) = \sigma_s(x, y)(\mathcal{Q}\psi)(x, y, \Omega) + S(x, y, \Omega) \quad (x, y, \Omega) \in \Gamma \tag{3.77a}$$

$$\psi(x, y, \Omega) = 0 \qquad\qquad\qquad\qquad\quad (x, y) \in \Gamma^-, \tag{3.77b}$$

where $S(x, y, \Omega) = \mathcal{L}v(x, y, \Omega) - \sigma_s(x, y)(\mathcal{Q}v)(x, y, \Omega)$ and $\sigma_t = \sigma_s = 1$.

We choose to simulate the solution of this equation with an $N = 16$ Lebedev quadrature with 185 ordinates. This quadrature is chosen to make the error in (3.40a) negligible with respect to the filter error (see Lemma 3.11). Additionally, a spatial resolution for a 2nd-order DG discretization is chosen fine enough to make the associated errors small: the spatial grid is 501×501 cells.

We use an exponential filter function of order α

$$f(\eta; \alpha) = \exp(\ln(\epsilon) \times \eta^\alpha), \tag{3.78}$$

where $\epsilon = 2^{-52}$ is the machine epsilon using double precision. Additionally, the filter strength is $\sigma_f = 1$ for all runs. In Table 3.2, the errors and convergence orders for two sets of filtered numerical solutions using filter orders of $\alpha = 2$ or $\alpha = 4$ are given.

We observe that the order of convergence is bounded by the order of the filter and increases as P increases. However it does not approach the filter order. We expect that if P

Table 3.2: Filter convergence test.

P	L^2 Error	Order
3	0.2650	-
5	0.2165	0.3962
7	0.1611	0.8783
9	0.1185	1.2224
11	0.0888	1.4347
13	0.0684	1.5671
15	0.0540	1.6544

(a) Second order filter

P	L^2 Error	Order
3	0.1889	-
5	0.0466	2.7407
7	0.0148	3.3997
9	0.0061	3.5320
11	0.0029	3.6319
13	0.0016	3.6931
15	0.0009	3.7359

(b) Fourth order filter

L^2 errors and order of convergence for numerical solutions $\mathrm{FS}_{16,P}$ to a steady-state problem using a manufactured solution. The filter strength for all runs is $\sigma_\mathrm{f} = 1$.

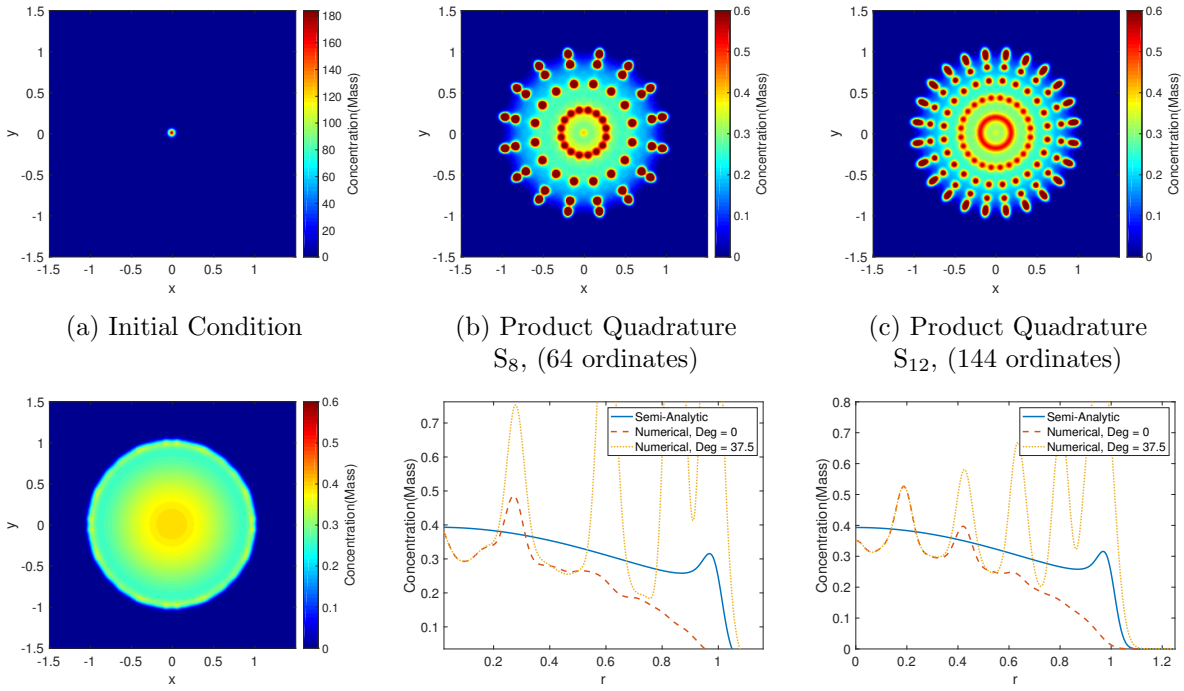
was increased further then it would approach the filter order. Due to limitations of resources we cannot confirm this conjecture.

3.3.2 Line source

Isotropic scattering

In this section we simulate the line source problem defined in Subsection 2.5.1 using a constant phase function defined in (3.75). We simulate the problem with a 301×301 grid on domain $[-1.5, 1.5] \times [-1.5, 1.5]$. We impose a zero boundary condition and choose the standard deviation in (2.53) to be $\beta = 0.03$. The time step is $\Delta t = 0.9\Delta x$. Although the implicit time-stepping algorithm allows for a much larger stable time step, we have observed that the filter does not perform well in this problem when exceeding the explicit hyperbolic time-step.

Figure 3.1(a) shows the cell-averaged particle concentration, $\boldsymbol{\phi}(0)$, for the initial condition (2.53). The semi-analytic solution to the line source problem with the aforementioned data is shown in Figure 3.1(d) and is computed with the code used in [24]. The remaining plots in Figure 3.1 show two (non-filtered) S_N solutions and their associated radial line-outs, directed along two different angles that are measured with respect to the x-axis. These line-outs



(a) Initial Condition

(b) Product Quadrature S_8, (64 ordinates)

(c) Product Quadrature S_{12}, (144 ordinates)

(d) Semi-analytic

(e) Product Quadrature S_8, (64 ordinates)

(f) Product Quadrature S_{12}, (144 ordinates)

Figure 3.1: Line-source solutions with line-outs.

Cell-averaged particle concentrations for the isotropic line-source problem on a 301×301 grid. Graphs (b)-(f) show solutions at $t = 1$. Graphs (b) and (c) demonstrate the presence of ray-effects in non-filtered discrete ordinates numerical solutions. Graphs (e) and (f) show plots along two angles: the x-axis, represented as Deg $= 0$, and in one of the direction where ray-effects appear in Graphs (b) and (c), represented as Deg $= 37.5$

clearly demonstrate the lack of rotational invariance and the occurrence of ray-effects.

Figures 3.2 and 3.3 show results using different quadratures with the filter while varying the filter index P in the filter operator. We observe that P must be chosen carefully with regards to N: if P is significantly different from N, then qualitatively poor solutions are produced as seen in Figures 3.2(a)-(c) and Figures 3.2(g)-(i); if P is approximately N, both the product and Lebedev quadratures give reasonable results as seen in Figures 3.2(d) and (e). This seems to be related to both having enough moments in the filter term to resolve the ray-effects, and having enough precision in the quadrature to accurately calculate those moments. This is demonstrated more clearly in Figures 3.3(d) and (e) where $P = N - 1$ resulting in numerical solutions that are more radially symmetric. The triangular quadrature solution 3.2(f) still shows non-physical defects even when $P \approx N$. This may be due to the triangular quadrature never having the necessary precision to calculate the higher degree moments.

Anisotropic scattering

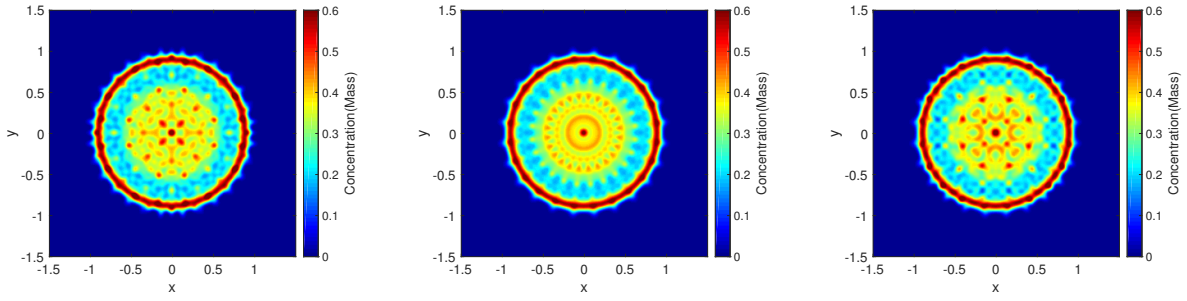
In this section we repeat the line source benchmark with a non-constant phase function. We use the Henyey-Greenstein phase function [31]:

$$g(\mu_0; \eta) = \frac{1 - \eta^2}{4\pi(1 + \eta^2 - 2\eta\mu_0)^{1.5}}, \quad \eta \in (-1, 1), \tag{3.79}$$

where η is called the anisotropy factor. With this choice of phase function

$$\hat{g}_\ell(\eta) = 2\pi \int_{-1}^{1} g(\mu_0; \eta) P_\ell(\mu_0)\, d\mu_0 = \eta^\ell. \tag{3.80}$$

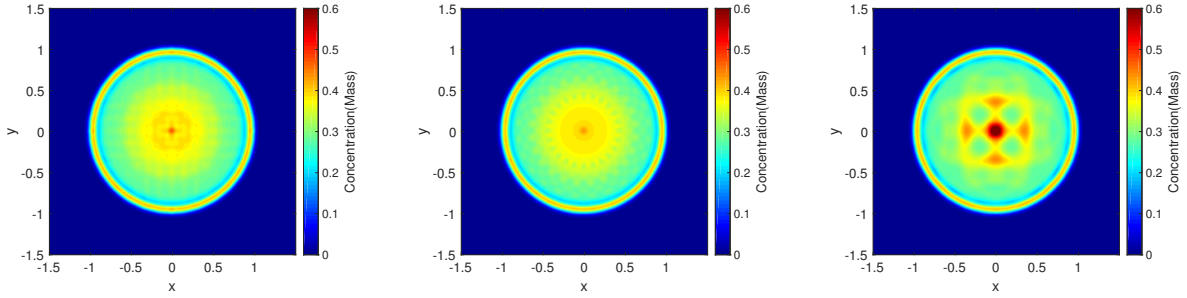
We perform simulations for two different anisotropy values: $\eta = 0.2$ and $\eta = 0.5$. The data for these simulations is given in Table 3.3. The line-out in Figure 3.4(c) shows how the solution changes when the initial condition and anisotropy changes.



(a) Lebedev Quadrature $\mathrm{FS}_{12,5}$, (105 ordinates, 21 moments)

(b) Product Quadrature $\mathrm{FS}_{12,5}$, (144 ordinates, 21 moments)

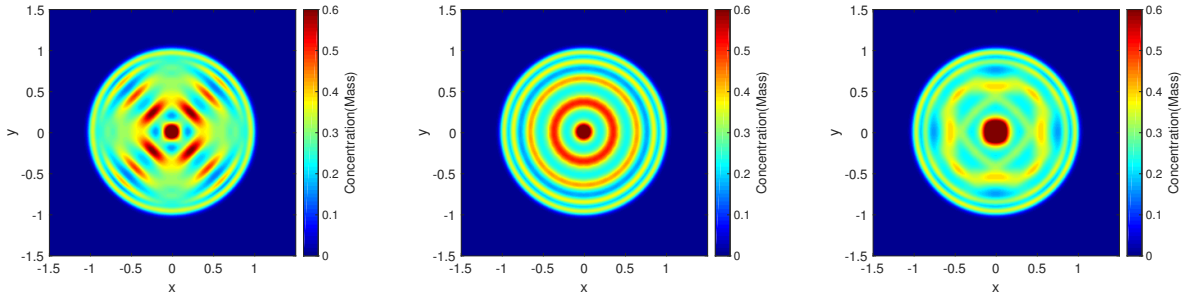
(c) Triangular Quadrature $\mathrm{FS}_{12,5}$, (84 ordinates, 21 moments)



(d) Lebedev Quadrature $\mathrm{FS}_{12,11}$, (105 ordinates, 78 moments)

(e) Product Quadrature $\mathrm{FS}_{12,11}$, (144 ordinates, 78 moments)

(f) Triangular Quadrature $\mathrm{FS}_{12,11}$, (84 ordinates, 78 moments)



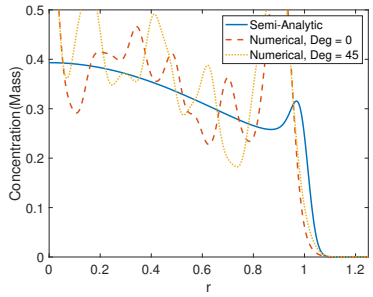
(g) Lebedev Quadrature $\mathrm{FS}_{12,15}$, (105 ordinates, 136 moments)

(h) Product Quadrature $\mathrm{FS}_{12,15}$, (144 ordinates, 136 moments)

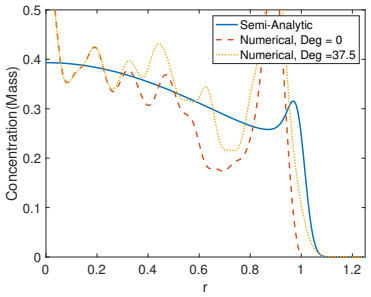
(i) Triangular Quadrature $\mathrm{FS}_{12,15}$, (84 ordinates, 136 moments)

Figure 3.2: Line-source $\mathrm{FS}_{N,P}$ solutions.

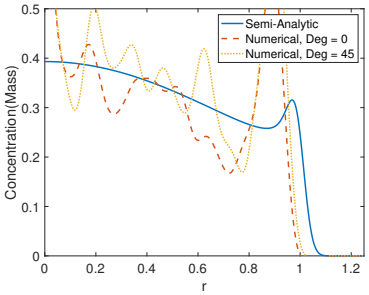
Cell-averaged particle concentrations for the isotropic line-source problem when $t = 1$ on a 301×301 grid. Numerical solutions are computed with order $N = 12$ quadrature rules of three different types and a varying number of moments (P) in the filter. Numerical solutions with $P \approx N$ give qualitatively reasonable results, with the exception of triangular quadrature solutions.



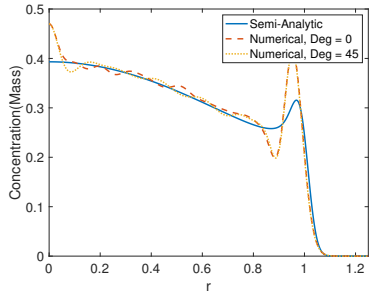
(a) Lebedev Quadrature $FS_{12,5}$, (105 ordinates, 21 moments)



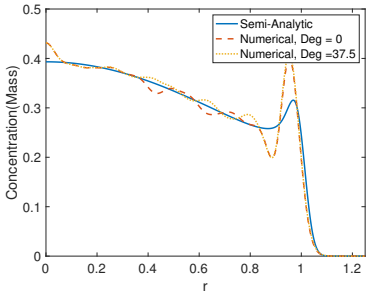
(b) Product Quadrature $FS_{12,5}$, (144 ordinates, 21 moments)



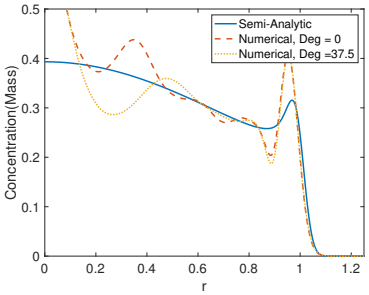
(c) Triangular Quadrature $FS_{12,5}$, (84 ordinates, 21 moments)



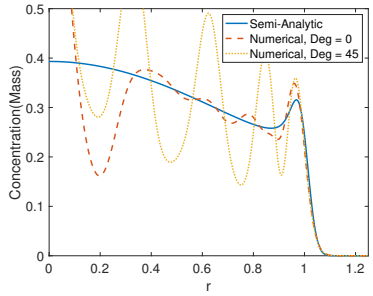
(d) Lebedev Quadrature $FS_{12,11}$, (105 ordinates, 78 moments)



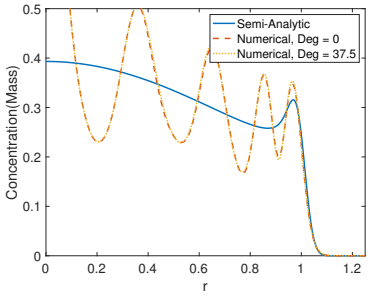
(e) Product Quadrature $FS_{12,11}$, (144 ordinates, 78 moments)



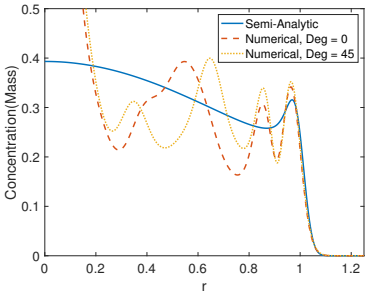
(f) Triangular Quadrature $FS_{12,11}$, (84 ordinates, 78 moments)



(g) Lebedev Quadrature $FS_{12,15}$, (105 ordinates, 136 moments)



(h) Product Quadrature $FS_{12,15}$, (144 ordinates, 136 moments)



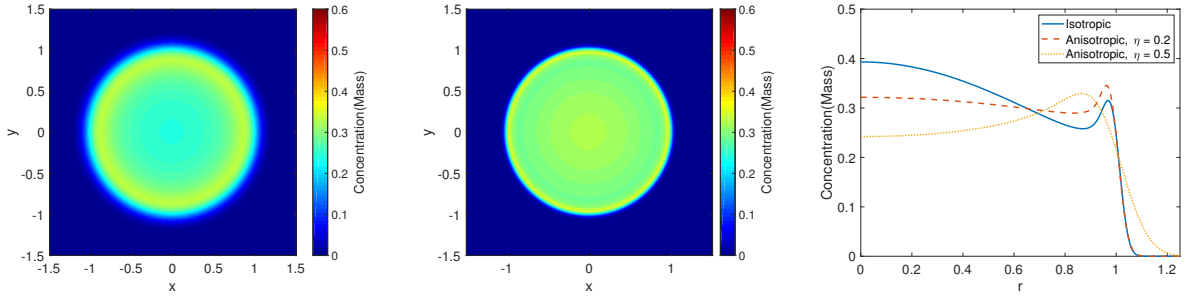
(i) Triangular Quadrature $FS_{12,15}$, (84 ordinates, 136 moments)

Figure 3.3: Radial line-outs of the solutions in Figure 3.2.

Line-outs show how rotationally invariant the numerical solutions are and how close they are to the semi-analytic solution. The line-outs show plots along two angles: the x-axis, represented as Deg = 0, and in one of the direction where ray-effects appear in the corresponding graph in Figure 3.2, represented as Deg = 45 or Deg = 37.5. Graphs (d) and (e) show a more radially symmetric solution and relatively good accuracy. Although graph (h) is nearly rotationally invariant, the additional oscillations reduces accuracy.

Table 3.3: Data for line-source problem with anisotropic scattering.

η	β	spatial mesh
0.2	0.03	504×504
0.5	0.09	505×505



(a) Reference Solution, $\eta = 0.5$ Lebedev S_{54}^{10}, 1981 ordinates

(b) Reference Solution, $\eta = 0.2$ Lebedev S_{66}^{20}, 2949 ordinates

(c) Line-out comparision

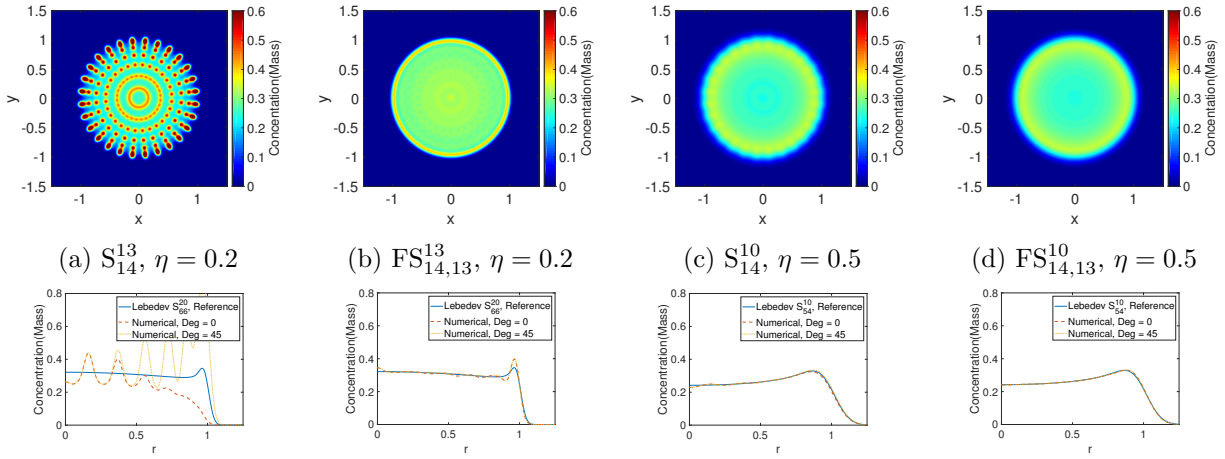
Figure 3.4: Reference solutions for line-source problem with anisotropic scattering.

Graphs (a) and (b) are generated using Lebedev quadrature S_{66} and S_{54} respectively with corresponding line-outs shown in (c). Solutions are at $t = 1$.

In Figure 3.5, we present lower resolution in angle simulations. We observe in Figure 3.5(b) that the filter is effective at reducing the occurrence of ray-effects when compared to Figure 3.5(a). We observe in Figure 3.5(c) that even when the standard deviation in the initial condition is increased to 0.09 ray-effects are still slightly apparent; in Figure 3.5(d) the filter is still effective at reducing even minor fluctuations.

Efficiency

In this section we explore the efficiency of the filter to the line source problem. For isotropic simulations, a filtered numerical solution will have a greater computational cost than its non-filtered equivalent. This is because the iterative solver takes up most of the computational time in the simulation, and it increases with the number of moments in the expansion of the scattering source term and filter term (see (3.7)). Since non-filtered simulations with isotropic scattering have only one moment, the solution process is relatively quick and may result in a more efficient method. However, an advantage of the filtered simulations is that they require fewer ordinates to get a comparable solution. This will result in smaller memory requirement



(a) S_{14}^{13}, $\eta = 0.2$ (b) $FS_{14,13}^{13}$, $\eta = 0.2$ (c) S_{14}^{10}, $\eta = 0.5$ (d) $FS_{14,13}^{10}$, $\eta = 0.5$

(e) S_{14}^{13}, $\eta = 0.2$ (f) $FS_{14,13}^{13}$, $\eta = 0.2$ (g) S_{14}^{10}, $\eta = 0.5$ (h) $FS_{14,13}^{10}$, $\eta = 0.5$

Figure 3.5: Numerical solutions to line-source problem with anisotropic scattering.

Solutions use product quadratures. Simulations are run to time $t = 1$. Top row: cell-averaged particle concentration. Bottom row: corresponding line-outs. The line-outs show plots along two angles.

for holding the solution at each time step, but may still require additional memory to store the iterated values in each Krylov solve.

We examine the relative errors of the cell-average particle concentration at $t = 1$, defined by

$$L^p \text{ Error} = \frac{\|\boldsymbol{\Phi}^{\Theta,h}(1) - \boldsymbol{\phi}(1)\|_{L^p(X)}}{\|\boldsymbol{\phi}(1)\|_{L^p(X)}}, \tag{3.81}$$

where $||\cdot||_{L^p(X)}$ is defined in (3.73) and (3.74) for $p = 2$ and $p = \infty$ respectively. Define the efficiency by

$$L^p \text{ Efficiency} = L^p \text{ Error} \times \text{Run Time}. \tag{3.82}$$

Thus a smaller value means that the method is more efficient. Table 3.4 shows the relative L^2 errors and computational time as well as the efficiency of the method.

We observe that $\text{FS}_{N,N-1}$ yields smaller errors than its non-filtered S_N counterpart. However a better result can be obtained more efficiently by increasing the number of ordinates in the non-filtered method. Table 3.5 lists the solutions from each respective category with the fastest computational time that meets the given tolerance. The most efficient method in each column is shown in bold.

Table 3.5 shows that in terms of effiecincy, the standard runs are able to achieve better results for the isotropic line source problem in most cases.

In Table 3.6, we show the efficiency of the filter with respect to the L^2 and L^∞ relative errors when applied to the anisotropic line source benchmark with anistropy factor $\eta = 0.2$ and the product quadrature. Table 3.7 is generated similarly and suggests that in every case the filtered methods are more efficient.

In Table 3.8, we report the results of a similar test with $\eta = 0.5$. For this experiment we use the same expansion in the collision operator for all simulations, truncating the expansion at $D = 10$.

We observe that in some situations, it is possible for the filtered run to complete faster than its non-filtered equivalent. This may be due to the iterative solver being able to converge faster when a filter is applied. From Tables 3.7 and 3.9 we observe that the filtered runs are much more efficient than the non-filtered runs in almost all cases.

Table 3.4: Line source (isotropic) results.

Method	Time	L^2 Error	L^2 Efficiency
S_4	4.7700	1.8520	8.8340
S_6	9.2200	1.3130	12.1058
S_8	10.0900	0.9689	9.7762
S_{10}	10.3800	0.7377	7.6573
S_{12}	10.1500	0.5656	5.7408
S_{14}	11.1600	0.5326	5.9438
S_{16}	14.2100	0.4182	5.9426
S_{18}	15.6800	0.2897	4.5425
S_{20}	22.7100	0.2062	4.6828
S_{22}	21.7600	0.1449	3.1530
S_{26}	28.1200	0.1011	2.8429
S_{30}	31.9300	0.0710	2.2670
$FS_{4,3}$	7.9970	0.9292	7.4308
$FS_{6,5}$	13.2478	0.4405	5.8357
$FS_{8,7}$	18.8374	0.2718	5.1200
$FS_{10,9}$	27.9849	0.1687	4.7211
$FS_{12,11}$	41.9300	0.1064	4.4613
$FS_{14,13}$	49.3566	0.0689	3.4007

(a) Lebedev Quadrature

Method	Time	L^2 Error	L^2 Efficiency
S_4	5.7700	1.8305	10.5620
S_6	8.8100	1.1708	10.3147
S_8	7.9200	0.8368	6.6275
S_{10}	14.0900	0.6512	9.1754
S_{12}	13.7600	0.5270	7.2515
S_{14}	16.8200	0.4426	7.4445
S_{16}	19.1200	0.3654	6.9864
S_{18}	22.2800	0.3063	6.8244
S_{20}	22.8900	0.2537	5.8072
S_{22}	26.1200	0.2113	5.5192
S_{26}	34.4300	0.1378	4.7445
S_{30}	45.3700	0.0907	4.1151
$FS_{4,3}$	9.8364	0.7232	7.1137
$FS_{6,5}$	13.3181	0.4251	5.6615
$FS_{8,7}$	19.8163	0.2693	5.3365
$FS_{10,9}$	35.7738	0.1696	6.0672
$FS_{12,11}$	39.9741	0.1063	4.2492
$FS_{14,13}$	56.8318	0.0685	3.8930

(b) Product Quadrature

Run times (in minutes), relative error, and efficiency for the line source problem for various filtered and non-filtered runs with isotropic scattering and solution time t=1.0. Results suggest that adding more ordinates is computationally more efficient than filtering with fewer ordinates.

Table 3.5: Efficiency comparison (isotropic).

	Tolerance					
Angular Method	1.0	0.8	0.6	0.4	0.2	0.1
S_N (Lebedev)	$S_8(9.8, 49, 1)$	$S_{10}(7.7, 77, 1)$	$S_{12}(5.7, 105, 1)$	$\mathbf{S_{18}}(4.5, 229, 1)$	$\mathbf{S_{22}}(3.2, 401, 1)$	$\mathbf{S_{30}}(2.3, 621, 1)$
$FS_{N,N-1}$ (Lebedev)	$FS_{4,3}(7.4, 17, 10)$	$\mathbf{FS_{6,5}}(5.8, 29, 21)$	$FS_{6,5}(5.8, 29, 21)$	$FS_{8,7}(5.1, 49, 36)$	$FS_{10,9}(4.7, 77, 55)$	$FS_{14,13}(3.4, 141, 105)$
S_N (Product)	$\mathbf{S_8}(6.6, 64, 1)$	$S_{10}(9.2, 100, 1)$	$S_{12}(7.3, 144, 1)$	$S_{16}(7.0, 256, 1)$	$S_{26}(4.7, 676, 1)$	$S_{30}(4.1, 900, 1)$
$FS_{N,N-1}$ (Product)	$FS_{4,3}(7.1, 16, 10)$	$FS_{4,3}(7.1, 16, 10)$	$\mathbf{FS_{6,5}}(5.66, 36, 21)$	$FS_{8,7}(5.3, 64, 36)$	$FS_{10,9}(6.1, 100, 55)$	$FS_{14,13}(3.9, 196, 105)$

Efficiency comparison between filtered and non-filtered solutions with error meets the given tolerance. Each row corresponds to a particular type of angular discretization. Each column corresponds to an error tolerance. Values in parentheses are (Efficiency, number of ordinates (N^*), number of moments (M^*)). Bold indicates the most efficient method (in terms of run time) for given tolerance.

Table 3.6: Line source (anisotropic) results.

Method	Time	L^2 Error	L^∞ Error	L^2 Efficiency	L^∞ Efficiency
S_4^3	32.5650	1.9614	17.7025	63.8730	576.4819
S_6^5	60.3333	1.2217	7.4867	73.7092	451.6973
S_8^7	88.6500	0.8672	3.9349	76.8773	348.8289
S_{10}^9	131.5500	0.6721	2.3562	88.4148	309.9581
S_{14}^{13}	267.9500	0.4537	1.5284	121.5689	409.5348
S_{16}^{15}	408.9833	0.3780	1.2231	154.5957	500.2275
$FS_{4,3}^3$	31.5833	0.7811	4.5019	24.6697	142.1849
$FS_{6,5}^5$	66.2333	0.4525	0.9529	29.9706	63.1137
$FS_{8,7}^7$	103.8833	0.2750	0.7060	28.5679	73.3416
$FS_{10,9}^9$	146.3167	0.1638	0.4499	23.9667	65.8279
$FS_{14,13}^{13}$	325.3667	0.0545	0.1623	17.7325	52.8070
$FS_{16,15}^{15}$	666.7833	0.0306	0.0935	20.4036	62.3442

Run times (in minutes), relative error, and efficiency for the line source problem for various filtered and non-filtered runs with anisotropic scattering, $\eta = 0.2$, and final time t=1.0. Results suggest that filtering with fewer ordinates is more efficient than adding ordinates without filtering.

Table 3.7: Efficiency comparision (anisotropic).

	Tolerance				
Angular Method	1.0	0.8	0.6	0.4	0.2
Product S_N^{N-1}	$S_8^7(77, 64, 36)$	$S_{10}^9(88, 100, 55)$	$S_{14}^{13}(122, 196, 105)$	$S_{16}^{15}(155, 256, 136)$	— —
Product $FS_{N,N-1}^{N-1}$	$\mathbf{FS_{4,3}^3}(25, 16, 10)$	$\mathbf{FS_{4,3}^3}(25, 16, 10)$	$\mathbf{FS_{6,5}^5}(30, 36, 21)$	$\mathbf{FS_{8,7}^7}(29, 64, 36)$	$\mathbf{FS_{10,9}^9}(24, 100, 55)$

Efficiency comparison between filtered and non-filtered solutions when error meets the given tolerance. Each row corresponds to a particular type of angular discretization. Each column corresponds to an error tolerance. Values in parentheses are (Efficiency, number of ordinates (N^*), number of moments (M^*)). Bold indicates the most efficient method for the given tolerance. Missing data in tolerance 0.2 column shows that a non-filtered method with many more ordinates would have been required to achieve the given tolerance.

Table 3.8: Line source (anisotropic, $D = 10$) results.

Method	Time	L^2 Error	L^∞ Error	L^2 Efficiency	L^∞ Efficiency
S_4^{10}	61.5068	0.7106	2.1486	43.7067	132.1535
S_6^{10}	77.0854	0.3509	0.6765	27.0493	52.1483
S_8^{10}	91.5962	0.1844	0.3558	16.8903	32.5899
S_{10}^{10}	121.4340	0.0889	0.1756	10.7955	21.3238
S_{12}^{10}	132.1840	0.0388	0.0858	5.1287	11.3414
$FS_{4,3}^{10}$	66.9561	0.5445	0.9841	36.4576	65.8915
$FS_{6,5}^{10}$	92.9296	0.1849	0.3280	17.1827	30.4809
$FS_{8,7}^{10}$	102.0970	0.0643	0.1197	6.5648	12.2210
$FS_{10,9}^{10}$	111.8940	0.0232	0.0445	2.5959	4.9793
$FS_{12,11}^{10}$	150.6420	0.0093	0.0181	1.4010	2.7266

Run times (in minutes), relative error, and efficiency for the line source problem for various filtered and non-filtered runs with anisotropic scattering, $\eta = 0.5$, and final time t=1.0.

Table 3.9: Efficiency comparison (anisotropic, $D = 10$).

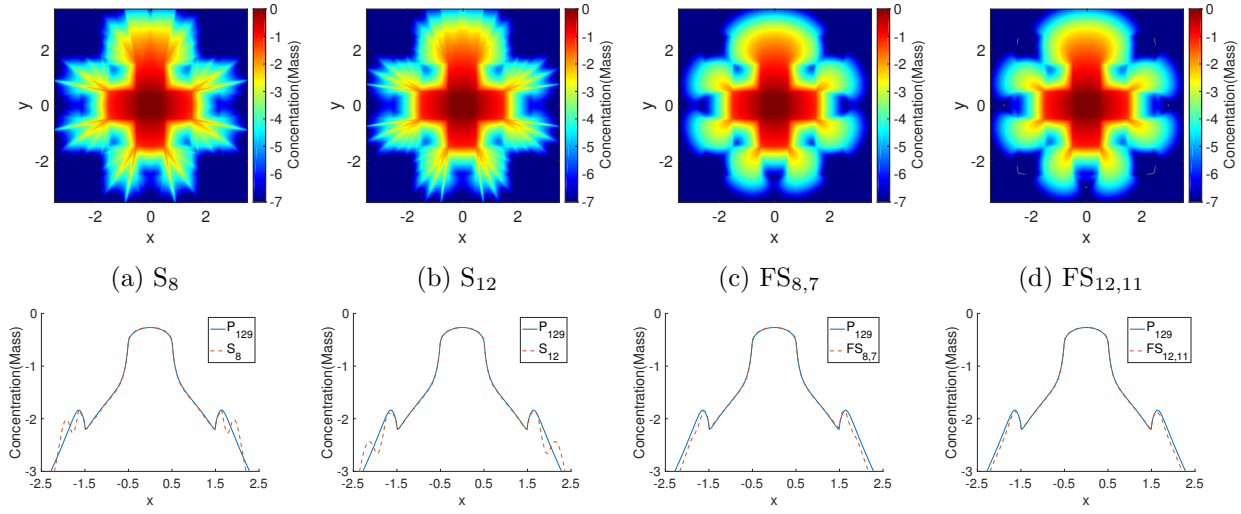
	Tolerance				
Angular Method	0.8	0.4	0.2	0.1	0.05
Product S_N^{10}	$S_4^{10}(44, 16, 66)$	$S_6^{10}(27, 36, 66)$	$\mathbf{S_8^{10}}(16.9, 64, 66)$	$S_{10}^{10}(11, 100, 66)$	$S_{12}^{10}(5.1, 144, 66)$
Product $FS_{N,N-1}^{10}$	$\mathbf{FS_{4,3}^{10}}(36, 16, 66)$	$\mathbf{FS_{6,5}^{10}}(17.2, 36, 66)$	$FS_{6,5}^{10}(17.2, 36, 66)$	$\mathbf{FS_{8,7}^{10}}(6.6, 64, 66)$	$\mathbf{FS_{10,9}^{10}}(2.6, 100, 66)$

Efficiency comparison between filtered and non-filtered solutions with error meets the given tolerance. Each row corresponds to a particular type of angular discretization. Each column corresponds to an error tolerance. Values in parentheses are (Efficiency, number of ordinates (N^*), number of moments (M^*)). Bold indicates the most efficient method for given tolerance.

3.3.3 Lattice

In the section we use the more real world lattice benchmark, (see Subsection 2.5.2), to demonstrate the effectiveness of the filtered discrete ordinates equations. We use a 504×504 spatial grid with domain $[-3.5, 3.5] \times [-3.5, 3.5]$, set $\Delta t = 10\Delta x$, and run to a final time $t = 2.8$. Unlike the line source, for this problem we are able to take advantage of the implicit discretization in time with a relatively large time step, without destroying the benefits of the filter. The filter strength and filter function are the same as for the the line source tests; see (2.54).

In Figures 3.6 (a) and (b), we have run the lattice test using a S_8 and S_{12} discrete ordinates method respectively. The line-outs in Figures 3.6 (e) and (f) show a cross-section of the solutions in Figures 3.6 (a) and (b) at the line $y = 0.6$. We observe in both cases that ray-effects are very apparent. When the test is run with a 4th-order filter as in Figures 3.6 (c), (d) we observe that the occurrence of ray-effects have been reduced significantly.



(a) S_8 (b) S_{12} (c) $FS_{8,7}$ (d) $FS_{12,11}$

(e) S_8 (f) S_{12} (g) $FS_{8,7}$ (h) $FS_{12,11}$

Figure 3.6: Non-filtered and filtered solutions of the lattice problem

Cell-averaged particle concentrations for lattice problem when time $t = 2.8$ using a logarithmic scale. Simulations are run on a 504×504 grid with a 4th-order filter. Line-outs are taken at the line $y = 0.6$.

Chapter 4

Hybrid solver for the radiative transport equation using finite volume and discontinuous Galerkin

In the following sections, we introduce and analyze a hybrid spatial discretization for the RTE. The underlying formulation is based on the idea of first collision source [3] and has been used in [59] as a way to combine different angular discretizations in a fully implicit time integration scheme. Here we combine DG discretization, which performs well in the diffusion limit, with FV discretization, which uses less memory per computational cell, into a single discretization strategy.

4.1 General hybrid formulation

The basic idea of first collision source is to separate ψ into the sum $\psi = \psi_{\mathrm{u}} + \psi_{\mathrm{c}}$ where the uncollided flux ψ_{u} and the collided flux ψ_{c} satisfy

$$(\epsilon\partial_t + \mathcal{L}_\epsilon)\psi_{\mathrm{u}} = \epsilon S, \tag{4.1a}$$

$$(\epsilon\partial_t + \mathcal{L}_\epsilon)\psi_{\mathrm{c}} = \mathcal{Q}_{\mathrm{s},\epsilon}\left(\psi_{\mathrm{u}} + \psi_{\mathrm{c}}\right). \tag{4.1b}$$

Equation (4.1) is an exact splitting of (2.48). The hybrid method evolves these equations for a time-step and then re-initalizes the values ψ_{u} and ψ_{c}. For illustration, with a backward Euler discretization these steps (in reverse order) give

$$\psi_{\mathrm{u}}^{n+1/2} = \psi_{\mathrm{u}}^{n} + \psi_{\mathrm{c}}^{n} \quad \text{and} \quad \psi_{\mathrm{c}}^{n+1/2} = 0 \tag{4.2}$$

and then

$$\epsilon \frac{\psi_{\mathrm{u}}^{n+1} - \psi_{\mathrm{u}}^{n+1/2}}{\Delta t} + \mathcal{L}_{\epsilon} \psi_{\mathrm{u}}^{n+1} = \epsilon S^{n+1}, \tag{4.3a}$$

$$\epsilon \frac{\psi_{\mathrm{c}}^{n+1} - \psi_{\mathrm{c}}^{n+1/2}}{\Delta t} + \mathcal{L}_{\epsilon} \psi_{\mathrm{c}}^{n+1} = \mathcal{Q}_{\mathrm{s},\epsilon} \left(\psi_{\mathrm{u}}^{n+1} + \psi_{\mathrm{c}}^{n+1} \right), \tag{4.3b}$$

where $S^{n+1} = S^{n+1}(r, \Omega) = S(t^{n+1}, r, \Omega)$. In practice, ψ_{u} is often approximated by a high-order angular discretization, while ψ_{c} uses a low-order discretization. Thus the point of the reinitialization step (4.2) is to maintain the benefits of the high-resolution discretization.

The reinitialization step (4.2) can be substituted into (4.3) to obtain a closed rule for updating ψ_{u}^{n} and ψ_{c}^{n}:

$$\mathcal{L}_{\epsilon}^{\Delta t} \psi_{\mathrm{u}}^{n+1} = \epsilon \left(\frac{\psi_{\mathrm{u}}^{n} + \psi_{\mathrm{c}}^{n}}{\Delta t} + S^{n+1} \right), \tag{4.4a}$$

$$\mathcal{L}_{\epsilon}^{\Delta t} \psi_{\mathrm{c}}^{n+1} = \mathcal{Q}_{\mathrm{s},\epsilon} \left(\psi_{\mathrm{u}}^{n+1} + \psi_{\mathrm{c}}^{n+1} \right), \tag{4.4b}$$

where

$$\mathcal{L}_{\epsilon}^{\Delta t} = \Omega \cdot \nabla_r + \frac{\sigma_{\mathrm{t}}(r)}{\epsilon} + \frac{\epsilon}{\Delta t}. \tag{4.5}$$

Although (4.4) is not exactly a backward Euler discretization of (4.1), adding (4.4a) and (4.4b) together does yield a backward Euler discretization of (2.48a):

$$\mathcal{L}_{\epsilon}^{\Delta t} \psi^{n+1} = \mathcal{Q}_{\mathrm{s},\epsilon} \psi^{n+1} + \epsilon \left(\frac{1}{\Delta t} \psi^{n} + S^{n+1} \right), \tag{4.6}$$

where $\psi^{n} = \psi_{\mathrm{u}}^{n} + \psi_{\mathrm{c}}^{n}$ and $\psi^{n+1} = \psi_{\mathrm{u}}^{n+1} + \psi_{\mathrm{c}}^{n+1}$. However, the hybrid strategy is to approximate (4.4a) and (4.4b) in angle and space using different discretizations, in which case the approximations for ψ_{u}^{n+1} and ψ_{c}^{n+1} cannot be added directly.

For the remainder of this section we fix n and Δt and focus on finding a hybrid approximation for ψ^{n+1}. Since it is implicit, this update can be reformulated as a steady-state problem. Let V^{u} and V^{c} be two finite dimensional vector spaces and let $f^{\mathrm{u}} \in V^{\mathrm{u}}$ and $f^{\mathrm{c}} \in V^{\mathrm{c}}$ satisfy

$$L_\epsilon^{\mathrm{u}} f^{\mathrm{u}} = \epsilon S, \tag{4.7a}$$

$$L_\epsilon^{\mathrm{c}} f^{\mathrm{c}} = Q_{\mathrm{s},\epsilon}^{\mathrm{c,u}} f^{\mathrm{u}} + Q_{\mathrm{s},\epsilon}^{\mathrm{c,c}} f^{\mathrm{c}}. \tag{4.7b}$$

Here $S \in V^{\mathrm{u}}$ is an approximation of $\left(\frac{1}{\Delta t}\psi^n + S^{n+1}\right)$; $Q_{\mathrm{s},\epsilon}^{\mathrm{c,u}} : V^{\mathrm{u}} \to V^{\mathrm{c}}$ and $Q_{\mathrm{s},\epsilon}^{\mathrm{c,c}} : V^{\mathrm{c}} \to V^{\mathrm{c}}$ are both approximations of $\mathcal{Q}_{\mathrm{s},\epsilon}$; and $L_\epsilon^{\mathrm{u}} : V^{\mathrm{u}} \to V^{\mathrm{u}}$ and $L_\epsilon^{\mathrm{c}} : V^{\mathrm{c}} \to V^{\mathrm{c}}$ are both approximations of $\mathcal{L}_\epsilon^{\Delta t}$.

The next step is to compute an approximation $f \in V^{\mathrm{u}}$ of ψ^{n+1} from f^{u} and f^{c}. The strategy in [18, 59] is to let $f = f^{\mathrm{u}} + \mathcal{R} f^{\mathrm{c}}$, where $\mathcal{R} : V^{\mathrm{c}} \to V^{\mathrm{u}}$ is a "relabeling operator". Here, we instead follow [17] and solve the following approximation of (4.6):

$$L_\epsilon^{\mathrm{u}} f = \left(Q_{\mathrm{s},\epsilon}^{\mathrm{u,u}} f^{\mathrm{u}} + Q_{\mathrm{s},\epsilon}^{\mathrm{u,c}} f^{\mathrm{c}}\right) + \epsilon S, \tag{4.8}$$

where $Q_{\mathrm{s},\epsilon}^{\mathrm{u,u}} \colon V^{\mathrm{u}} \to V^{\mathrm{u}}$ and $Q_{\mathrm{s},\epsilon}^{\mathrm{u,c}} \colon V^{\mathrm{c}} \to V^{\mathrm{u}}$ both approximate $\mathcal{Q}_{\mathrm{s},\epsilon}$.

While the formulation above uses backward Euler for the temporal discretization, we use a second-order diagonally implicit Runge-Kutta method [4] for the time dependent numerical experiments in Section 4.4. However, like backward Euler, each stage can be written into a steady state form.

4.2 Finite Volume / Discontinuous Galerkin Hybrid

One of the important features of the hybrid method is that it is allows for different discretizations of each component of (4.7) as well as (4.8). While the focus of this work is on the hybridization in space, we also allow for discrete ordinate angular discretizations of different orders. Specifically, we use finite volume (FV) with (possibly) high-order discrete ordinates to solve (4.7a) and (4.8), and we use discontinuous Galerkin (DG) with (possibly) low-order discrete ordinates to solve (4.7b). The FV and DG discretizations will both be

formally second-order and, to allow for sweeping, will use upwinding to define numerical traces at cell interfaces.

Let the high-order quadrature be of order N_u with N_u^* points and weights. Likewise, let the low-order quadrature be of order N_c with N_c^* points and weights. For the remainder of this thesis, we will use the simplified notation

$$\{\Omega_i^\mathrm{u}, w_i^\mathrm{u}\}_{i=1}^{N_\mathrm{u}^*} := \{\Omega_i^{N_\mathrm{u}}, w_i^{N_\mathrm{u}}\}_{i=1}^{N_\mathrm{u}^*}, \quad \{\Omega_i^\mathrm{c}, w_i^\mathrm{c}\}_{i=1}^{N_\mathrm{c}^*} := \{\Omega_i^{N_\mathrm{c}}, w_i^{N_\mathrm{c}}\}_{i=1}^{N_\mathrm{c}^*}, \tag{4.9}$$

and define the component boundaries

$$\partial X_i^{\mathrm{n},\pm} = \{r \in \partial X \colon \pm \Omega_i^\mathrm{n} \cdot n(r) > 0\}, \quad \text{for } \mathrm{n} \in \{\mathrm{u}, \mathrm{c}\}. \tag{4.10}$$

We denote a hybrid angular discretization by $\mathrm{S}_{N_\mathrm{u}}\mathrm{S}_{N_\mathrm{c}}$ and by X-Y a spatial discretization that uses method X to discretize the uncollided component ψ_u and method Y to discretize the collided component ψ_c. In Table 4.1, we summarize the leading order terms in the flop count and degrees of freedom, for both Cartesian and triangular meshes. The values are given per iteration per spatial mesh cell and their derivation is explained in greater detail in Appendix C.

We formulate the spatial discretization using bilinear and linear operators common to DG discretization [31]. Let $\{\mathcal{T}_h\}_{h>0}$ be a regular family of partitions of X into open elements K, with $h_K = \mathrm{diam}(K)$ and $h = \max_{K\in\mathcal{T}_h} h_K$. Let $Q_i(K)$ be the set of polynomials with support K with maximum degree i in each spatial dimension, and let $P_i(K)$ be the set of

Table 4.1: Comutational scaling for Cartesian and triangular mesh cells.

	Cartesian		Triangular	
	Flops	DOF	Flops	DOF
FV	N^*	N^*	N^*	N^*
DG	$(2^d + 2^{2d})N^*$	$2^d N^*$	$(d+1)^2 N^*$	$(d+1)N^*$
DG-DG	$(2^d + 2^{2d})(N_\mathrm{u}^* + N_\mathrm{c}^*)$	$2^d(N_\mathrm{u}^* + N_\mathrm{c}^*)$	$(d+1)\,(d+2)\,(N_\mathrm{u}^* + N_\mathrm{c}^*)$	$(d+1)(N_\mathrm{u}^* + N_\mathrm{c}^*)$
FV-DG	$N_\mathrm{u}^* + (2^d + 2^{2d})N_\mathrm{c}^*$	$N_\mathrm{u}^* + 2^d N_\mathrm{c}^*$	$N_\mathrm{u}^* + (d+1)\,(d+2)\,N_\mathrm{c}^*$	$N_\mathrm{u}^* + (d+1)N_\mathrm{c}^*$

Leading order operations (Flops) and degrees of freedom (DOF) per iteration per cell. Here, d is the dimension of the spatial domain. DG methods on Cartesian grids use polynomials in Q_1 for each cell, and on triangular grids use polynomials in P_1. Further details can be found in Appendix C.

polynomial with support K with total degree less than or equal to i. Let $\mathcal{E}_h^{\text{in}}$ be the set of all interior edges of $\mathcal{T}_h$; let $\mathcal{E}_h^{\text{ex}}$ be the set of exterior edges; and let $\mathcal{E}_{h,i}^{\text{n},\pm} = \mathcal{E}_h^{\text{ex}} \cap \partial X_i^{\text{n},\pm}$ for $\text{n} \in \{\text{u}, \text{c}\}$. For each edge e, let n_e be a fixed normal vector with respect to e. For interior edges, the direction of n_e is chosen by convention. For exterior edges, we assume that n_e points outward from the domain. Given an edge e, let v be any scalar-valued function that is smooth on the cells adjacent to e. Then

$$v^+(r) = \lim_{\varepsilon \to 0+} v(r + \varepsilon n_e), \qquad v^-(r) = \lim_{\varepsilon \to 0+} v(r - \varepsilon n_e), \tag{4.11}$$

and the jump of v at r is $[\![v]\!](r) = v^+(r) - v^-(r)$.

For $i \in \{0, 1\}$, define

$$\mathcal{X}_{h,i} = \{v \in L^2(X) : \forall K \in \mathcal{T}_h,\ v|_K \in Z_i(K)\} \tag{4.12}$$

where $Z = Q$ for Cartesian grids and $Z = P$ for triangular grids. Let $\mathcal{W}_{h,0}^{\text{u}} = (\mathcal{X}_{h,0})^{N_{\text{u}}^*}$ and $\mathcal{W}_{h,1}^{\text{c}} = (\mathcal{X}_{h,1})^{N_{\text{c}}^*}$, and for all $v^{\text{u}} \in \mathcal{W}_{h,0}^{\text{u}}$ or $v^{\text{c}} \in \mathcal{W}_{h,1}^{\text{c}}$

$$v^{\text{u}} = [v_1^{\text{u}}, \ldots, v_{N_{\text{u}}^*}^{\text{u}}]^T, \quad v^{\text{c}} = [v_1^{\text{c}}, \ldots, v_{N_{\text{c}}^*}^{\text{c}}]^T. \tag{4.13}$$

For finite volume methods, polynomial approximations are generated from cell averages in neighboring cells. In particular, approximations in $\mathcal{X}_{h,1}$ for the uncollided equations are generated using elements of $\mathcal{X}_{h,0}$.

For the uncollided equation,the bilinear form $\mathcal{B}_\epsilon^{\text{u}} \colon \mathcal{W}_{h,0}^{\text{u}} \times \mathcal{W}_{h,0}^{\text{u}} \to \mathbb{R}$, corresponding to the left-hand side of (4.7a), is given by

$$\mathcal{B}_\epsilon^{\text{u}}(g, v) = \sum_{i=1}^{N_{\text{u}}^*} w_i^{\text{u}} \mathcal{B}_{\epsilon,i}^{\text{u}}(g, v), \tag{4.14}$$

where $\mathcal{B}^{\mathrm{u}}_{\epsilon,i}\colon \mathcal{W}^{\mathrm{u}}_{h,0} \times \mathcal{W}^{\mathrm{u}}_{h,0} \to \mathbb{R}$ is given by

$$\mathcal{B}^{\mathrm{u}}_{\epsilon,i}(g,v) = \sum_{K\in\mathcal{T}_h} \int_K \frac{\sigma_{\mathrm{t}}}{\epsilon} g_i v_i \, dr + \sum_{e\in\mathcal{E}_h^{\mathrm{in}}} \int_e \Omega_i^{\mathrm{u}} \cdot n_e (Rg)_i^{\uparrow} [\![v_i]\!] \, dr + \sum_{e\in\mathcal{E}_{h,i}^{\mathrm{u},+}} \int_e \Omega_i^{\mathrm{u}} \cdot n_e (Rg)_i^{\uparrow} v_i \, dr, \tag{4.15}$$

$R\colon \mathcal{W}^{\mathrm{u}}_{h,0} \to \mathcal{W}^{\mathrm{u}}_{h,1}$ is a reconstruction operator, and for any $v \in \mathcal{W}^{\mathrm{n}}_{h,1}$,

$$v_i^{\uparrow} = \begin{cases} v_i^-, & \Omega_i^{\mathrm{n}} \cdot n_e > 0 \\ v_i^+, & \Omega_i^{\mathrm{n}} \cdot n_e < 0, \end{cases} \qquad \mathrm{n} \in \{\mathrm{u}, \mathrm{c}\}. \tag{4.16}$$

Details of the operator R, for the case $Z_i(K) = Q_0(K)$ and $d = 2$, are given in Appendix D.

The linear operator $\mathcal{F}^{\mathrm{u}}_i\colon \mathcal{W}^{\mathrm{u}}_{h,0} \to \mathbb{R}$, corresponding to the right-hand side of (4.7a), is given by

$$\mathcal{F}^{\mathrm{u}}(v) = \sum_{i=1}^{N_{\mathrm{u}}^*} w_i^{\mathrm{u}} \mathcal{F}_i^{\mathrm{u}}(v), \tag{4.17}$$

where $\mathcal{F}^{\mathrm{u}}_i\colon \mathcal{W}^{\mathrm{u}}_{h,0} \to \mathbb{R}$ is given by

$$\mathcal{F}_i^{\mathrm{u}}(v) = \sum_{K\in\mathcal{T}_h} \int_K \epsilon S_i v_i \, dr + \sum_{e\in\mathcal{E}_{h,i}^{\mathrm{u},-}} \int_e |\Omega_i^{\mathrm{u}} \cdot n_e| \psi_{\mathrm{b},i} v_i \, dr, \tag{4.18}$$

$S_i = S_i^{N_{\mathrm{u}}}(r)$, and $\psi_{\mathrm{b},i} = \psi_{\mathrm{b}}(r, \Omega_i^{\mathrm{u}})$.

For the collided equations, the bilinear form $\mathcal{B}^{\mathrm{c}}_{\epsilon}\colon \mathcal{W}^{\mathrm{c}}_{h,1} \times \mathcal{W}^{\mathrm{c}}_{h,1} \to \mathbb{R}$ corresponding to the left-hand side of (4.7b) is given by

$$\mathcal{B}^{\mathrm{c}}_{\epsilon}(g,v) = \sum_{i=1}^{N_{\mathrm{c}}^*} w_i^{\mathrm{c}} \mathcal{B}^{\mathrm{c}}_{\epsilon,i}(g,v), \tag{4.19}$$

where $\mathcal{B}^{\mathrm{c}}_{\epsilon,i}\colon \mathcal{W}^{\mathrm{c}}_{h,1}\times\mathcal{W}^{\mathrm{c}}_{h,1}\to\mathbb{R}$ is given by

$$\mathcal{B}^{\mathrm{c}}_{\epsilon,i}(g,v)=\sum_{K\in\mathcal{T}_h}\int_K -g_i\Omega^{\mathrm{c}}_i\cdot\nabla_r v_i+\frac{\sigma_{\mathrm{t}}}{\epsilon}g_i v_i\,dr \tag{4.20a}$$
$$+\sum_{e\in\mathcal{E}^{\mathrm{in}}_h}\int_e \Omega^{\mathrm{c}}_i\cdot n_e g_i^{\uparrow}[\![v_i]\!]\,dr+\sum_{e\in\mathcal{E}^{\mathrm{c},+}_{h,i}}\int_e \Omega^{\mathrm{c}}_i\cdot n_e g_i^{\uparrow}v_i\,dr.$$

Additionally, the bilinear form $\mathcal{C}^{\mathrm{c,c}}_{\epsilon}\colon \mathcal{W}^{\mathrm{c}}_{h,1}\times\mathcal{W}^{\mathrm{c}}_{h,1}\to\mathbb{R}$ corresponding to the first term of the right-hand side of (4.7b) is

$$\mathcal{C}^{\mathrm{c,c}}_{\epsilon}(g,v)=\sum_{i=1}^{N^*_{\mathrm{c}}} w^{\mathrm{c}}_i\mathcal{C}^{\mathrm{c,c}}_{\epsilon,i}, \tag{4.21}$$

where $\mathcal{C}^{\mathrm{c,c}}_{\epsilon,i}\colon \mathcal{W}^{\mathrm{c}}_{h,1}\times\mathcal{W}^{\mathrm{c}}_{h,1}\to\mathbb{R}$ is given by

$$\mathcal{C}^{\mathrm{c,c}}_{\epsilon,i}(g,v)=\sum_{K\in\mathcal{T}_h}\int_K \left(Q^{\mathrm{c,c}}_{\mathrm{s},\epsilon}g\right)_i v_i\,dr, \tag{4.22}$$

and $Q^{\mathrm{m,n}}_{\mathrm{s},\epsilon}:\mathcal{W}^{\mathrm{n}}_{h,1}\to\mathcal{W}^{\mathrm{m}}_{h,1}$ is a linear operator given by (cf. (2.50))

$$Q^{\mathrm{m,n}}_{\mathrm{s},\epsilon}v=\sigma_{\mathrm{s},\epsilon}(A^{\mathrm{n}}v)\mathbb{1}^{\mathrm{m}},\quad A^{\mathrm{n}}v:=\frac{1}{4\pi}\left\langle v,\mathbb{1}^{\mathrm{n}}\right\rangle_N,\ \forall\ v\in\mathcal{W}^{\mathrm{n}}_{h,1},\ \mathrm{m},\mathrm{n}\in\{\mathrm{u},\mathrm{c}\}. \tag{4.23}$$

Here $\sigma_{\mathrm{s},\epsilon}=(\sigma_{\mathrm{t}}/\epsilon-\epsilon\sigma_{\mathrm{a}})$ and $\mathbb{1}^{\mathrm{m}}\in\mathbb{R}^{N^*_{\mathrm{m}}}$ is a vector whose component are all one. Finally, the bilinear form $\mathcal{C}^{\mathrm{u,c}}_{\epsilon}\colon \mathcal{W}^{\mathrm{u}}_{h,1}\times\mathcal{W}^{\mathrm{c}}_{h,1}\to\mathbb{R}$ and the generic bilinear form $\mathcal{C}^{\mathrm{n,u}}_{\epsilon}\colon \mathcal{W}^{\mathrm{n}}_{h,1}\times\mathcal{W}^{\mathrm{u}}_{h,0}\to\mathbb{R}$ are given by

$$\mathcal{C}^{\mathrm{u,c}}_{\epsilon}(g,v)=\sum_{i=1}^{N^*_{\mathrm{c}}} w^{\mathrm{c}}_i\mathcal{C}^{\mathrm{u,c}}_{\epsilon,i},\quad \mathcal{C}^{\mathrm{n,u}}_{\epsilon}(g,v)=\sum_{i=1}^{N^*_{\mathrm{u}}} w^{\mathrm{u}}_i\mathcal{C}^{\mathrm{n,u}}_{\epsilon,i},\quad \mathrm{n}\in\{\mathrm{u},\mathrm{c}\}, \tag{4.24}$$

where $\mathcal{C}^{\mathrm{u,c}}_{\epsilon,i}\colon \mathcal{W}^{\mathrm{u}}_{h,1}\times\mathcal{W}^{\mathrm{c}}_{h,1}\to\mathbb{R}$ and $\mathcal{C}^{\mathrm{n,u}}_{\epsilon,i}\colon \mathcal{W}^{\mathrm{n}}_{h,1}\times\mathcal{W}^{\mathrm{u}}_{h,0}\to\mathbb{R}$ are given by

$$\mathcal{C}^{\mathrm{u,c}}_{\epsilon,i}(g,v)=\sum_{K\in\mathcal{T}_h}\int_K \left(Q^{\mathrm{c,u}}_{\mathrm{s},\epsilon}g\right)_i v_i\,dr,\quad \mathcal{C}^{\mathrm{n,u}}_{\epsilon,i}(g,v)=\sum_{K\in\mathcal{T}_h}\int_K \left(Q^{\mathrm{u,n}}_{\mathrm{s},\epsilon}g\right)_i v_i\,dr,\quad \mathrm{n}\in\{\mathrm{u},\mathrm{c}\}. \tag{4.25}$$

Using these bilinear forms, our method is to find $f_h^{\mathrm{u}} \in \mathcal{W}_{h,0}^{\mathrm{u}}$, $f_h^{\mathrm{c}} \in \mathcal{W}_{h,1}^{\mathrm{c}}$, and $f_h \in \mathcal{W}_{h,0}^{\mathrm{u}}$ such that the following holds:

$$\mathcal{B}_\epsilon^{\mathrm{u}}(f_h^{\mathrm{u}}, v) = \mathcal{F}^{\mathrm{u}}(v), \quad \forall\ v \in \mathcal{W}_{h,0}^{\mathrm{u}}, \tag{4.26a}$$

$$\mathcal{B}_\epsilon^{\mathrm{c}}(f_h^{\mathrm{c}}, v) - \mathcal{C}_\epsilon^{\mathrm{c,c}}(f_h^{\mathrm{c}}, v) = \mathcal{F}^{\mathrm{c}}(v), \quad \forall\ v \in \mathcal{W}_{h,1}^{\mathrm{c}}, \tag{4.26b}$$

$$\mathcal{B}_\epsilon^{\mathrm{u}}(f_h, v) = \mathcal{F}(v), \quad \forall\ v \in \mathcal{W}_{h,0}^{\mathrm{u}}, \tag{4.26c}$$

where the linear operators $\mathcal{F}^{\mathrm{c}} \colon \mathcal{W}_{h,1}^{\mathrm{c}} \to \mathbb{R}$ and $\mathcal{F} \colon \mathcal{W}_{h,0}^{\mathrm{u}} \to \mathbb{R}$ are given by

$$\mathcal{F}^{\mathrm{c}}(v) = \mathcal{C}_\epsilon^{\mathrm{u,c}}(f_h^{\mathrm{u}}, v), \tag{4.27a}$$

$$\mathcal{F}(v) = \mathcal{F}^{\mathrm{u}}(v) + \mathcal{C}_\epsilon^{\mathrm{u,u}}(f_h^{\mathrm{u}}, v) + \mathcal{C}_\epsilon^{\mathrm{c,u}}(f_h^{\mathrm{c}}, v). \tag{4.27b}$$

To assemble the matrix components for the operator form of (4.26), let $M_{\mathrm{u}} = \dim(\mathcal{X}_{h,0})$ and $M_{\mathrm{c}} = \dim(\mathcal{X}_{h,1})$. Then $\dim(\mathcal{W}_{h,0}^{\mathrm{u}}) = M_{\mathrm{u}} N_{\mathrm{u}}^*$ and $\dim(\mathcal{W}_{h,0}^{\mathrm{c}}) = M_{\mathrm{c}} N_{\mathrm{c}}^*$. Let

$$\{\mathbf{b}^{\mathrm{u},(i,k)} : i = 1, \dots N_{\mathrm{u}}^*, k = 1, \dots M_{\mathrm{u}}\} \quad \text{and} \quad \{\mathbf{b}^{\mathrm{c},(i,k)} : i = 1, \dots N_{\mathrm{c}}^*, k = 1, \dots M_{\mathrm{c}}\} \tag{4.28}$$

be two sets of vector-valued basis functions for $\mathcal{W}_{h,0}^{\mathrm{u}}$ and $\mathcal{W}_{h,0}^{\mathrm{c}}$, respectively, and set

$$f_h^{\mathrm{u}} = \sum_{i=1}^{N_{\mathrm{u}}^*} \sum_{k=1}^{M_{\mathrm{u}}} \alpha_{(i,k)}^{\mathrm{u}} \mathbf{b}^{\mathrm{u},(i,k)}, \qquad f_h^{\mathrm{c}} = \sum_{i=1}^{N_{\mathrm{c}}^*} \sum_{k=1}^{M_{\mathrm{c}}} \alpha_{(i,k)}^{\mathrm{c}} \mathbf{b}^{\mathrm{c},(i,k)} \quad \text{and} \quad f_h = \sum_{i=1}^{N_{\mathrm{u}}^*} \sum_{k=1}^{M_{\mathrm{u}}} \alpha_{(i,k)} \mathbf{b}^{\mathrm{u},(i,k)}. \tag{4.29}$$

Then the matrix form of (4.26) is an equation for the coefficient vectors

$$\boldsymbol{\alpha}_h^{\mathrm{u}} = [\alpha_{(1,1)}^{\mathrm{u}}, \alpha_{(1,2)}^{\mathrm{u}}, \dots, \alpha_{(1,M_{\mathrm{u}})}^{\mathrm{u}}, \alpha_{(2,1)}^{\mathrm{u}}, \dots, \alpha_{(N_{\mathrm{u}}^*,M_{\mathrm{u}})}^{\mathrm{u}}]^T \in \mathbb{R}^{N_{\mathrm{u}}^* M_{\mathrm{u}}}, \tag{4.30}$$

$$\boldsymbol{\alpha}_h^{\mathrm{c}} = [\alpha_{(1,1)}^{\mathrm{c}}, \alpha_{(1,2)}^{\mathrm{c}}, \dots, \alpha_{(1,M_{\mathrm{c}})}^{\mathrm{c}}, \alpha_{(2,1)}^{\mathrm{c}}, \dots, \alpha_{(N_{\mathrm{c}}^*,M_{\mathrm{c}})}^{\mathrm{c}}]^T \in \mathbb{R}^{N_{\mathrm{c}}^* M_{\mathrm{c}}}, \text{and} \tag{4.31}$$

$$\boldsymbol{\alpha}_h = [\alpha_{(1,1)}, \alpha_{(1,2)}, \dots, \alpha_{(1,M_{\mathrm{u}})}, \alpha_{(2,1)}, \dots, \alpha_{(N^*,M_{\mathrm{u}})}]^T \in \mathbb{R}^{N_{\mathrm{u}}^* M_{\mathrm{u}}}, \tag{4.32}$$

that takes the form

$$L_\epsilon^{\mathrm{u}} \boldsymbol{\alpha}_h^{\mathrm{u}} = \mathbf{S}_h^{\mathrm{u}}, \tag{4.33a}$$

$$L_\epsilon^{\mathrm{c}} \boldsymbol{\alpha}_h^{\mathrm{c}} - C_\epsilon^{\mathrm{c,c}} \boldsymbol{\alpha}_h^{\mathrm{c}} = C_\epsilon^{\mathrm{c,u}} \boldsymbol{\alpha}_h^{\mathrm{u}}, \tag{4.33b}$$

$$L_\epsilon^{\mathrm{u}} \boldsymbol{\alpha}_h = C_\epsilon^{\mathrm{u,c}} \boldsymbol{\alpha}_h^{\mathrm{c}} + C_\epsilon^{\mathrm{u,u}} \boldsymbol{\alpha}_h^{\mathrm{u}} + \mathbf{S}_h^{\mathrm{u}} \tag{4.33c}$$

where the components of $L_\epsilon^{\mathrm{n}} : \mathbb{R}^{(N_{\mathrm{n}}^* M_{\mathrm{n}})} \to \mathbb{R}^{(N_{\mathrm{n}}^* M_{\mathrm{n}})}$ for $\mathrm{n} \in \{\mathrm{u}, \mathrm{c}\}$ are

$$(L_\epsilon^{\mathrm{n}})_{(i,k),(i',k')} = \mathcal{B}_\epsilon^{\mathrm{n}}(\mathbf{b}^{\mathrm{n},(i',k')}, \mathbf{b}^{\mathrm{n},(i,k)}), \qquad 1 \le i, i' \le N_{\mathrm{n}}^*, \quad 1 \le k, k' \le M_{\mathrm{n}}, \tag{4.34}$$

the components of $C_\epsilon^{\mathrm{n,m}} : \mathbb{R}^{(N_{\mathrm{m}}^* M_{\mathrm{m}})} \to \mathbb{R}^{(N_{\mathrm{n}}^* M_{\mathrm{n}})}$, for $\mathrm{m}, \mathrm{n} \in \{\mathrm{u}, \mathrm{c}\}$ are

$$(C_\epsilon^{\mathrm{n,m}})_{(i,k),(i',k')} = \mathcal{C}_\epsilon^{\mathrm{m,n}}(\mathbf{b}^{\mathrm{m},(i',k')}, \mathbf{b}^{\mathrm{n},(i,k)}), \tag{4.35}$$

for $1 \le i' \le N_{\mathrm{m}}^*,\ 1 \le i \le N_{\mathrm{n}}^*, \quad 1 \le k' \le M_{\mathrm{m}},\ 1 \le k \le M_{\mathrm{n}}$, and the (i,k)-th component of the source vector $\mathbf{S}_h^{\mathrm{u}}$ is

$$S_{(i,k)}^{\mathrm{u}} = \mathcal{F}^{\mathrm{u}}(\mathbf{b}^{\mathrm{u},(i,k)}), \quad \forall\ i \le N_{\mathrm{u}},\ k \le M_{\mathrm{u}}. \tag{4.36}$$

Solving (4.33a) and (4.33c) can be solved easily by inverting the streaming operator L_ϵ^{u}, assuming the right-hand side of each respective equation is known. To solve (4.33b), we reformulate it into a Krylov framework by inverting the streaming operator $L_{\epsilon,h}^{\mathrm{c}}$ and applying the discrete average operator A^{c} to both sides. Here we overload the operator so that $A^{\mathrm{c}} : \mathbb{R}^{M_{\mathrm{c}} N_{\mathrm{c}}^*} \to \mathbb{R}^{M_{\mathrm{c}}}$ where

$$(A^{\mathrm{c}}\mathbf{v})_k = \frac{1}{4\pi} \sum_{i=1}^{N_{\mathrm{c}}^*} w_i^{\mathrm{c}} v_{(i,k)},\ \forall\ 1 \le k \le M_{\mathrm{c}},\ \forall\ \mathbf{v} \in \mathbb{R}^{M_{\mathrm{c}} N_{\mathrm{c}}^*}. \tag{4.37}$$

Additionally, let β_k^{u} and β_k^{c} be basis functions such that $\mathrm{span}\{\beta_k^{\mathrm{u}}\}_{k=1}^{M_{\mathrm{u}}} = \mathcal{X}_{h,0}$ and $\mathrm{span}\{\beta_k^{\mathrm{c}}\}_{k=1}^{M_{\mathrm{c}}} = \mathcal{X}_{h,1}$. We assume that

$$\mathbf{b}^{\mathrm{n},(i,k)} = [b_1^{\mathrm{n},(i,k)}, \ldots, b_{N_{\mathrm{n}}^*}^{\mathrm{n},(i,k)}]^T, \quad \mathrm{n} \in \{\mathrm{u}, \mathrm{c}\}, \tag{4.38}$$

where $b^{\mathrm{n},(i,k)}_{i'} = \beta^{\mathrm{n}}_k \delta_{i,i'}$, in which case

$$(C^{\mathrm{c,c}}_\epsilon \boldsymbol{\alpha}^{\mathrm{c}}_h)_{i,k} = \sum_{k'=1}^{M_{\mathrm{c}}} w^{\mathrm{c}}_i \int_X \sigma_{\mathrm{s},\epsilon} \beta^{\mathrm{c}}_{k'} \beta^{\mathrm{c}}_k \left(A^{\mathrm{c}} \boldsymbol{\alpha}^{\mathrm{c}}_h\right)_{k'} . \tag{4.39}$$

Let $\boldsymbol{\varphi}^{\mathrm{n}}_h = A^{\mathrm{n}} \boldsymbol{\alpha}^{\mathrm{n}}_h$, and let $\Sigma^{\mathrm{n,m}}_\epsilon : \mathbb{R}^{M_{\mathrm{m}}} \to \mathbb{R}^{(N^*_{\mathrm{n}} M_{\mathrm{n}})}$ for $\mathrm{m}, \mathrm{n} \in \{\mathrm{u}, \mathrm{c}\}$, where

$$(\Sigma^{\mathrm{n,m}}_\epsilon)_{(i,k),k'} = w^\ell_i \int_X \sigma_{\mathrm{s},\epsilon} \beta^{\mathrm{m}}_{k'} \beta^{\mathrm{n}}_k, \quad 1 \le i \le N^*_{\mathrm{n}},\ 1 \le k \le M_{\mathrm{n}},\ 1 \le k' \le M_{\mathrm{m}}, \tag{4.40}$$

Then

$$C^{\mathrm{n,m}}_\epsilon \boldsymbol{\alpha}^{\mathrm{m}}_h = \Sigma^{\mathrm{n,m}}_\epsilon \boldsymbol{\varphi}^{\mathrm{m}}_h, \quad \mathrm{m}, \mathrm{n} \in \{\mathrm{u}, \mathrm{c}\}. \tag{4.41}$$

Using (4.41), we invert the transport operator L^{c}_ϵ in (4.33b) and apply the discrete average operator A^{c} to both sides of the equation. Then (4.33b) can be written in the following form:

$$\left(I^{\mathrm{c}} - A^{\mathrm{c}} (L^{\mathrm{c}}_\epsilon)^{-1} \Sigma^{\mathrm{c,c}}_\epsilon\right) \boldsymbol{\varphi}^{\mathrm{c}}_h = A^{\mathrm{c}} (L^{\mathrm{c}}_{\epsilon,h})^{-1} \Sigma^{\mathrm{c,u}}_\epsilon \boldsymbol{\varphi}^{\mathrm{u}}_h, \tag{4.42}$$

where $I^{\mathrm{c}} \in \mathbb{R}^{(N^*_{\mathrm{c}} M_{\mathrm{c}}) \times (N^*_{\mathrm{c}} M_{\mathrm{c}})}$ is the identity matrix, and (4.33c) simplifies to

$$L^{\mathrm{u}}_\epsilon \boldsymbol{\alpha}_h = \Sigma^{\mathrm{u,c}}_\epsilon \boldsymbol{\varphi}^{\mathrm{c}}_h + \Sigma^{\mathrm{u,u}}_\epsilon \boldsymbol{\varphi}^{\mathrm{u}}_h + \mathbf{q}^{\mathrm{u}}_h. \tag{4.43}$$

In summary, (4.33) can be solved using (4.42) and (4.43) as outlined in Algorithm 1 below.

Algorithm 1 Steady State Spatial Hybrid Solution Algorithm.

Initial Data: $\sigma_{\mathrm{t}} \ge 0$, $\sigma_{\mathrm{a}} \ge 0$, $\epsilon > 0$, ψ_{b}.

$q^{\mathrm{u}}_{(i,k)} \leftarrow \mathcal{F}^{\mathrm{u}}(\mathbf{b}^{\mathrm{u},(i,k)})$, $\forall\ i \le N^*_{\mathrm{u}}$, $\forall\ k \le M_{\mathrm{u}}$, // Initialize Source, (4.36)

$\boldsymbol{\alpha}^{\mathrm{u}}_h \leftarrow (L^{\mathrm{u}}_\epsilon)^{-1} \mathbf{q}^{\mathrm{u}}_h$, // Solve (4.33a)

$\boldsymbol{\varphi}^{\mathrm{u}}_h \leftarrow A^{\mathrm{u}} \boldsymbol{\alpha}^{\mathrm{u}}_h$, // Form $\boldsymbol{\varphi}^{\mathrm{u}}_h$

$\boldsymbol{\varphi}^{\mathrm{c}}_h \leftarrow \left(I^{\mathrm{c}} + A^{\mathrm{c}} (L^{\mathrm{c}}_\epsilon)^{-1} \Sigma^{\mathrm{c,c}}_\epsilon\right)^{-1} A^{\mathrm{c}} (L^{\mathrm{c}}_\epsilon)^{-1} \Sigma^{\mathrm{c,u}}_\epsilon \boldsymbol{\varphi}^{\mathrm{u}}_h$, // Solve (4.42) using GMRES

$\boldsymbol{\alpha}_h \leftarrow (L^{\mathrm{u}}_\epsilon)^{-1} (\Sigma^{\mathrm{u,c}}_\epsilon \boldsymbol{\varphi}^{\mathrm{c}}_h + \Sigma^{\mathrm{u,u}}_\epsilon \boldsymbol{\varphi}^{\mathrm{u}}_h + \mathbf{q}^{\mathrm{u}}_h)$, // Solve (4.33c)

4.3 Diffusion limit, steady state

In this section, we show that (4.26) converges to a consistent discretization of the steady-state form of the diffusion limit (2.52). The analysis here closely follows [28]. The scaling of time in (2.48a) ensures that the steady-state analysis applies also to the time-dependent problem (4.4).

We expand the solutions $f_h^{\mathrm{u}}, f_h^{\mathrm{c},}$, and f_h of (4.26) in formal Hilbert expansions:

$$f_h^{\mathrm{u}} = f_h^{\mathrm{u},(0)} + \epsilon f_h^{\mathrm{u},(1)} + \epsilon^2 f_h^{\mathrm{u},(2)} + O(\epsilon^3), \tag{4.44a}$$

$$f_h^{\mathrm{c}} = f_h^{\mathrm{c},(0)} + \epsilon f_h^{\mathrm{c},(1)} + \epsilon^2 f_h^{\mathrm{c},(2)} + O(\epsilon^3), \tag{4.44b}$$

$$f_h = f_h^{(0)} + \epsilon f_h^{(1)} + \epsilon^2 f_h^{(2)} + O(\epsilon^3), \tag{4.44c}$$

where $f_h^{\mathrm{u},(j)}, f_h^{(j)} \in \mathcal{X}_{h,0}$ and $f_h^{\mathrm{u},(j)} \in \mathcal{X}_{h,1}$ for all $j \in \{0,1,2\}$. The problem solved by the leading term $f_h^{(0)}$ is obtained by substituting the expansions (4.44a) into (4.26) and matching powers of ϵ. In order to perform the analysis, we assume that $\psi_b = 0$, that $\sigma_{\mathrm{t}} \in \mathcal{X}_{h,0}$, and that all quadratures employed use positive weights and are exact for polynomials up to degree two:

$$\sum_{i=1}^{N_{\mathrm{n}}^*} w_i^{\mathrm{n}} = 4\pi, \quad \sum_{i=1}^{N_{\mathrm{n}}^*} w_i^{\mathrm{n}} \Omega_i^{\mathrm{n}} = 0, \quad \sum_{i=1}^{N_{\mathrm{n}}^*} w_i^{\mathrm{n}} \left(\Omega_i^{\mathrm{n}} \otimes \Omega_i^{\mathrm{n}}\right) = \frac{4\pi}{3} I, \quad \mathrm{n} \in \{\mathrm{u},\mathrm{c}\},\ \mathrm{j} \in \{1,2,3\}, \tag{4.45}$$

where $\otimes$ is the outer product and $I \in \mathbb{R}^{3\times 3}$ is the identity matrix. To simply the presentation, we set $\bar{f}_h^{\mathrm{n},(j)} = A^{\mathrm{n}} f_h^{\mathrm{n},(j)}$ for $j \in \{0,1,2\}$ and $\mathrm{n} \in \{\mathrm{u},\mathrm{c}\}$, we denote test functions in $\mathcal{W}_{h,0}^{\mathrm{u}}$ and $\mathcal{W}_{h,1}^{\mathrm{c}}$ by v^{u} and v^{c}, respectively, and we assume that jumps across edges in $\mathcal{E}_h^{\mathrm{ex}}$ are computed assuming a zero value on the exterior of X. This last assumption allows us to combine terms over $\mathcal{E}_h = \mathcal{E}_h^{\mathrm{in}} \cup \mathcal{E}_h^{\mathrm{ex}}$. After substituting the expansions (4.44a) into (4.26), the terms that

balance at ϵ^{-1} are

$$\sum_{i=1}^{N_{\mathrm{u}}^*} w_i^{\mathrm{u}} \int_X \sigma_{\mathrm{t}} f_{h,i}^{\mathrm{u},(0)} v_i^{\mathrm{u}} \, dr = 0, \tag{4.46a}$$

$$\sum_{i=1}^{N_{\mathrm{c}}^*} w_i^{\mathrm{c}} \int_X \sigma_{\mathrm{t}} \left(f_{h,i}^{\mathrm{c},(0)} - \bar{f}_h^{\mathrm{c},(0)} \right) v_i^{\mathrm{c}} \, dr = \sum_{i=1}^{N_{\mathrm{u}}^*} w_i^{\mathrm{u}} \int_X \sigma_{\mathrm{t}} \bar{f}_h^{\mathrm{u},(0)} v_i^{\mathrm{c}} \, dr, \tag{4.46b}$$

$$\sum_{i=1}^{N_{\mathrm{u}}^*} w_i^{\mathrm{u}} \int_X \sigma_{\mathrm{t}} f_{h,i}^{(0)} v_i^{\mathrm{u}} \, dr = \sum_{i=1}^{N_{\mathrm{u}}^*} w_i^{\mathrm{u}} \int_X \sigma_{\mathrm{t}} \left(\bar{f}_h^{\mathrm{u},(0)} + \bar{f}_h^{\mathrm{c},(0)} \right) v_i^{\mathrm{u}} \, dr, \tag{4.46c}$$

the terms that balance at ϵ^0 are

$$\sum_{i=1}^{N_{\mathrm{u}}^*} w_i^{\mathrm{u}} \int_X \sigma_{\mathrm{t}} f_{h,i}^{\mathrm{u},(1)} v_i^{\mathrm{u}} \, dr + \sum_{e \in \mathcal{E}_h} \int_e \Omega_i^{\mathrm{u}} \cdot n_e \left(R f_h^{\mathrm{u},(0)} \right)_i^{\uparrow} [\![v_i^{\mathrm{u}}]\!] \, dr = 0, \tag{4.47a}$$

$$\begin{aligned} &\sum_{i=1}^{N_{\mathrm{c}}^*} w_i^{\mathrm{c}} \int_X -f_{h,i}^{\mathrm{c},(0)} \Omega_i^{\mathrm{c}} \cdot \nabla v_i^{\mathrm{c}} + \sigma_{\mathrm{t}} \left(f_{h,i}^{\mathrm{c},(1)} - \bar{f}_h^{\mathrm{c},(1)} \right) v_i^{\mathrm{c}} \, dr \\ &\quad + \sum_{e \in \mathcal{E}_h} \int_e \Omega_i^{\mathrm{c}} \cdot n_e \left(f_{h,i}^{\mathrm{c},(0)} \right)^{\uparrow} [\![v_i^{\mathrm{c}}]\!] \, dr = \sum_{i=1}^{N_{\mathrm{c}}^*} w_i^{\mathrm{c}} \int_X \sigma_{\mathrm{t}} \bar{f}_h^{\mathrm{u},(1)} v_i^{\mathrm{c}} \, dr, \end{aligned} \tag{4.47b}$$

and the terms that balance at ϵ are

$$\sum_{i=1}^{N_{\mathrm{u}}^*} w_i^{\mathrm{u}} \int_X \sigma_{\mathrm{t}} f_{h,i}^{\mathrm{u},(2)} v_i^{\mathrm{u}} + \sum_{e \in \mathcal{E}_h} \int_e \Omega_i^{\mathrm{u}} \cdot n_e \left(R f_h^{\mathrm{u},(1)} \right)_i^{\uparrow} [\![v_i^{\mathrm{u}}]\!] \, dr = \sum_{i=1}^{N_{\mathrm{u}}^*} w_i^{\mathrm{u}} \int_X S_i v_i^{\mathrm{u}}, \tag{4.48a}$$

$$\begin{aligned} &\sum_{i=1}^{N_{\mathrm{c}}^*} w_i^{\mathrm{c}} \int_X -f_{h,i}^{\mathrm{c},(1)} \Omega_i^{\mathrm{c}} \cdot \nabla v_i^{\mathrm{c}} + \sigma_{\mathrm{t}} \left(f_{h,i}^{\mathrm{c},(2)} - \bar{f}_h^{\mathrm{c},(2)} \right) v_i^{\mathrm{c}} + \sigma_{\mathrm{a}} \bar{f}_h^{\mathrm{c},(0)} v_i^{\mathrm{c}} \, dr \\ &\quad + \sum_{e \in \mathcal{E}_h} \int_e \Omega_i^{\mathrm{c}} \cdot n_e \left(f_{h,i}^{\mathrm{c},(1)} \right)^{\uparrow} [\![v_i^{\mathrm{c}}]\!] \, dr = \sum_{i=1}^{N_{\mathrm{c}}^*} w_i^{\mathrm{c}} \int_X \left(\sigma_{\mathrm{t}} \bar{f}_h^{\mathrm{u},(2)} - \sigma_{\mathrm{a}} \bar{f}_h^{\mathrm{u},(0)} \right) v_i^{\mathrm{c}} \, dr. \end{aligned} \tag{4.48b}$$

The contributions of (4.26c) to the balance equations at order ϵ^0 and ϵ are omitted in (4.47) and (4.48) as they will not be used in the analysis that follows.

Define $\mathcal{C}_{h,1}$ to be the subspace of $\mathcal{X}_{h,1}$ where every element is continuous, and define the following:

$$J_{\mathrm{c},h}^{(0)} = \sum_{i=1}^{N_{\mathrm{c}}^*} w_i^{\mathrm{c}} \Omega_i^{\mathrm{c}} f_{h,i}^{\mathrm{c},(1)}. \tag{4.49}$$

Let $S = [S_1, S_2, \dots, S_{N_{\mathrm{u}}^*}]^T$ and $\bar{S} = A^{\mathrm{u}} S$. Let P_0 be the orthogonal projection from $L^2(X)$ onto $\mathcal{X}_{h,0}$ with respect to the usual inner product. Then our main result is the following.

Theorem 4.1. *Let $f_h^{\mathrm{u},(j)}, f_h^{(0)} \in \mathcal{X}_{h,0}$ and $f_h^{\mathrm{c},(j)} \in \mathcal{X}_{h,1}$ solve (4.46), (4.47), and (4.48) for $j \in \{0,1,2\}$. Additionally, let $\sigma_{\mathrm{t}} \in \mathcal{X}_{h,0}$ and $\sigma_{\mathrm{t}} \geq \sigma_{\mathrm{a}} > 0$. Then $f_{h,i}^{(0)} = P_0 \bar{f}_h^{\mathrm{c},(0)}$ for all $i \leq N_{\mathrm{u}}^*$ where P_0 is the orthogonal projection from $\mathcal{X}_{h,1}$ onto $\mathcal{X}_{h,0}$ with respect to the inner product on $L^2(X)$. Moreover, for all $\vartheta \in \mathcal{C}_{h,1}$ and $\varphi \in (\mathcal{X}_{h,1})^3$, $J_{\mathrm{c},h}^{(0)}$ and $\bar{f}_h^{\mathrm{c},(0)}$ satisfy*

$$\int_X -J_{\mathrm{c},h}^{(0)} \cdot \nabla \vartheta + 4\pi \sigma_{\mathrm{a}} \bar{f}_h^{\mathrm{c},(0)} \vartheta \, dr = 4\pi \int_X \left(P_0 \bar{S}\right) \vartheta \, dr, \tag{4.50a}$$

$$\int_X \left(\frac{4\pi}{3} \nabla \bar{f}_h^{\mathrm{c},(0)} + \sigma_{\mathrm{t}} J_{\mathrm{c},h}^{(0)} \right) \cdot \varphi \, dr = 0. \tag{4.50b}$$

Theorem 4.1 is a consistent discretization of the first-order, steady-state form of the diffusion limit (2.52). To prove it we first require some preliminary lemmas.

Lemma 4.2. *Let $f_h^{\mathrm{u},(0)}, f_h^{\mathrm{c},(0)}$, and $f_h^{(0)}$ solve (4.46). Then $f_{h,i}^{\mathrm{u},(0)} = 0$ for all $i \leq N_{\mathrm{u}}^*$, $f_{h,i}^{\mathrm{c},(0)} = \bar{f}_h^{\mathrm{c},(0)}$ for all $i \leq N_{\mathrm{c}}^*$, and $f_{h,i}^{(0)} = P_0 \bar{f}_h^{\mathrm{c},(0)}$ for all $i \leq N_{\mathrm{u}}^*$ where P_0 is the orthogonal projection from $\mathcal{X}_{h,1}$ onto $\mathcal{X}_{h,0}$ with respect to the inner product on $L^2(X)$.*

Proof. Let $v_i^{\mathrm{u}} = f_{h,i}^{\mathrm{u},(0)}$ for all $i \leq N_{\mathrm{u}}^*$ in (4.46a). Then

$$\sum_{i=1}^{N_{\mathrm{u}}^*} w_i^{\mathrm{u}} \int_X \sigma_{\mathrm{t}} \left(f_{h,i}^{\mathrm{u},(0)} \right)^2 \, dr = 0. \tag{4.51}$$

Since $\sigma_{\mathrm{t}} > 0$ and $w_i^{\mathrm{u}} > 0$ for all $i \leq N_{\mathrm{u}}^*$, it follows from (4.51) that $f_{h,i}^{\mathrm{u},(0)} = 0$ for all $i \leq N_{\mathrm{u}}^*$.

We now show that $f_{h,i}^{\mathrm{c},(0)} = \bar{f}_h^{\mathrm{c},(0)}$ for all $i \leq N_{\mathrm{c}}^*$. Let $v_i^{\mathrm{c}} = f_{h,i}^{\mathrm{c},(0)} - \bar{f}_h^{\mathrm{c},(0)}$ for all $i \leq N_{\mathrm{c}}^*$. Then (4.46b) becomes

$$\sum_{i=1}^{N_{\mathrm{c}}^*} w_i^{\mathrm{c}} \int_X \sigma_{\mathrm{t}} \left(f_{h,i}^{\mathrm{c},(0)} - \bar{f}_h^{\mathrm{c},(0)} \right)^2 \, dr = 0. \tag{4.52}$$

Since $\sigma_{\mathrm{t}} > 0$ and $w_i^{\mathrm{c}} > 0$ for all $i \leq N_{\mathrm{c}}^*$, (4.52) implies $f_{h,i}^{\mathrm{c},(0)} = \bar{f}_h^{\mathrm{c},(0)}$ for all $i \leq N_{\mathrm{c}}^*$.

We now show that $f_{h,i}^{(0)} = P_0 \bar{f}_h^{\mathrm{c},(0)}$ for all $i \leq N_{\mathrm{u}}^*$. Let $v_i^{\mathrm{u}} = f_{h,i}^{(0)} - P_0 \bar{f}_h^{\mathrm{c},(0)}$ for all $i \leq N_{\mathrm{u}}^*$. Since $\bar{f}_h^{\mathrm{u},(0)} = 0$, $\sigma_{\mathrm{t}} \in \mathcal{X}_{h,0}$, and P_0 is the orthogonal projection onto $\mathcal{X}_{h,0}$, it follows then that

$$\sum_{i=1}^{N_{\mathrm{u}}^*} w_i^{\mathrm{u}} \int_X \sigma_{\mathrm{t}} \left(f_{h,i}^{(0)} - P_0 \bar{f}_h^{\mathrm{c},(0)} \right)^2 \, dr = 0. \tag{4.53}$$

Since $\sigma_{\mathrm{t}} > 0$ and $w_i^{\mathrm{u}} > 0$ for all $i \leq N_{\mathrm{u}}^*$, (4.53) implies $f_{h,i}^{(0)} = P_0 \bar{f}_h^{\mathrm{c},(0)}$ for all $i \leq N_{\mathrm{u}}^*$. □

Lemma 4.3. *Let $f_h^{\mathrm{u},(1)}$ and $f_h^{\mathrm{c},(1)}$ solve (4.47) where $f_h^{\mathrm{u},(0)}$ and $f_h^{\mathrm{c},(0)}$ are solutions to (4.46a) and (4.46b) respectively. Then $f_{h,i}^{\mathrm{u},(1)} = 0$ for all $i \leq N_{\mathrm{u}}^*$, and $\bar{f}_h^{\mathrm{c},(0)}$ is continuous on X.*

Proof. Because $f_h^{\mathrm{u},(0)} = 0$, (4.47a) becomes

$$\sum_{i=1}^{N_{\mathrm{u}}^*} w_i^{\mathrm{u}} \sum_{K \in \mathcal{T}_h} \int_K \sigma_{\mathrm{t}} f_{h,i}^{\mathrm{u},(1)} v_i^{\mathrm{u}} = 0. \tag{4.54}$$

Similar to the proof of Lemma 4.2, this implies $f_{h,i}^{\mathrm{u},(1)} = 0$ for all $i \leq N_{\mathrm{u}}^*$.

We now show that $\bar{f}_h^{\mathrm{c},(0)}$ is a continuous function on X, in particular by showing that it is continuous on the cell edges. Let $v_i^{\mathrm{c}} = \bar{f}_h^{\mathrm{c},(0)}$ for all $1 \leq i \leq N_{\mathrm{c}}^*$. From Lemma 4.2, $f_{h,i}^{\mathrm{c},(0)} = \bar{f}_h^{\mathrm{c},(0)}$ for all $i \leq N_{\mathrm{c}}^*$. Using the accuracy of the quadrature, which calculates the integral of degree one polynomials exactly, the first term in (4.47b) becomes

$$\sum_{i=1}^{N_{\mathrm{c}}^*} w_i^{\mathrm{c}} \int_X -\bar{f}_h^{\mathrm{c},(0)} \Omega_i^{\mathrm{c}} \cdot \nabla \bar{f}_h^{\mathrm{c},(0)} \, dr = \int_X -\bar{f}_h^{\mathrm{c},(0)} \left(\sum_{i=1}^{N_{\mathrm{c}}^*} w_i^{\mathrm{c}} \Omega_i^{\mathrm{c}} \right) \cdot \nabla \bar{f}_h^{\mathrm{c},(0)} \, dr = 0. \tag{4.55}$$

The next term in (4.47b) becomes

$$\sum_{i=1}^{N_{\mathrm{c}}^*} w_i^{\mathrm{c}} \int_X \sigma_{\mathrm{t}} \left(f_{h,i}^{\mathrm{c},(1)} - \bar{f}_h^{\mathrm{c},(1)} \right) \bar{f}_h^{\mathrm{c},(0)} \, dr = \int_X \sigma_{\mathrm{t}} \left(\sum_{i=1}^{N_{\mathrm{c}}^*} w_i^{\mathrm{c}} \left(f_{h,i}^{\mathrm{c},(1)} - \bar{f}_h^{\mathrm{c},(1)} \right) \right) \bar{f}_h^{\mathrm{c},(0)} = 0. \tag{4.56}$$

With the results in (4.55) and (4.56), (4.47b) becomes

$$0 = \sum_{i=1}^{N_{\mathrm{c}}^*} w_i^{\mathrm{c}} \sum_{e \in \mathcal{E}_h} \int_e \Omega_i^{\mathrm{c}} \cdot n_e \left(\bar{f}_h^{\mathrm{c},(0)} \right)^{\uparrow} \left[\!\left[\bar{f}_h^{\mathrm{c},(0)} \right]\!\right] \, dr = \sum_{e \in \mathcal{E}_h} \int_e \left(\sum_{i=1}^{N_{\mathrm{c}}^*} w_i^{\mathrm{c}} |\Omega_i^{\mathrm{c}} \cdot n_e| \right) \left[\!\left[\bar{f}_h^{\mathrm{c},(0)} \right]\!\right]^2 dr. \tag{4.57}$$

Since $w_i^{\mathrm{c}} > 0$, it follows from (4.57) that $\left[\!\left[\bar{f}_h^{\mathrm{c},(0)} \right]\!\right] = 0$ for every edge in $\mathcal{E}_h$. □

Proof of Theorem 4.1. We first show that $f_{h,i}^{\mathrm{u},(2)} = \frac{P_0 S_i}{\sigma_\mathrm{t}}$ for all $i \leq N_\mathrm{u}^*$. By Lemma 4.2 and 4.3, $f_h^{\mathrm{u},(0)} = 0$ and $f_h^{\mathrm{u},(1)} = 0$. This implies that (4.48a) becomes

$$\sum_{i=1}^{N_\mathrm{u}^*} w_i^\mathrm{u} \int_X \left(\sigma_\mathrm{t} f_{h,i}^{\mathrm{u},(2)} - S_i\right) v_i^\mathrm{u} = 0. \tag{4.58}$$

Let $v_i^\mathrm{u} = \sigma_\mathrm{t} f_{h,i}^{\mathrm{u},(2)} - P_0 S_i$ for all $i \leq N_\mathrm{u}^*$. Since P_0 is the orthogonal projection onto $\mathcal{X}_{h,0}$, (4.58) implies

$$\sum_{i=1}^{N_\mathrm{u}^*} w_i^\mathrm{u} \int_X \left(\sigma_\mathrm{t} f_{h,i}^{\mathrm{u},(2)} - P_0 S_i\right)^2 = 0. \tag{4.59}$$

Similar to the proof of Lemma 4.2, (4.59) implies

$$f_{h,i}^{\mathrm{u},(2)} = \frac{P_0 S_i}{\sigma_\mathrm{t}}. \tag{4.60}$$

As a consequence of this,

$$\bar{f}_h^{\mathrm{u},(2)} = \frac{P_0 \bar{S}}{\sigma_\mathrm{t}}. \tag{4.61}$$

We now show that $\bar{f}_h^{\mathrm{c},(0)}$ and $J_{\mathrm{c},h}^{(0)}$ satisfy (4.50a). Lemma 4.3 implies $\bar{f}_h^{\mathrm{c},(0)} \in \mathcal{C}_{h,1}$. Let $v_i^\mathrm{c} = \vartheta \in \mathcal{X}_{h,1}$ for all $1 \leq i \leq N_\mathrm{c}^*$. Then the first term of (4.48b) becomes:

$$\sum_{i=1}^{N_\mathrm{c}^*} w_i^\mathrm{c} \int_X -f_{h,i}^{\mathrm{c},(1)} \Omega_i^\mathrm{c} \cdot \nabla\vartheta = \int_X -\left(\sum_{i=1}^{N_\mathrm{c}^*} w_i^\mathrm{c} f_{h,i}^{\mathrm{c},(1)} \Omega_i^\mathrm{c}\right) \cdot \nabla\vartheta = \int_X -J_{\mathrm{c},h}^{(0)} \cdot \nabla\vartheta. \tag{4.62}$$

The subsequent terms involving $f_{h,i}^{\mathrm{c},(2)}$ and $\bar{f}_h^{\mathrm{c},(2)}$ in (4.48b) will cancel owing to the definition of $\bar{f}_h^{\mathrm{c},(2)}$. The next term involving $\bar{f}_h^{\mathrm{c},(0)}$ is simply

$$\sum_{i=1}^{N_\mathrm{c}^*} w_i^\mathrm{c} \int_X \sigma_\mathrm{a} \bar{f}_h^{\mathrm{c},(0)} \vartheta \; dr = \int_X 4\pi\sigma_\mathrm{a} \bar{f}_h^{\mathrm{c},(0)} \vartheta \; dr. \tag{4.63}$$

We now restrict the test functions to be continuous on X. As a result

$$\sum_{e \in \mathcal{E}_h} \int_e \Omega_i^\mathrm{c} \cdot n_e \left(f_{h,i}^{\mathrm{c},(1)}\right)^\uparrow [\![\vartheta]\!] \; dr = 0. \tag{4.64}$$

Using the fact that $f_{h,i}^{\mathrm{u},(0)} = 0$, (from Lemma 4.2), and (4.61), the right side of (4.48b) becomes

$$\sum_{i=1}^{N_{\mathrm{c}}^*} w_i^{\mathrm{c}} \int_X \sigma_{\mathrm{t}} \bar{f}_h^{\mathrm{u},(2)} \vartheta \, dr = \int_X 4\pi P_0 \bar{S} \vartheta \, dr. \tag{4.65}$$

Combining these results, (4.48b) implies

$$\int_X -J_{\mathrm{c},h}^{(0)} \cdot \nabla \vartheta + 4\pi \sigma_{\mathrm{a}} \bar{f}_h^{\mathrm{c},(0)} \vartheta \, dr = \int_X 4\pi \sigma_{\mathrm{t}} P_0 \bar{S} \vartheta \, dr. \tag{4.66}$$

Now we show that $\bar{f}_h^{\mathrm{c},(0)}$ and $J_{\mathrm{c},h}^{(0)}$ satisfy (4.50b). Let $v_i^{\mathrm{c}} = \varphi \cdot \Omega_i^{\mathrm{c}}$ for all $1 \leq i \leq N_{\mathrm{c}}^*$, where $\varphi \in (\mathcal{X}_{h,1})^3$ is arbitrary. Using integration by parts and recalling that $f_{h,i}^{\mathrm{c},(0)} = \bar{f}_h^{\mathrm{c},(0)}$ for all $1 \leq i \leq N_{\mathrm{c}}^*$, (from Lemma 4.2), and $f_{h,i}^{\mathrm{u},(1)} = 0$ for all $i \leq N_{\mathrm{u}}^*$, (from Lemma 4.3), we can rewrite (4.47b) as the following:

$$\begin{aligned} \sum_{i=1}^{N_{\mathrm{c}}^*} w_i^{\mathrm{c}} \int_X \left(\Omega_i^{\mathrm{c}} \cdot \nabla \bar{f}_h^{\mathrm{c},(0)} + \sigma_{\mathrm{t}} \left(f_{h,i}^{\mathrm{c},(1)} - \bar{f}_h^{\mathrm{c},(1)} \right) \right) (\varphi \cdot \Omega_i^{\mathrm{c}}) \; dr \\ + \sum_{e \in \mathcal{E}_h} \int_F -\Omega_i^{\mathrm{c}} \cdot n_1 \left[\!\!\left[\bar{f}_h^{\mathrm{c},(0)} \right]\!\!\right] (\varphi \cdot \Omega_i^{\mathrm{c}})^{\downarrow} \; dr = 0. \end{aligned} \tag{4.67}$$

We can infer that the first term of (4.67) is

$$\sum_{i=1}^{N_{\mathrm{c}}^*} w_i^{\mathrm{c}} \int_X \nabla \bar{f}_h^{\mathrm{c},(0)} \cdot (\Omega_i^{\mathrm{c}} \otimes \Omega_i^{\mathrm{c}}) \, \varphi \, dr = \int_X \nabla \bar{f}_h^{\mathrm{c},(0)} \cdot \left(\sum_{i=1}^{N_{\mathrm{c}}^*} w_i^{\mathrm{c}} \Omega_i^{\mathrm{c}} \otimes \Omega_i^{\mathrm{c}} \right) \varphi \, dr = \int_X \frac{4\pi}{3} \nabla \bar{f}_h^{\mathrm{c},(0)} \cdot \varphi \, dr. \tag{4.68}$$

Computing the contribution of $f_h^{\mathrm{c},(1)}$ we have,

$$\sum_{i=1}^{N_{\mathrm{c}}^*} w_i^{\mathrm{c}} \int_X \sigma_{\mathrm{t}} \left(f_{h,i}^{\mathrm{c},(1)} - \bar{f}_h^{\mathrm{c},(1)} \right) (\varphi \cdot \Omega_i^{\mathrm{c}}) \; dr = \int_X \sigma_{\mathrm{t}} \left(\sum_{i=1}^{N_{\mathrm{c}}^*} w_i^{\mathrm{c}} f_{h,i}^{\mathrm{c},(1)} \Omega_i^{\mathrm{c}} \right) \cdot \varphi \; dr = \int_X \sigma_{\mathrm{t}} J_{\mathrm{c},h}^{(0)} \cdot \varphi \; dr. \tag{4.69}$$

All edge flux terms are zero since $\bar{f}_h^{\mathrm{c},(0)} \in \mathcal{C}_{h,1}$. Combining these results give (4.50b). □

4.4 Numerical results

In this section, we compare the performance of the spatial hybrid to standard DG and FV approaches, as well as the angular hybrid DG-DG approach. We also investigate the benefits of hybridization in both the angular and spatial variable. In Section 4.4.1 we examine the diffusion limit. In the remaining subsections, we use benchmark problems to assess efficiency and accuracy of the method.

Remark 3. *For the angular hybrid DG-DG approach, some experiments were run with the angular resolution set to the same value for both the uncollided and collided equations. We acknowledge that this selection is mathematically equivalent to a standard DG method using the same angular resolution, but that the resultant hybrid does not have any computational advantages. It is included in some of the figures for the purpose of consistency.*

All numerical simulations use the discrete ordinates method with the product quadrature [6] to discretize the angular components of (4.1) or (2.48a). The number of ordinates used in any S_N simulation is $N^* = N^2$. All of the numerical simulations are performed on a reduced spatial geometry that assumes no variations in the z direction. In all cases, the domain is Cartesian, the mesh is square, and the DG elements are Q_1. The finite volume discretization uses a second-order reconstruction with slopes computed using only upwind information (see Appendix D). For time-dependent problems, a second-order strongly S-stable DIRK scheme is used (see Appendix A).

4.4.1 Diffusion Limit Test

We solve a steady-state version of (2.48a), for standard discretization, or (4.1), for hybrid discretizaiton, in x-y geometry with zero boundary condition on Γ^-, $\sigma_\mathrm{t} = 4.0$, $\sigma_\mathrm{a} = 0.5$, and $S(x, y, \Omega) = 900x^2y^2(1-x)^2(1-y)^2 m_{1,1}^2(\Omega)$, where $m_{1,1}(\Omega) = \sqrt{\frac{3}{4\pi}}\Omega_x$.

Using an S_8 angular discretization, we compare numerical results using standard DG, FV, and a DG-FV hybrid. We examine errors and order of convergence with respect to the spatial mesh h as ϵ varies, using a DG spatial discretization with $h = 1/256$ as a numerical reference. Results are shown in Tables 4.2–4.4. The DG-DG scheme maintains second-order convergence in h for large and small ϵ, although it loses order for intermediate values of ϵ.

Table 4.2: Diffusion limit: DG S_8.

	$h = 1/8$		$h = 1/16$		$h = 1/32$		$h = 1/64$		$h = 1/128$	
ϵ	Error	Ord.	Error	Ord.	Error	Ord.	Error	Ord.	Error	Ord.
1	8.55E-3	-	2.24E-3	1.93	5.90E-4	1.92	1.52E-4	1.95	3.71E-5	2.04
2^{-1}	8.05E-3	-	2.18E-3	1.88	6.34E-4	1.78	1.78E-4	1.83	4.51E-5	1.98
2^{-5}	1.22E-2	-	2.79E-3	2.12	6.63E-4	2.07	1.90E-4	1.80	6.57E-5	1.54
2^{-9}	1.40E-2	-	3.45E-3	2.02	8.41E-4	2.04	1.99E-4	2.08	4.28E-5	2.22
2^{-13}	1.42E-2	-	3.52E-3	2.01	8.72E-4	2.01	2.11E-4	2.05	4.62E-5	2.19

Table 4.3: Diffusion limit: FV S_8.

	$h = 1/8$		$h = 1/16$		$h = 1/32$		$h = 1/64$		$h = 1/128$	
ϵ	Error	Ord.	Error	Ord.	Error	Ord.	Error	Ord.	Error	Ord.
1	6.84E-2	-	1.69E-2	2.02	4.35E-3	1.96	1.10E-3	1.98	2.79E-4	1.98
2^{-1}	9.18E-2	-	2.17E-2	2.08	5.29E-3	2.03	1.31E-3	2.01	3.34E-4	1.98
2^{-5}	3.98E-1	-	9.53E-2	2.06	1.67E-2	2.51	2.79E-3	2.59	5.14E-4	2.44
2^{-9}	9.03E-1	-	5.67E-1	0.67	1.57E-1	1.86	1.40E-1	0.16	1.93E-1	-0.46
2^{-13}	9.93E-1	-	9.54E-1	0.06	8.39E-1	0.19	6.40E-1	0.39	8.84E-1	-0.47

Table 4.4: Diffusion limit: FV-DG S_8 S_8.

	$h = 1/8$		$h = 1/16$		$h = 1/32$		$h = 1/64$		$h = 1/128$	
ϵ	Error	Ord.	Error	Ord.	Error	Ord.	Error	Ord.	Error	Ord.
1	3.65E-2	-	1.01E-2	1.86	2.54E-3	1.98	6.35E-4	2.00	1.58E-4	2.00
2^{-1}	2.65E-2	-	6.53E-3	2.02	1.56E-3	2.07	3.81E-4	2.03	9.36E-5	2.02
2^{-5}	1.59E-2	-	3.36E-3	2.24	7.36E-4	2.19	1.97E-4	1.90	6.62E-5	1.58
2^{-9}	1.71E-2	-	3.96E-3	2.11	9.19E-4	2.11	2.09E-4	2.14	4.39E-5	2.25
2^{-13}	1.72E-2	-	4.03E-3	2.10	9.51E-4	2.08	2.22E-4	2.10	4.75E-5	2.22

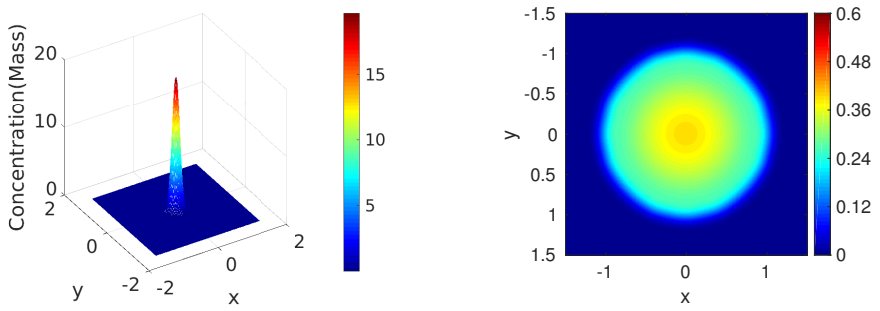
Reductions in order of this type are common in multiscale problems [37]. The finite volume method performs well for large values of ϵ, but the convergence is lost as ϵ gets smaller. As expected, the new FV-DG hybrid performs similarly to the DG-DG method, with a similar drop in convergence order for intermiate values of ϵ. However, errors for the FV-DG hybrid are 2–3 times larger than the DG-DG scheme for larger ϵ.

4.4.2 Linesource Benchmark

Efficiency of the spatial hybrid method

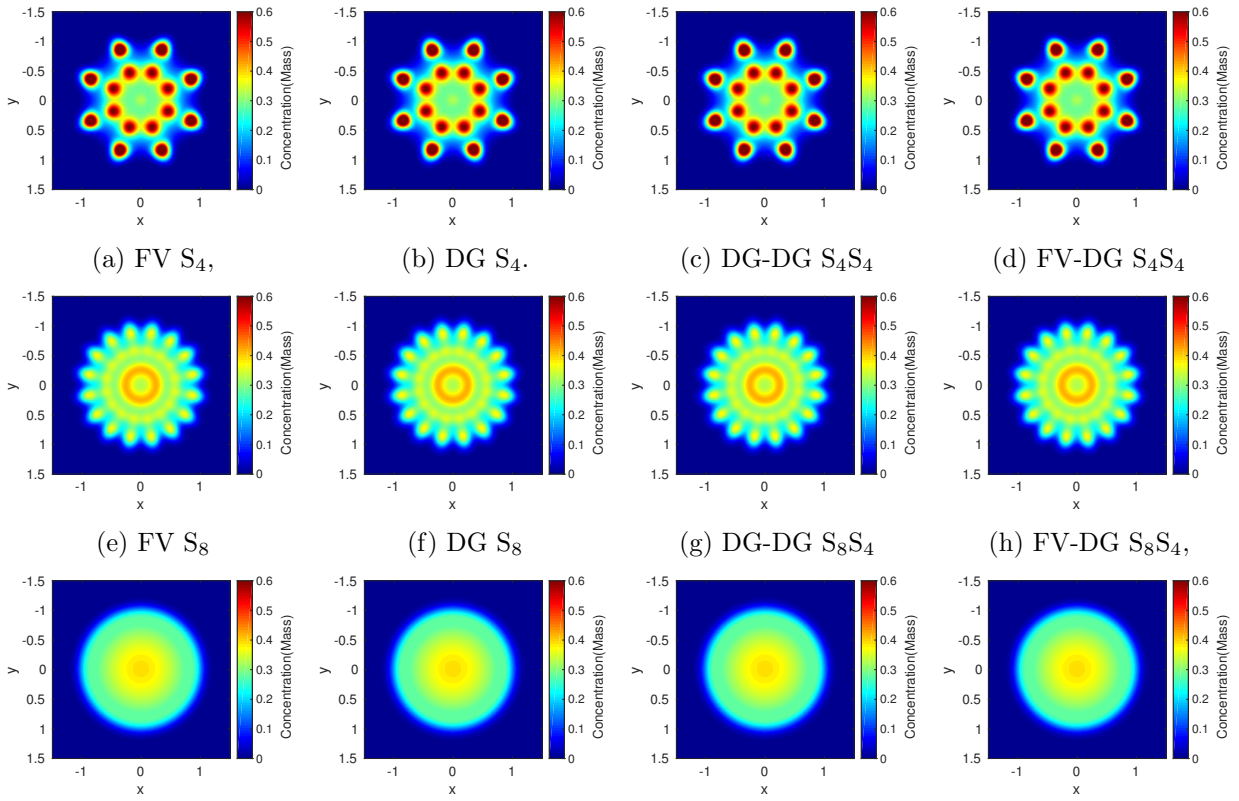
This experiment is meant to demonstrate the efficiency gains of spatial hybridization. We solve (2.48a) or (4.1) with $\epsilon = 1$ and approximate the initial condition (2.53) with small standard deviation $\beta = 0.09$. We consider problem with an absorption cross section $\sigma_a = 0$, scattering cross section $\sigma_s = 1$, source $q = 0$, and boundary condition $\psi_b = 0$. For reference, a semi-analytic solution is computed using the algorithm described in [24]; see Figure 4.1. We simulate the problem using a 301×301 grid on domain $[-1.5, 1.5] \times [-1.5, 1.5]$. The time step is $\Delta t = 5\Delta x$ and the final time $t = 1$. Several different orders of angular discretization are considered.

The results in Figure 4.2 show that the numerical solution changes dramatically based on the number discrete ordinates used. However, the choice of spatial discretization makes little difference in the qualitative solution. What is different is the computational time and memory usage. The quantities, both real and predicted, are depicted in Figure 4.3. Further details for the predicted values in Figure 4.3 are shown in Appendix C.



(a) Initial Condition (b) Semi-Analytic, $t = 1$. [24]

Figure 4.1: Initial Condition and Semi-Analytic Solution for $t = 1.0$, $\beta = 0.9$.



(a) FV S_4, (b) DG S_4. (c) DG-DG S_4S_4 (d) FV-DG S_4S_4

(e) FV S_8 (f) DG S_8 (g) DG-DG S_8S_4 (h) FV-DG S_8S_4,

(i) FV S_{24} (j) DG S_{24}, (k) DG-DG $S_{24}S_4$, (l) FV-DG $S_{24}S_4$

Figure 4.2: Discrete ordinates solutions of the line-source problem with various spatial discretizations.

Solutions to the line source problem with initial condition (2.53). Simulations are run on a 301×301 grid with time step $\Delta t = 5\Delta x$. Occurrence of ray-effects is strongly related to the angular resolution of the uncollided equations. Qualitative results of each method are similar, regardless of the angular resolution of the collided equations in the hybrid methods. Figure (c) is included here for consistency. See Remark 3.

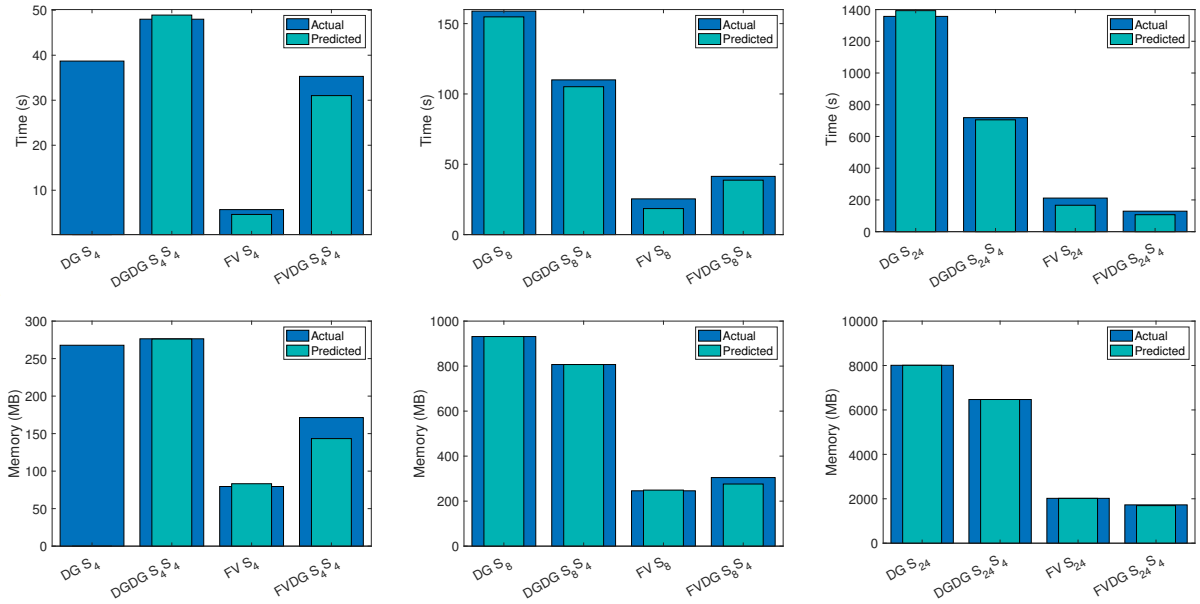


Figure 4.3: Wall time and maximum memory usages for solutions in Figure 4.2.

Wider bars (blue) represent the actual (measured) quantities in each graph for each method. For graphs showing time, internal timings were used within each program. An external script [72] was used to query the machine as to how much memory was being used for the method in question. The thinner bars (teal) were generated based on the predicted values that the method should take when compared to the DG equivalent. All predictions are based on the low resolution DG S_4 values. Timings and memory usage for DGDG S_4S_4 are included here for consistency. See Remark 3.

Accuracy of spatial hybrid method

In this experiment, we demonstrate that the computational advantages of the spatial hybrid demonstrated in Figure 4.3 can be leveraged using hybridization in angle to produce solutions with errors comparable to established methods in less run time. The numerical parameters used are the same as in Subsection 4.4.2, except $\beta = 0.045$ in (2.53) and $\Delta t = 3\Delta x$. Here we are able to use the computational speed gains that the spatial hybrid has over a DG method to produce a more accurate solution in less time.

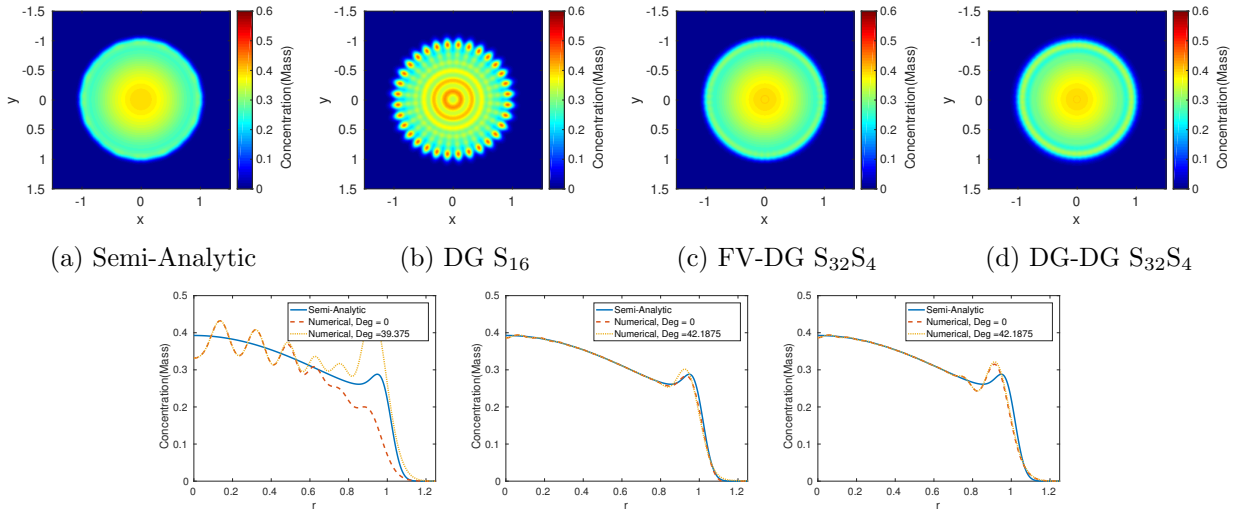
The results from Figure 4.2 have shown that the angular hybrid allows for more resolution in the uncollided equation and less resolution in the collided equation. We have used this to our advantage as shown in Figure 4.4. Table 4.5 shows that this strategy produces a better solution, but the run time is still a factor of two greater. The spatial hybrid brings down the run time significantly with comparable errors. In this case the errors are smaller, this result may not be the case in different experiments using different ratios of Δt to Δx.

4.4.3 Lattice Problem

In this experiment, we use a more real world example (see Subsection 2.5.2) to demonstrate that the spatial hybrid method is able to use hybridization in angle to produce solutions with errors comparable to established methods but in less time.

The initial data is void and the and boundary conditions are absorbing, i.e., $\psi_0 = 0$ and $\psi_\mathrm{b} = 0$. We simulate the problem with 504×504 spatial grid on the domain $[-3.5, 3.5] \times [-3.5, 3.5]$, set $\Delta t = 10\Delta x$, and run to a final time $t = 2.8$.

The solutions shown in Figure 4.5 and the results in Table 4.6 show again that the angular hybrid is capable of producing a better solution by increasing the resolution in the uncollided equation, without increasing the resolution of the collided equation. However, this comes at the cost of increased run time. The spatial hybrid is able to reduce the run time significantly while producing a solution with comparable errors.



(a) Semi-Analytic (b) DG S_{16} (c) FV-DG $S_{32}S_4$ (d) DG-DG $S_{32}S_4$

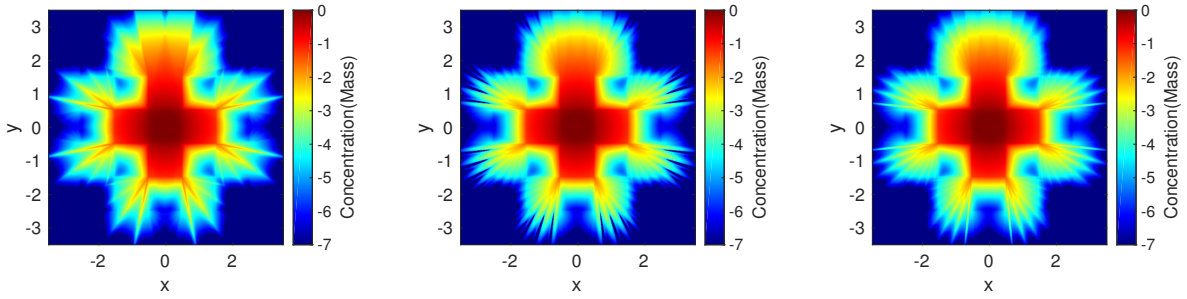
(e) DG S_{16} (f) FV-DG $S_{32}S_4$ (g) DG-DG $S_{32}S_4$

Figure 4.4: Accuracy comparison for solutions of the line-source problem.

Simulations are run on a 301×301 grid to $t = 1$. Top row: Particle densities. Bottom row: corresponding line-outs for numerical solutions only. The line-outs show plots along two angles: the x-axis, represented as Deg $= 0$, and in one of the direction where ray-effects appear in the corresponding graphs.

Table 4.5: Run times and errors for numerical solutions in Figure 4.4.

Method	DG S_{16}	DG-DG $S_{32}S_4$	FV-DG $S_{32}S_4$
Run time (mins)	14.8	30.6	5.0
L^2 Error	0.18	0.067	0.031
L^∞ Error	0.46	0.18	0.11



(a) DG S_8 (b) FV-DG $S_{16}S_4$ (c) DG-DG $S_{16}S_4$

Figure 4.5: Accuracy comparison for solutions of the lattice problem.

Solutions are on a logarithmic scale. Simulations are run on a 504×504 grid to $t = 2.8$.

Table 4.6: Run times and errors for numerical solutions in Figure 4.5.

Method	Reference	DG S_8	DG-DG $S_{16}S_4$	FV-DG $S_{16}S_4$
Run time (mins)	569.1	7.5	14.5	4.2
L^2 Error	-	0.0094	0.0029	0.0032
L^∞ Error	-	0.015	0.0062	0.011

Chapter 5

Future Works

Our current analysis for the convergence of the filtered discrete ordinates equations is based on the global L^2 norm. Our results also suggest convergence in the L^∞ norm; therefore, future work would include analysis in this norm. Also, the filter process has been based on the modified equations in [67], where the filtering depends on a polynomial expansion of the numerical solution. Determining a filtering process without the use of an angular polynomial basis is also the scope of future work.

As in [22], the current analysis shows only how the filter affects the asymptotic error in the solution. It does not show how or why the filter helps for low-order angular approximations, where it is most needed and useful. To this end, a more refined estimate can be derived in Theorem 3.8 by using the filter term on the right-hand side of (3.41a) to improve the stability of the operator on the left-hand side, rather than just removing the term altogether. Such an estimate would highlight the effect of the filter at low-order. The choice of filter strength is then a matter of balancing between damping and the size of the consistency error introduced in the source term of (3.42). This issue is common for regularized inverse problems; see for example [32]. A more detailed analysis of this balance, including strategies for local, automated tuning is the subject of future work. Currently the filter strength is first tuned by hand at low resolution (so that numerical solution are cheap) and then applied to simulations at higher resolution.

Our current analysis for the diffusion limit of the spatial hybrid method suggests that the choice of spatial discretization is inconsequential to the overall scheme achieving the diffusion

limit. Future work would include testing other spatial discretizations (finite difference, diamond difference, finite element) to see if these new combinations would be more efficient.

Additionally, although not shown in this dissertation, attempts have been made to combine a filtering process with the hybrid method. Results showed that the combination is ill-advised in the case of the isotropic line source benchmark where the order of the quadrature used in the uncollided equations N_{u} is significantly greater than the order of the quadrature used in the collided equations N_{c}. In the case where the degree of the filter operator is close the order of the quadrature in the collided equations $P \approx N_{\mathrm{c}} - 1$, the resultant solutions were similar to the solutions shown in Figure 3.2 (a), (b), and (c). However, this may not be the case with other problems, especially those with non-point sources or anisotropic scattering. Tests involving the conditions described would be the subject of future work.

Bibliography

[1] Abramowitz, M. and Stegun, I., editors (1964). *Handbook of mathematical functions with formulas, graphs, and mathematical tables*. National Bureau of Standards. 14

[2] Adams, M. L. (2001). Discontinuous finite element transport solutions in thick diffusive problems. *Nuclear science and engineering*, 137(3):298–333. 5, 22, 23, 24

[3] Alcouffe, R., Dautray, R., Forster, A., Ledanois, G., and (Eds.), B. M. (1985). A first collision source method for coupling monte carlo and discrete ordinates for localized source problems. *Photonics and Statistical Physics*, 240(352-366). 5, 59

[4] Alexander, R. (1977). Diagonally implicit runge-kutta method for stiff o.d.e.'s. *SIAM Journal on Numerical Analysis*, 14(6):1006–1021. 23, 61, 95

[5] Askew, J. (1972). Review of the status of collision probability methods. Technical report, UKAEA Reactor Group. 2

[6] Atkinson, K. (1982). Numerical intergration on the sphere. *H. Austral. Math. Soc.*, 23:332–347. 43, 75

[7] Atkinson, K. and Han, W. (2012). *Spherical Harmonics and Approximations on the Unit Sphere: An Introduction*. Springer. 3, 13, 14, 33, 41

[8] Aubin, T. (1982). *Nonlinear analysis on manifolds, Monge-Ampere equations*. Grundlehren der mathematischen Wissenschaften. 252. Springer-Verlag, New York. 33

[9] Bell, G. I. and Glasstone, S. (1970). Nuclear reactor theory. Technical report, US Atomic Energy Commission, Washington, DC (United States). 2

[10] Briggs, L., Jr., W. M., and Lewis, E. (1975). Ray-effect mitigation in discrete ordinate-like angular finite element approximations in neutron transport. *Nuclear Science and Engineering*, 57(3):205–217. 3, 4

[11] Brunner, T. and Holloway, J. (2005). Two-dimensional time-dependent riemann solvers for neutron transport. *Journal of Computational Physics*, 210:386–399. 25

[12] Canuto, C., Hussaini, M., Quarteroni, A., and Zang, T. (2006). *Spectral Methods: Fundamentals in Single Domain*. Springer. 20

[13] Carlson, B. (1970). Transport theory: Discrete ordinates quadrature over the unit sphere. Technical report, Los Alamos Scientific Lab. 43

[14] Carlson, B. G., Lathrop, K. D., et al. (1965). *Transport theory: the method of discrete ordinates*. Los Alamos Scientific Laboratory of the University of California. 3

[15] Case, K. and Zweifel, P. (1967). *Linear Transport Theory*. Addison-Wesley, Reading, MA. 1

[16] Chandrasekhar, S. (1960). *Radiative Transfer*. Dover Publications, Inc. 12

[17] Crockatt, M. (2018). Hybrid methods for radiation transport using integral deferred correction. 61

[18] Crockatt, M., Christlieb, A., Garrett, C. K., and Hauck, C. (2017). An arbitrary-order, fully implicit, hybrid kinetic solver for linear radiative transport using integral deferred correction. *Journal of Computational Physics*, 346:212–241. 5, 61, 96

[19] Dautray, R. and Lions, J.-L. (1984). *Mathematical Analysis and Numerical Methods for Science and Technology*, volume 6. Springer-Verlag, Berlin. 13

[20] Dolejsi, V. and Feistauer, M. (2015). *Discontinuous Galerkin Method: Analysis and Applications to Compressible Flow*. Springer. 96

[21] Emendorfer, D. (1974). Physics assumptions and applications of collision probabilities methods. In *ANS Conference on Mathematical Models and Computational Techniques for Analysis of Nuclear Systems, Ann Arbor, Michigan, CONF-730414-P2, US Atomic Energy Commission*. 2

[22] Frank, M., Hauck, C., and Kupper, K. (2016). Convergence of filtered spherical harmonic equations for radiation transport. *Communications in Mathematical Sciences*, 14(5):1443–1465. 4, 7, 8, 19, 20, 21, 40, 83

[23] Ganapol, B., Baker, R., Dahl, J., and Alcouffe, R. (2001). Homogeneous infinite media time-dependent analytical benchmarks. Technical report, Los Alamos National Laboratory. 3, 24

[24] Garrett, C. K. and Hauck, C. (2014). A comparison of moment closures for linear kinetic transport equations: the line source benchmark. *Transport Theory and Statistical Physics*, 42(6-7):203–235. 24, 25, 45, 77

[25] Gelbard, E. (1968). Spherical harmonics methods: P_L and double P_L approximations. *Computing methods in reactor physics*, page 271. 2

[26] Goertzel, G. and Kalos, M. H. (1958). Monte carlo methods in transport problems. *Progress in nuclear energy*, 2:315–369. 2

[27] Gosse, L. and Toscani, G. (2004). Asymptotic-preserving & well-balanced schemes for radiative transfer and the rosseland approximation. *Numerische Mathematik*, 98(2):223–250. 22

[28] Guermond, J.-L. and Kanschat, G. (2010). Asymptotic analysis of upwind discontinuous Galerkin approximation of the radiative transport equation in the diffusion limit. *SIAM Journal of Numerical Analysis*, 48(1):53–78. 5, 9, 22, 23, 24, 69

[29] Guo, B.-Y. (1998). *Spectral Methods and Their Applications*. World Scientific. 18, 32, 34

[30] Habetler, G. J. and Matkowsky, B. J. (1975). Uniform asymptotic expansions in transport theory with small mean free paths, and the diffusion approximation. *Journal of Mathematical Physics*, 16. 22

[31] Han, W., Huang, J., and Eichholz, J. A. (2010). Discrete-ordinate discontinuous Galerkin methods for solving the radiative transfer equation. *SIAM Journal of Scientific Computing*, 32(2). 3, 47, 62

[32] Hansen, P. C. and O'Leary, D. P. (1993). The use of the L-curve in the regularization of discrete ill-posed problems. *SIAM Journal on Scientific Computing*, 14(6):1487–1503. 83

[33] Hauck, C. D. and Lowrie, R. B. (2009). Temporal regularization of the P_N equations. *Multiscale Modeling & Simulation*, 7(4):1497–1524. 22

[34] Hesthaven, J. S., Gottlieb, S., and Gottlieb, D. (2007). *Spectral Methods for Time-Dependent Problems*, volume 21. Cambridge University Press. 18, 20

[35] Hill, T. (1975). Onetran: a discrete ordinates finite element code for the solution of the one-dimensional multigroup transport equation. Technical report, Los Alamos Scientific Lab., N. Mex.(USA). 3

[36] Hill, T. and Reed, W. H. (1976). Timex: A time-dependent explicit discrete ordinates program for the solution of multigroup transport equations with delayed neutrons. Technical report, Los Alamos Scientific Lab., N. Mex.(USA). 2

[37] Jin, S. (2010). Asymptotic preserving (ap) schemes for multiscale kinetic and hyperbolic equations: a review. *Lecture Notes for Summer School on "Methods and Models of Kinetic Theory"(M&MKT), Porto Ercole (Grosseto, Italy)*, pages 177–216. 77

[38] Jin, S. and Levermore, C. D. (1996). Numerical schemes for hyperbolic conservation laws with stiff relaxation terms. *Journal of computational physics*, 126(2):449–467. 22, 24

[39] Jin, S., Pareschi, L., and Toscani, G. (2000). Uniformly accurate diffusive relaxation schemes for multiscale transport equations. *SIAM Journal on Numerical Analysis*, 38(3):913–936. 22

[40] Kaplan, S. (1968). A new derivation of discrete ordinate approximations. *Nuclear Science and Engineering*, 34(1):76–82. 4

[41] Kaplan, S. (1969). Variational methods in nuclear engineering. In *Advances in Nuclear Science and Technology*, pages 185–221. Elsevier. 2

[42] Kaplan, S. and Davis, J. A. (1967). Canonical and involutory transformations of the variational problems of transport theory. *Nuclear Science and Engineering*, 28(2):166–176. 2

[43] Kavenoky, A. (1981). Status of integral transport theory. Technical report, CEA Centre d'Etudes Nucleaires de Saclay. 2

[44] Keepin, G. R. (1965). *Physics of nuclear kinetics*. Addison-Wesley Pub. Co. 1

[45] Keller, H. B. (1960). Approximate solutions of transport problems. II.convergence and applications of the discrete ordinate method. *J. Soc. Indust. Appl. Math.*, 8(1):43–73. 3

[46] Larsen, E. W. and Keller, J. B. (1974). Asymptotic solution of neutron transport problems for small mean free paths. *Journal of Mathematical Physics*, 15:75. 22

[47] Larsen, E. W. and Morel, J. E. (1989). Asymptotic solutions of numerical transport problems in opticaly thick, diffusive regimes II. *Journal of Computational Physics*, 83:212–236. 5, 22, 23, 24

[48] Larsen, E. W. and Morel, J. E. (2010). *Nuclear Computional Science*, chapter Advances in discrete-ordinates methodology. Springer. 3, 16

[49] Lathrop, K. (1971). Remedies for ray effects. *Nuclear Science and Engineering*, 45(3):255–268. 3, 4

[50] Lathrop, K. and Carlson, B. (1964). *Discrete Ordinates Angular Quadrature of the Neutron Transport Equation*. Los Alamos Scentific Laboratory. 3, 16

[51] Lebedev, V. (1976). Quadratures on a sphere. *USSR Computational Mathematics and Mathematical Physics*, 16(2):10–24. 43

[52] Lewis, E. (1981). Finite element approximation to the even-parity transport equation. In *Advances in Nuclear Science and Technology*, pages 155–225. Springer. 2

[53] Lewis, E. E. and W.F. Miller, J. (1993). *Computational Methods of Neutron Transport*. American Nuclear Society, La Grange Park, IL. 1, 2, 3, 10, 15, 16, 18

[54] Li, H.-S., Flamant, G., and Lu, J.-D. (2003). Mitigation of ray effects in the discrete ordinates method. *Numerical Heat Transfer, Part B: Fundamentals*, 43(5):445–466. 4

[55] Martin, W. R. and Duderstadt, J. J. (1977). Finite element solutions of the neutron transport equation with applications to strong heterogeneities. *Nuclear Science and Engineering*, 62(3):371–390. 2

[56] Martin, W. R., Yehnert, C. E., Lorence, L., and Duderstadt, J. J. (1981). Phase-space finite element methods applied to the first-order form of the transport equation. *Annals of Nuclear Energy*, 8(11-12):633–646. 2

[57] McClarren, R. G., Evans, T. M., Lowrie, R. B., and Densmore, J. D. (2008). Semi-implicit time integration for pn thermal radiative transfer. *Journal of Computational Physics*, 227(16):7561–7586. 23

[58] McClarren, R. G. and Hauck, C. D. (2010). Robust and accurate filtered spherical harmonics expansions for radiative transfer. *Journal of Computational Physics*. 3, 4, 7, 20

[59] McClarren, R. G. and Hauck, C. D. (2013). A collision-based hybrid method for time-dependent, linear, kinetic transport equations. *Multiscale Modeling and Simulation*, 11(4):1197–1227. 5, 9, 23, 59, 61

[60] Mezzacappa, A. and Bruenn, S. W. (1993a). A numerical method for solving the neutrino boltzmann equation coupled to spherically symmetric stellar core collapse. *The Astrophysical Journal*, 405:669–684. 1

[61] Mezzacappa, A. and Bruenn, S. W. (1993b). Stellar core collapse-a boltzmann treatment of neutrino-electron scattering. *The Astrophysical Journal*, 410:740–760.

[62] Mezzacappa, A. and Bruenn, S. W. (1993c). Type ii supernovae and boltzmann neutrino transport-the infall phase. *The Astrophysical Journal*, 405:637–668. 1

[63] Mihalis, D. and Weibel-Mihalis, B. (1999). *Foundations of Radiation Hydrodynamics*. Dover, Mineola, New York. 1

[64] Morel, J. E. (1989). A hybrid collocation-Galerkin-S_N method for solving the boltzmann transport equation. *Nuclear Science and Engineering*, 101:72–87. 3, 43

[65] Muller, C. (1966). *Spherical Harmonics*. Springer. 13

[66] Pomraning, G. C. (1973). *Radiation Hydrodynamics*. Pergamon Press, New York. 1

[67] Radice, D., Abdikamalov, E., Rezzolla, L., and Ott, C. D. (2013). A new spherical harmonics scheme for multi-dimensional radiation transport I. static matter configurations. *Journal of Computational Physics*, 242:648 – 669. 4, 7, 20, 21, 83

[68] Reed, W. and Hill, T. (1973). Triangular mesh methods for the neutron transport equation. 2, 3, 5

[69] Saad, Y. and Schultz, M. H. (1986). Gmres: A generalized minimal residual algorithm for solving nonsymmetric linear systems. *SIAM Journal on Scientific and Statistical Computing*, 7(3):856–869. 97

[70] Sanchez, R. and McCormick, N. J. (1982). A review of neutron transport approximations. *Nuclear Science and Engineering*, 80(4):481–535. 2

[71] Seibold, B. and Frank, M. (2014). StaRMAP - a second order staggered grid method for spherical harmonics moment equation of radiative transfer. *ACM Trans. Math. Softw.* 25

[72] Shin, J. (2010). memusg. https://gist.github.com/netj/526585. 79

[73] Spanier, J. and Gelbard, E. M. (2008). *Monte Carlo principles and neutron transport problems*. Courier Corporation. 2

[74] Sumiyoshi, K. and Yamada, S. (2012). Neutrino transfer in three dimensions for core-collapse supernovae. i. static configurations. *The Astrophysical Journal Supplement Series*, 199(1):17. 1

[75] Thurgood, C., Pollard, A., and Becker, H. (1995). The T_N quadrature set for the discrete ordinates method. *Journal of Heat Transfer*, 117:1068–1070. 4

[76] Vandeven, H. (1991). Family of spectral filters for discontinuous problems. *Journal of Scientific Computing*, 6(2):159–192. 20

[77] Victory, Jr, H. D. (1980). Convergence properties of discrete-ordinates solutions for neutron transport in three-dimensional media. *SIAM Journal of Numerical Analysis*, 17(1):71–83. 3

[78] Vladimirov, V. (1963). Mathematical problems in the one-velocity theory of particle transport. Technical report, Atomic Energy of Canada Limited. 2

[79] Zheng-Ming, L. and Brahme, A. (1993). An overview of the transport theory of charged particles. *Radiation Physics and Chemistry*, 41(4):673 – 703. 1

Appendices

A Second order diagonally implicit Runge-Kutta

We begin with the time discretization of (3.2). For simplicity we suppress the indices associated with the superscript Θ as well as the variable r, and we assume that $\mathbf{S}$, σ_{a}, and σ_{s} are independent of t. Then (3.2) can be written as

$$\partial_t \boldsymbol{\psi}(t) = \mathbf{G}(t, \boldsymbol{\psi}), \quad \mathbf{G}(t, \boldsymbol{\psi}) := \mathbf{S} - \Upsilon \boldsymbol{\psi}(t) - \sigma_{\mathrm{a}} \boldsymbol{\psi}(t) - \sigma_{\mathrm{s}}(R - Q)\boldsymbol{\psi}(t) + \sigma_{\mathrm{f}} F \boldsymbol{\psi}(t). \tag{1}$$

The temporal domain $[0, T]$ is divided into uniform intervals with time step Δt. Let n be the time step index. The numerical solution at time $t^n = n\Delta t$ is given by $\boldsymbol{\psi}^n \approx \boldsymbol{\psi}(t^n)$. It is obtained by using a second-order, diagonally implicit, Runge-Kutta (DIRK) method [4] with the following Butcher tableau:

$$\begin{array}{c|cc} c_1 & D_{1,1} & D_{1,2} \\ c_2 & D_{2,1} & D_{2,2} \\ \hline & b_1 & b_2 \end{array} = \begin{array}{c|cc} \gamma & \gamma & 0 \\ 1 & 1-\gamma & \gamma \\ \hline & 1-\gamma & \gamma \end{array}, \quad \gamma = 1 - \frac{1}{\sqrt{2}}. \tag{2}$$

Given $\boldsymbol{\psi}^n$, the stages are

$$\boldsymbol{\psi}^{(s)} = \boldsymbol{\psi}^n + \Delta t \sum_{r=1}^{s} D_{s,r} \mathbf{G}(t^n + c_r \Delta t, \boldsymbol{\psi}^{(r)}), \quad s = 1, 2, \tag{3}$$

and $\boldsymbol{\psi}^{n+1} = \boldsymbol{\psi}^{(2)}$. Rearranging (3) yields the steady-state form

$$\Upsilon \boldsymbol{\psi}^{(s)} - \sigma_{\mathrm{s}} Q \boldsymbol{\psi}^{(s)} - \sigma_{\mathrm{f}} F \boldsymbol{\psi}^{(s)} + \Lambda^{(s)} \boldsymbol{\psi}^{(s)} = \mathbf{q}^{(s)}, \tag{4}$$

where

$$\mathbf{q}^{(s)} = \mathbf{S} + \frac{1}{D_{s,s}} \left(\frac{\boldsymbol{\psi}^n}{\Delta t} + \sum_{r=1}^{s-1} D_{s,r} \mathbf{G}(t^n + c_r \Delta t, \boldsymbol{\psi}^{(r)}) \right), \quad \Lambda^{(s)} = \sigma_{\mathrm{s}} R + \left(\sigma_{\mathrm{a}} + \frac{1}{D_{s,s} \Delta_t} \right) I. \tag{5}$$

B Discontinuous Galerkin

Let (3.2) be discretized in time using the second-order DIRK method, (See Appendix A). For simplicity we drop the stage index (s) and right (4) in the form

$$\mathcal{L}\boldsymbol{\psi} = C\boldsymbol{\psi} + \mathbf{q}, \tag{6}$$

where $\mathcal{L} = \Upsilon + \Lambda$ and $C = \sigma_{\mathrm{s}} Q + \sigma_{\mathrm{f}} F$.

Let $M = \max\{D, P\}$, and let M^* be the number of moments associated with M. For a single ordinate Ω_i^N, (6) can be written as

$$\mathcal{L}_i \psi_i = (\mathcal{S}\mathcal{P}\boldsymbol{\psi})_i + q_i, \tag{7}$$

where $\mathcal{S}\mathcal{P} = C$ and

$$\mathcal{P}\colon \mathbb{R}^{N^*} \to \mathbb{R}^{M^*}, \quad \text{s.t.} \quad \mathcal{P}_{(\ell,k),i} = w_i^N m_{\ell,k}(\Omega_i^N), \tag{8a}$$

$$\mathcal{S}\colon \mathbb{R}^{M^*} \to \mathbb{R}^{N^*}, \quad \text{s.t.} \quad \mathcal{S}_{i,(\ell,k)} = (\, \sigma_{\mathrm{s}} \hat{g}_\ell + \sigma_{\mathrm{f}} f_\ell^P \,)\, m_{\ell,k}(\Omega_i^N). \tag{8b}$$

We discretize (7) using the discontinuous Galerkin (DG) method [20] with upwind numerical fluxes. Our presentation follows [18] closely. We assume X is a rectangular domain and create a uniform partition $\mathcal{T}_h$ of X with rectangular cells $\mathcal{K}$ of maximum length h and a fixed aspect ratio. We let $\mathcal{E}_h$ be the collection of interior edges in $\mathcal{T}_h$, and for each $e \in \mathcal{E}_h$, we let n_e be the associated normal vector with orientation chosen beforehand.

Let V_h be a finite dimensional function space over the spatial variable such that for each $v \in V_h$, $v|_{\mathcal{K}}$ is a polynomial for each $\mathcal{K} \in \mathcal{T}_h$. The DG method is then to find $\psi_i^h \in V_h$ for each $i \leq N^*$ satisfying

$$\mathcal{B}_i^h(\psi_i^h, v^h) = \bar{\mathcal{B}}^h(\mathcal{S}\mathcal{P}\boldsymbol{\psi}^h, v^h) + \mathcal{F}_i^h(v^h), \quad \forall\, v^h \in V_h, \tag{9}$$

where $\boldsymbol{\psi}^h = [\psi_1^h, \psi_2^h, \ldots, \psi_{N^*}^h]^T$. The bilinear and linear forms are given by

$$\begin{aligned}\mathcal{B}_i^h(u,v) = &-\sum_{\mathcal{K}\in\mathcal{T}_h}\int_{\mathcal{K}}\left(\Omega_i^N\cdot\nabla_x v\right)u\,dx+\sum_{\mathcal{K}\in\mathcal{T}_h}\int_{\mathcal{K}}\Lambda_{i,i}vu\,dx \qquad (10a)\\ &-\sum_{e\in\mathcal{E}_h}\int_e(\Omega_i^N\cdot n_e)[v]\widehat{u}\,dx+\sum_{e\in\partial X_i^+}\int_e(\Omega_i^N\cdot n_e)v^-u^-\,ds,\\ \bar{\mathcal{B}}^h(u,v) = &\sum_{\mathcal{K}\in\mathcal{T}_h}\int_{\mathcal{K}}vu\,dx, \qquad (10b)\\ \mathcal{F}_i^h(v) = &\sum_{\mathcal{K}\in\mathcal{T}_h}\int_{\mathcal{K}}q_i v\,dx-\sum_{e\in\partial X_i^-}\int_e\left(\Omega_i^N\cdot n_e\right)\psi_i^B v^-\,dx, \qquad (10c)\end{aligned}$$

where $\psi_i^B = \psi_i^B(x)$ is the inflow on ∂X_i^-,

$$v^{\pm}(x) = \lim_{\epsilon\to 0^+} v(x\pm\epsilon n_e),\quad [v] = v^+ - v^-,\quad \text{and}\quad \widehat{u}_i = \lim_{\epsilon\to 0^+} u_i(x-\epsilon\Omega_i^N). \qquad (11)$$

To formulate a matrix equation from (9), let $\mathbf{b}_h$ be a vector whose components form a basis of V_h, and let $\boldsymbol{\alpha}_i^h$ be a real vector such that $\psi_i^h = \mathbf{b}_h^T\boldsymbol{\alpha}_i^h$. Then (9) can be reduced to

$$\mathcal{L}^h\boldsymbol{\Psi}^h = \mathcal{S}^h\mathcal{P}^h\boldsymbol{\Psi}^h + \mathbf{q}^h, \qquad (12)$$

where $\mathcal{L}^h$ is a block diagonal matrix with blocks $\{\mathcal{L}_i^h\}_{i=1}^{N^*}$, $(\boldsymbol{\Psi}^h)^T = \left[(\boldsymbol{\alpha}_1^h)^T, (\boldsymbol{\alpha}_2^h)^T, \ldots, (\boldsymbol{\alpha}_{N^*}^h)^T\right]$, $(\mathbf{q}^h)^T = \left[(\mathbf{q}_1^h)^T, (\mathbf{q}_2^h)^T, \ldots, (\mathbf{q}_{N^*}^h)^T\right]$, $\mathcal{S}^h = \mathcal{S}\otimes\bar{\mathcal{S}}^h$, and

$$\mathcal{L}_i^h = \mathcal{B}_i^h(\mathbf{b}_h^T,\mathbf{b}_h),\quad \mathbf{q}_i^h = \mathcal{F}_i^h(\mathbf{b}_h),\quad \bar{\mathcal{S}}^h = \bar{\mathcal{B}}^h(\mathbf{b}_h^T,\mathbf{b}_h),\quad (\mathcal{P}^h\boldsymbol{\Psi}^h)_{\ell,k} = \sum_{i=1}^{N^*} w_i^N\boldsymbol{\alpha}_i^h m_{\ell,k}(\Omega_i^N). \qquad (13)$$

With $\boldsymbol{\Phi}^h := \mathcal{P}^h\boldsymbol{\Psi}^h$, (12) is equivalent to

$$\left(\mathbf{1}-\mathcal{P}^h(\mathcal{L}^h)^{-1}\mathcal{S}^h\right)\boldsymbol{\Phi}^h = \mathcal{P}^h(\mathcal{L}^h)^{-1}\mathbf{q}^h;\quad \boldsymbol{\Psi}^h = (\mathcal{L}^h)^{-1}\left(\mathcal{S}^h\boldsymbol{\Phi}^h+\mathbf{q}^h\right). \qquad (14)$$

Here we use GMRES [69], a Krylov solver for non-symmetric systems, to solve the first equation in (14) and obtain $\boldsymbol{\Phi}^h$. We then recover $\boldsymbol{\Psi}^h$ using the second equation.

C Computational scaling

In this appendix, we explain the details behind the numbers in Table 4.1 and the charts in Figure 4.3. All computational methods rely on four main subroutines: `source`(4.36), `integrate`(4.37), `copy`(4.40), and `sweep`(4.34). In the `source` subroutine, a known source function is used to compute a coefficient for every unknown in a mesh cell and every angle. In the `integrate over angle` subroutine, the angular unknowns associated to each spatial unknown are mapped to a single value. In the `copy` routine, a single value of each spatial unknown is copied across all angles. The `sweep` routine requires a matrix-vector product and the inversion of a linear system (both of a size equal to the number of unknowns) for every angle and mesh cell. When the cross-sections are constant, which we assume for the experiment in Subsection 4.4.2, the matrix used in the inversion can be pre-factored. The result is that the usual $O(n^3)$ operation count for an $n \times n$ matrix is reduced to $O(n^2)$, where n is the number of unknowns.

The cost of each of the subroutines above depends on the number of angles, number of mesh cells, and number of unknowns per mesh cell. In standard DG or FV codes, we use N^* angles and M cells. In hybrid DG-DG and FV-DG we use N_{u}^* and N_{c}^* points for the uncollided equations and collided equations respectively on M cells. FV methods will use one unknown per angle per cell for both quadrilateral and triangular cells, and DG methods will use 2^d unknowns for quadrilateral cells and $(d+1)$ unknowns for triangular cells where d is the dimension of the spatial domain. With these values the number of flops for each subroutine is given in Table C.1 and C.2. The results in Table 4.1 are obtained by summing across each row in Table C.1 or C.2.

Table C.1: Computational scaling leading orders per rectangular element.

	Source	Integrate	Copy	Sweep
FV	N^*	N^*	N^*	N^*
DG	$2^d N^*$	$2^d N^*$	$2^d N^*$	$2^{2d} N^*$
DG-DG	$2^d N_{\text{u}}^*$	$2^d (N_{\text{u}}^* + N_{\text{c}}^*)$	$2^d (N_{\text{u}}^* + N_{\text{c}}^*)$	$2^{2d} (N_{\text{u}}^* + N_{\text{c}}^*)$
FV-DG	N_{u}^*	$N_{\text{u}}^* + 2^d N_{\text{c}}^*$	$N_{\text{u}}^* + 2^d N_{\text{c}}^*$	$N_{\text{u}}^* + 2^{2d} N_{\text{c}}^*$

Table C.2: Computational scaling leading orders per triangular element.

	Source	Integrate	Copy	Sweep
FV	N^*	N^*	N^*	N^*
DG	$(d+1)N^*$	$(d+1)N^*$	$(d+1)N^*$	$(d+1)^2N^*$
DG-DG	$(d+1)N_{\mathrm{u}}^*$	$(d+1)(N_{\mathrm{u}}^*+N_{\mathrm{c}}^*)$	$(d+1)(N_{\mathrm{u}}^*+N_{\mathrm{c}}^*)$	$(d+1)^2(N_{\mathrm{u}}^*+N_{\mathrm{c}}^*)$
FV-DG	N_{u}^*	$N_{\mathrm{u}}^*+(d+1)N_{\mathrm{c}}^*$	$N_{\mathrm{u}}^*+(d+1)N_{\mathrm{c}}^*$	$N_{\mathrm{u}}^*+(d+1)^2N_{\mathrm{c}}^*$

To generate the predictions in Figure 4.3, we use the leading orders in Table C.1 and the knowledge of how many times each subroutine is called within a program, which is shown in Table C.3. Let T_{ref} be the minutes it takes to compute the standard DG reference, and let $\mathtt{n}_{\mathrm{so}}$, $\mathtt{n}_{\mathrm{int}}$, $\mathtt{n}_{\mathrm{cp}}$ and $\mathtt{n}_{\mathrm{sw}}$ be the total occurrences of the `source`, `integrate`, `copy`, and `sweep` subroutines respectively in the reference simulation. These are acquired by knowing either how many times these subroutines are performed in the code per iteration of the iterative solver or per time step. We assume the total number of time steps and iterations of the iterative solver are known. Then $T_{\mathrm{ref}} = (2^d\mathtt{n}_{\mathrm{so}} + 2^d\mathtt{n}_{\mathrm{int}} + 2^d\mathtt{n}_{\mathrm{cp}} + 2^{2d}\mathtt{n}_{\mathrm{sw}})kN^*M$, where k is an unknown conversion constant that is assumed to be independent of the type of method used. This implies that

$$kN^*M = \frac{T_{\mathrm{ref}}}{(2^d\mathtt{n}_{\mathrm{so}} + 2^d\mathtt{n}_{\mathrm{int}} + 2^d\mathtt{n}_{\mathrm{cp}} + 2^{2d}\mathtt{n}_{\mathrm{sw}})}, \tag{15}$$

Let $\mathtt{n}_{\mathrm{int,m}}$, $\mathtt{n}_{\mathrm{cp,m}}$ and $\mathtt{n}_{\mathrm{sw,n}}$ be the total occurrences of the `integrate`, `copy`, and `sweep` subroutines respectively with a loop structure involving N_{m}^* angles for $\mathrm{m} \in \{\mathrm{u}, \mathrm{c}\}$. With the constant k determined, we assume the total number of time steps and iterations of the iterative solver in the other simulations are the same as in the reference simulation. The predicted times for the other three methods in Figure 4.3 are calculated as follows:

$$T_{FV} = (\mathtt{n}_{\mathrm{so}} + \mathtt{n}_{\mathrm{int}} + \mathtt{n}_{\mathrm{cp}} + \mathtt{n}_{\mathrm{sw}})kN^*M \tag{16}$$

$$T_{\text{DG-DG}} = \left((\mathtt{n}_{\mathrm{so}} + \mathtt{n}_{\mathrm{int,u}} + \mathtt{n}_{\mathrm{cp,u}} + 2^d\mathtt{n}_{\mathrm{sw,u}})N_{\mathrm{u}}^* + (\mathtt{n}_{\mathrm{int,c}} + \mathtt{n}_{\mathrm{cp,c}} + 2^d\mathtt{n}_{\mathrm{sw,c}})N_{\mathrm{c}}^*\right)2^dkM \tag{17}$$

$$T_{\text{FV-DG}} = \left((\mathtt{n}_{\mathrm{so}} + \mathtt{n}_{\mathrm{int,u}} + \mathtt{n}_{\mathrm{cp,u}} + \mathtt{n}_{\mathrm{sw,u}})N_{\mathrm{u}}^* + 2^d(\mathtt{n}_{\mathrm{int,c}} + \mathtt{n}_{\mathrm{cp,c}} + 2^d\mathtt{n}_{\mathrm{sw,c}})N_{\mathrm{c}}^*\right)kM, \tag{18}$$

where N_{u}^* and N_{c}^* are known proportions of N^*. For the simulations in Figure 4.3, the values for the number of times each subroutine is performed is as follows.

Table C.3: Number of occurrences of subroutines used to compute solutions of the simulations in Figure 4.3.

$\mathtt{n}_{\mathrm{so}}$	$\mathtt{n}_{\mathrm{int}}$	$\mathtt{n}_{\mathrm{cp}}$	$\mathtt{n}_{\mathrm{sw}}$	$\mathtt{n}_{\mathrm{int,u}}$	$\mathtt{n}_{\mathrm{cp,u}}$	$\mathtt{n}_{\mathrm{sw,u}}$	$\mathtt{n}_{\mathrm{int,c}}$	$\mathtt{n}_{\mathrm{cp,c}}$	$\mathtt{n}_{\mathrm{sw,c}}$
42	124	124	166	42	42	84	124	124	124

The memory predictions are easier to compute. Using the DG S_4 simulation as a reference, we measure the maximum memory expenditure during the run of the simulation. We assume that the majority of the memory expenditure is taken up by the largest vectors in the code and we know ahead of time how many vectors are needed to run the simulation. During the run of the code, we require 4 vectors of size $2^d N^* M$ and 4 vectors of size $2^d M$ to hold various forms of the solution and source at every time step. During the iterative solver step a number of temporary vectors are created, one of size $2^d N^* M$ and 2 of size $2^d M$. Additionally, at iteration k of the GMRES solver it requires $k+1$ vectors of size $2^d M$ to construct the Krylov space. The maximum iterations the solver took was 2 throughout all our runs, so the code required an additional 3 vectors. The codes used eight bytes of memory (7.63×10^{-6} MB) for every entry in a vector and the product quadrature has $N^* = N^2$ ordinates. The computational domain uses $301 \times 301 = M$ mesh cells and $d = 2$. This leads to the following:

$$267.8 \text{ MB} = \left(5 * 2^d N^* M + 9 * 2^d M\right) * 7.63 \times 10^{-6} \text{ MB} + x \implies x = 21.7 \text{ MB}. \tag{19}$$

This x value is attributed to the overhead of the code and various other values that are held in memory that does not scale with M or N^*. We assume that this x value is relatively constant for every simulation we ran. To predict the other values in Figure 4.3 we simply count all the total entries from all relevant vectors, multiply by 8 bytes, ($7.63 * 10^{-6}$ MB), and then add x. The number of relevant vectors for each method is shown in Table C.4

Table C.4: Number and type of relevant vectors in each method.

Method	Relevant vectors
DG	5 vectors of size $2^d N^* M$, 9 vectors of size $2^d M$
FV	5 vectors of size $N^* M$, 9 vectors of size M
DG-DG	4 vectors of size $2^d N_{\mathrm{u}}^* M$, 1 vector of size $2^d N_{\mathrm{c}}^* M$, 12 vectors of size $2^d M$
FV-DG	4 vectors of size $N_{\mathrm{u}}^* M$, 1 vector of size $2^d N_{\mathrm{c}}^* M$, 12 vectors of size $2^d M$

D Finite volume second-order reconstruction

In this section, we give the specific form of reconstruction operator $\mathcal{R}$, (cf. (4.15)) for the calculations performed in Subsection 4.2, which were done in two-dimensional geometries using quadrilateral elements. Let $\mathcal{T}_h$ be a partition of $X \subset \mathbb{R}^2$ into $J^* \times K^*$ mesh cells. Let $C_{j,k} \in \mathcal{T}_h$ be a quadrilateral with cell center (x_j, y_k) and cell size $\Delta x \Delta y$ for all $1 \leq j \leq J^*$, $1 \leq k \leq K^*$. Let $N^* \in \mathbb{N}$ and let $\{\Omega_i\}_{i=1}^{N^*} \subset \mathbb{S}^2$ where $\Omega_i := (\Omega_{i,x}, \Omega_{i,y}, \Omega_{i,z})$. Let $f = [f_1, f_2, \ldots, f_{N^*}]^T \in (\mathcal{X}_{h,0})^{N^*}$, where $f_i \in \mathcal{X}_{h,0}$ for all $1 \leq i \leq N^*$. Denote the value of f_i on cell $C_{j,k}$ as $f_{i,j,k}$, and suppose $f_{\mathrm{b},i} \in L^2(\partial X)$ is a function on the boundary of X for each $1 \leq i \leq N^*$. Then

$$(Rf)_i(x,y)|_{C_{j,k}} = f_{i,j,k} + s^x_{i,j,k}(x - x_j) + s^y_{i,j,k}(y - y_k), \ \forall\ (x,y) \in C_{j,k}, \tag{20}$$

where

$$s^x_{i,j,k} = \begin{cases} \frac{f_{i,j,k} - f_{i,j-1,k}}{\Delta x}, & \Omega_{i,x} \geq 0,\ j > 1, \\ \frac{2(f_{i,j,k} - f_{i,j-1/2,k})}{\Delta x}, & \Omega_{i,x} \geq 0,\ j = 1, \\ \frac{f_{i,j,k} - f_{i,j+1,k}}{\Delta x}, & \Omega_{i,x} \leq 0,\ j < J^*, \\ \frac{2(f_{i,j,k} - f_{i,j+1/2,k})}{\Delta x}, & \Omega_{i,x} \leq 0,\ j = J^*, \end{cases} \tag{21}$$

and $s^y_{i,j,k}$ is defined similarly. The boundary terms are defined as

$$f_{i,1/2,k} = f_{\mathrm{b,i}}(x_{1/2}, y_k), \quad f_{i,J^*+1/2,k} = f_{\mathrm{b,i}}(x_{J^*+1/2}, y_k), \tag{22a}$$

$$f_{i,j,k-1/2} = f_{\mathrm{b,i}}(x_j, y_{1/2}), \quad f_{i,j,K^*+1/2} = f_{\mathrm{b,i}}(x_j, y_{K^*1/2}). \tag{22b}$$

Vita

Vincent Heningburg was born in Tokyo, Japan and raised in Prattville, AL to parents Darryl and Joanne Heningburg. He is the second child of four: Katherine, Octavia, and Avory. He attended Prattville High School in Prattville, AL. After graduation, he attended the Georgia Institute of Technology until December 2004. He soon after joined the United States Marine Corps. In 2008, he enrolled in Auburn University in Montgomery. He obtained a Bachelors of Science degree from AUM in December 2010 in Mathematics with an option in Computer Science. In 2012, he accepted a Distinguished Graduate Fellowship from the Bredesen Center to attend the University of Tennessee, Knoxville, in the Mathematics program. In 2013, he was honorably discharged from the Marine Corps. Vincent graduated with a Masters of Science degree in Mathematics in December 2014. He is set to graduate with a Doctorate of Philosophy in Mathematics in August 2019.