A Family of Independent Variable Eddington Factor Methods with Efficient Linear Solvers

Samuel Olivier\textsuperscript{a,}\*, Will Pazner\textsuperscript{b}, Terry S. Haut\textsuperscript{b}, Ben C. Yee\textsuperscript{b}

\textsuperscript{a}Applied Science \& Technology, University of California, Berkeley, Berkeley, CA 94708, United States of America
\textsuperscript{b}Lawrence Livermore National Laboratory, 7000 East Avenue, Livermore, CA 94550, United States of America

Abstract

We present a family of discretizations for the Variable Eddington Factor (VEF) equations that have high-order accuracy on curved meshes and efficient preconditioned iterative solvers. The VEF discretizations are combined with the Discontinuous Galerkin transport discretization from [1] to form an effective high-order, linear transport method. The VEF discretizations are derived by extending the unified analysis of Discontinuous Galerkin methods for elliptic problems presented by Arnold et al. [2] to the VEF equations. This framework is used to define analogs of the interior penalty, second method of Bassi and Rebay, minimal dissipation local Discontinuous Galerkin, and continuous finite element methods. The analysis of subspace correction preconditioners [3], which use a continuous operator to iteratively precondition the discontinuous discretization, is extended to the case of the non-symmetric VEF system. Numerical results demonstrate that the VEF discretizations have arbitrary-order accuracy on curved meshes, preserve the thick diffusion limit, and are effective on a proxy problem from thermal radiative transfer in both outer transport iterations and inner preconditioned linear solver iterations. In addition, a parallel weak scaling study of the interior penalty VEF discretization demonstrates the scalability of the method out to 1152 processors.

Keywords: Variable Eddington Factor, Discontinuous Galerkin

1. Introduction

The Variable Eddington Factor (VEF) method [4], also known as Quasidiffusion (QD) [5], is a rapidly-converging nonlinear scheme for solving the Boltzmann transport equation, a crucial component of high energy density physics (HEDP) simulations, nuclear reactor analysis, and medical physics. VEF has been applied to a wide range of transport and multiphysics problems including (but not limited to) nuclear reactor eigenvalue problems [6], nuclear reactor kinetics [7], and thermal radiative transfer (TRT) [8]. It performs well in problems having both optically thick and thin regions and treats anisotropic scattering equally well [9, 10]. Robust convergence is achieved by iteratively coupling the transport equation to the VEF equations, a moment-based equivalent reformulation of transport. The exact closures used to form the VEF equations are weak functions of the solution meaning even simple iterative schemes, such as fixed-point iteration, converge rapidly.

VEF offers significant algorithmic flexibility in that any valid discretization of the VEF equations will yield a rapidly-converging algorithm. This is in stark contrast to Diffusion Synthetic Acceleration (DSA) which places severe restrictions on the discretization of the moment equations in order to guarantee stability [11]. In the case where the VEF and transport discretizations are not algebraically consistent, referred to as a VEF method with an “independent” discretization [12, 13], the discrete solutions of the transport and VEF equations will differ on the order of the discretization error and will be equivalent only in the limit as the mesh is refined. However, even in an under-resolved problem, VEF still produces a “transport

\*Corresponding author

Preprint submitted to Elsevier November 25, 2021
solution" in that the solution of the VEF method is a discrete solution of an equivalent reformulation of the transport equation. Furthermore, VEF methods generally preserve the thick diffusion limit [14] and have conservation even if the transport discretization in isolation does not. These properties are particularly useful in multiphysics calculations since the lower-dimensional VEF equations can be directly coupled to the other physics components in place of the high-dimensional transport equation. In addition, discretizations for the transport and VEF equations can be designed independently so that they are in some sense optimal for their intended uses.

This flexibility has been exploited to improve efficiency in relation to all seven dimensions of the transport equation. Ghassemi and Anistratov [15] showed that different order temporal discretizations can be applied to the transport and VEF equations. Ongoing work suggests that time-stepping stability and accuracy can be maintained when just one transport inversion is performed per time step [16]. Anistratov and Coale [17] used data compression techniques to reduce storage costs in time-dependent calculations. In astrophysics, VEF is used to simplify the implementation of coupling TRT to hydrodynamics and to avoid the memory cost of solving the time-dependent transport equation [18, 19, 20]. Davis et al. [21] used a short characteristics discretization of the transport equation. Olivier and Morel [22] and Lou et al. [23] designed a spatial discretization of the VEF equations to increase multiphysics compatibility. Yee et al. [24] showed that robust convergence is maintained even when positivity-preserving methods are used inside the iteration. Anistratov [25] solved the multigroup TRT equations by using a VEF method with multiple levels in frequency. It is also well-known that the multigroup eigenvalue problem can be solved with only the need for eigenvalue iterations on the one-group VEF equations [26].

The above techniques rely on the efficient solution of the discretized VEF equations. VEF methods reduce the overall cost of the simulation by trading inversions of the high-dimensional transport equation with inversions of the lower-dimensional VEF equations. In all of VEF’s applications, the solution of the discretized VEF equations is buried under multiple nested loops corresponding to time integration, Newton iterations, eigenvalue iterations, multi-group iterations, and/or fixed-point iterations. The efficient iterative solution of the VEF equations is then crucial to the efficiency of the overall algorithm and is a prerequisite for the practicality of any VEF method.

The unusual structure of the VEF equations and their lack of self-adjointness make the development of discretizations and their corresponding preconditioned iterative solvers difficult. While considerable effort has been placed into the discretization of the VEF equations, to our knowledge, existing methods either rely on expensive and unscaleable preconditioners such as block incomplete LU (BILU) factorization, cannot be solved with iteration counts independent of the mesh size, or do not mention solvers entirely. Previous work on discretizing the VEF equations includes finite volume [9, 12, 27, 18, 28], finite difference [29], mixed finite element [30, 22, 31, 23], continuous finite element [32, 13], and discontinuous finite element [33] techniques. Most VEF methods are designed to be algebraically consistent with their application’s discretized transport equation which typically requires discretizing the first-order form of the VEF equations. Such discretizations solve for both the zeroth and first moment of the solution and thus have significantly more unknowns than discretizations of the second-order form. In addition, block preconditioners [34] are required to efficiently solve discretizations of the first-order form. Such solvers can require nested iteration for robustness (see [35] for a radiation diffusion example).

Warsa and Anistratov [13] showed that VEF methods with and without algebraic consistency converge equivalently as long as the transport data is properly represented. In particular, computing the Eddington tensor and boundary factor using finite element interpolation and Discrete Ordinates (S\text{N}) angular quadrature enables rapid convergence for any valid discretization of the VEF equations. An independent discretization of the second-order form of the VEF equations then has the potential to provide the rapid convergence of a consistent VEF method while solving for fewer unknowns and avoiding the need for block preconditioners. Such a method also has the flexibility to discretize the VEF equations in a manner that can leverage existing linear solver technology.

Our motivation for this research is in the context of HEDP experiments where the tightly-coupled simulation of hydrodynamics and TRT is required, the latter of which typically includes the S\text{N} transport equation. For hydrodynamics, it has been shown that, compared to low-order methods, high-order methods on curved meshes have improved robustness, symmetry preservation, and strong scaling on emerging high
performance computer architectures [36, 37, 38]. Transport methods compatible with this multiphysics framework are desired. Haut et al. [1] showed that adequately approximating realistic meshes generated from a high-order hydrodynamics code as straight-edged required a significant number of mesh refinements leading to an impractical increase in transport unknowns. It is also possible that high-order accurate transport methods could be beneficial in terms of multiphysics compatibility with high-order hydrodynamics. High-order transport methods compatible with curved meshes have been developed recently in [39, 1] with corresponding consistent DSA discretizations in [40, 41]. However, high-order discretizations of the VEF equations compatible with curved meshes have not yet been developed.

In this paper, we design a family of independent VEF discretizations for the linear, steady-state transport problem that can be efficiently and scalably solved with high-order accuracy, in multiple dimensions, and on curved meshes. Our approach is to begin with discretization techniques known to have effective preconditioners on the simpler case of radiation diffusion (i.e. the model Poisson problem) and adapt them to the VEF equations. By using the Eddington tensor and boundary factor interpolation procedure established in [13], these methods achieve both rapid convergence in outer fixed-point iterations and in inner linear solver iterations when paired with a high-order Discontinuous Galerkin (DG) discretization of $S_N$ transport.

In particular, we extend the unified analysis of DG methods developed for elliptic problems presented by Arnold et al. [2] to the VEF equations to derive analogues of the interior penalty (IP), second method of Bassi and Rebay (BR2), Minimal Dissipation Local Discontinuous Galerkin (MDLDG), and continuous finite element (CG) techniques. We show that the IP and BR2 VEF methods are effectively preconditioned by the subspace correction method from Pazner and Kolev [3] and that Algebraic Multigrid (AMG) is effective for the CG and MDLDG discretizations. Anistratov and Warsa [33] also applied DG techniques to the VEF equations but they discretize the first-order form of the VEF equations and only consider lowest-order elements in one dimension. We note that our CG operator is equivalent to extending the continuous finite element VEF discretization in [13] to multiple dimensions, arbitrary-order, and curved meshes.

The paper proceeds as follows. First, we describe the VEF method analytically and discuss iterative schemes to solve the coupled transport-VEF system. Then, we provide background on representing high-order meshes and finite element solutions and present the mathematical notation that will be used in the remainder of the paper. We derive the extension of the unified framework for DG to the VEF equations. The IP, BR2, MDLDG, and CG VEF discretizations are derived from this framework. Section 6 discusses the design and analysis of the subspace correction preconditioners, and extends their analysis to the case of non-symmetric linear systems. We next give computational results. We show that all the methods presented achieve high-order accuracy on curved meshes through the method of manufactured solutions, preserve the thick diffusion limit both on orthogonal and a severely distorted curved mesh, and are effective on the linearized, steady-state crooked pipe problem, a challenging proxy problem from TRT, in both outer fixed-point iterations and inner linear solver iterations. We also present a parallel weak scaling study for the IP discretization which demonstrates the scalability of the algorithm out to 1152 processors and 10 million VEF scalar flux unknowns. Finally, we give conclusions and recommendations for future work.

2. The VEF Method

The steady-state, mono-energetic, fixed-source transport problem with isotropic scattering and inflow boundary conditions is:

\begin{align}
\nabla \cdot \psi + \sigma_t \psi &= \frac{\sigma_s}{4\pi} \int \psi \, d\Omega' + q, \quad x \in D, \\
\psi(x, \Omega) &= f(x, \Omega), \quad x \in \partial D \text{ and } \Omega \cdot \hat{n} < 0,
\end{align}

where $\psi(x, \Omega)$ is the angular flux, $\Omega \in S^2$ the direction of particle flow, $D$ the spatial domain of the problem with $\partial D$ its boundary, $\sigma_t(x)$ and $\sigma_s(x)$ the total and scattering macroscopic cross sections, respectively, $q(x, \Omega)$ the fixed-source, and $f(x, \Omega)$ the inflow boundary function. The VEF equations are given by

\begin{align}
\nabla \cdot J + \sigma_t \varphi &= Q_0, \\
\nabla \cdot (E \varphi) + \sigma_t J &= Q_1,
\end{align}

where $J = J(x, \Omega)$ is the current density, $E = E(x, \Omega)$ is the electric field,

$\varphi = \psi |_{\partial D}$ is the outer flux.
where $\sigma_a(x) = \sigma_t(x) - \sigma_s(x)$ is the absorption macroscopic cross section, $\varphi(x)$ and $J(x)$ the zeroth and first angular moments of the angular flux, and

$$E(x) = \frac{\int \Omega \otimes \Omega \psi \, d\Omega}{\int \psi \, d\Omega}$$

is the Eddington tensor. We refer to $\varphi(x)$ as the scalar flux and $J(x)$ as the current. In addition, $Q_i = \int \Omega' q \, d\Omega$ are the angular moments of the fixed-source, $q$. The VEF equations are derived by taking the zeroth and first angular moments of the transport equation and closing the second moment of the angular flux, $P = \int \Omega \otimes \Omega \psi \, d\Omega$, with

$$P = E \varphi.$$ (4)

By eliminating the current, the VEF equations can be cast as a drift-diffusion equation:

$$-\nabla \cdot \left( \frac{1}{\sigma_t} \nabla (E \varphi) \right) + \sigma_a \varphi = Q_0 - \nabla \cdot \frac{Q_1}{\sigma_t}.$$ (5)

In both the first-order form (Eq. 2) and second-order form (Eq. 5), the presence of the Eddington tensor inside the divergence leads to diffusion, advection, and reaction-like terms that make applying existing discretization techniques difficult.

The Miften-Larsen transport-consistent boundary conditions [27] are

$$J \cdot \hat{n} = 2g + E_b \varphi, \quad x \in \partial D$$

(6)

where

$$g(x) = \int_{\Omega \cdot \hat{n} < 0} \Omega \cdot \hat{n} f(x, \Omega) \, d\Omega$$

(7)

is the incoming partial current computed from the transport boundary inflow function and

$$E_b = \frac{\int |\Omega \cdot \hat{n}| \psi \, d\Omega}{\int \psi \, d\Omega}$$

(8)

is the Eddington boundary factor. This boundary condition is derived by manipulating partial currents and using an analogous nonlinear closure. In equations, with the partial currents defined as $J_n^\pm = \int_{\Omega \cdot \hat{n} \gtrless 0} \Omega \cdot \hat{n} \psi \, d\Omega$,

$$J \cdot \hat{n} = J_n^+ + J_n^- = (J_n^+ - J_n^-) + 2J_n^- = \int |\Omega \cdot \hat{n}| \psi \, d\Omega + 2J_n^- \rightarrow E_b \varphi + 2J_n^-,$$ (9)

where $g$ in Eq. 6 plays the role of $J_n^-$ using the transport equation’s inflow boundary condition.

If the Eddington tensor and boundary factor are known, the VEF equations define the zeroth and first moments of the angular flux. In other words, the VEF equations with Miften-Larsen boundary conditions are an equivalent reformulation of the transport equation. However, this is a trivial closure in that the solution to the transport equation must already be known to define the VEF data. VEF methods rely on the fact that the VEF data are weak functions of the angular flux and thus simple iterative schemes can converge rapidly.

Note that when an independent discretization is used for the VEF equations, the discretized VEF scalar flux and VEF current will not be equivalent to the zeroth and first angular moments of the discrete angular flux; the two solutions will differ on the order of the spatial discretization error. To notationally separate the two scalar flux solutions, we use $\varphi$ (varphi) to denote the VEF scalar flux and $\phi = \int \psi \, d\Omega$ (phi) as the zeroth moment of the angular flux.
VEF methods seek the solution of the nonlinearly coupled system of equations:

\[ \Omega \cdot \nabla \psi + \sigma_t \psi = \frac{\sigma_s}{4\pi} \varphi + q, \quad (10a) \]

\[ -\nabla \cdot \frac{1}{\sigma_t} \nabla \cdot (E \varphi) + \sigma_a \varphi = Q_0 - \nabla \cdot \frac{Q_1}{\sigma_t}, \quad (10b) \]

where the drift-diffusion form of VEF is used for brevity. Boundary conditions are specified by Eqs. 1b and 6 for the transport and VEF drift-diffusion equation, respectively. Here, the transport equation’s scattering source is now coupled to the VEF drift-diffusion equation and the data for the VEF drift-diffusion equation are nonlinearly coupled to the transport equation. We have increased the complexity of the problem by adding the VEF scalar flux as an additional unknown and by casting the linear transport problem as nonlinear. However, properties of the VEF data allow this nonlinear, coupled system to be solved more efficiently.

Let

\[ L\psi = \Omega \cdot \nabla \psi + \sigma_t \psi, \quad (11) \]

\[ R(\psi)\varphi = -\nabla \cdot \frac{1}{\sigma_t} \nabla \cdot (E(\psi)\varphi) + \sigma_a \varphi, \quad (12) \]

be the streaming and collision operator and VEF drift-diffusion operator, respectively, where (·) indicates a nonlinear dependence on the argument. By linearly eliminating the angular flux, the transport-VEF system is equivalent to

\[ R \left( L^{-1} \left( \frac{\sigma_s}{4\pi} \varphi + q \right) \right) \varphi = Q_0 - \nabla \cdot \frac{Q_1}{\sigma_t}. \quad (13) \]

Applying the inverse of the drift-diffusion operator, we see that the solution of the coupled transport-VEF system is the fixed-point:

\[ \varphi = G(\varphi) \quad (14) \]

where

\[ G(\varphi) = R \left( L^{-1} \left( \frac{\sigma_s}{4\pi} \varphi + q \right) \right)^{-1} \left( Q_0 - \nabla \cdot \frac{Q_1}{\sigma_t} \right). \quad (15) \]

The fixed-point operator \( G \) is applied in two stages: 1) solve the transport equation using a scattering source formed from the VEF scalar flux and 2) solve the VEF drift-diffusion equation using the VEF data computed from the angular flux from stage 1).

The simplest algorithm to solve Eq. 14 is fixed-point iteration:

\[ \varphi^{k+1} = G(\varphi^k) \quad (16) \]

where \( \varphi^0 \) is an initial guess for the solution. This process is repeated until the difference between successive iterates is small enough. Since the Eddington tensor and boundary factor are weak functions of the angular flux even this simple iteration strategy will converge rapidly.

Iterative efficiency can be improved with the use of Anderson acceleration. Anderson acceleration defines the next iterate as the linear combination of the previous \( m \) iterates that minimizes the residual \( \varphi - G(\varphi) \). For the storage cost of \( m \) previous iterates, Anderson acceleration increases the convergence rate and improves robustness. While it is not practical to store multiple copies of the angular flux, it is reasonable to expect that a small set of scalar flux-sized vectors can be stored. The process of linearly eliminating the transport equation, codified in Eq. 13, allows the Anderson space to be built from the much smaller scalar flux-sized vectors only. In the case where a subset of the angular flux unknowns are not eliminated, such as when a parallel block Jacobi sweep is used to avoid communication costs or when mesh cycles or reentrant faces are present, the solution vector can be augmented with these un-eliminated unknowns so that they are included in the Anderson space. This is the nonlinear analog to the ideas used for Krylov-accelerated source iteration [42].
In addition, defining the nonlinear residual as

$$F(\varphi) = \varphi - G(\varphi) = 0,$$

(17)

root-finding methods such as Jacobian-free Newton Krylov (JFNK) can be used. JFNK builds a new Krylov space to approximate the gradient of $F$ at each iteration meaning information across iterations is not kept. JFNK typically required significantly more evaluations of $G$ than Anderson-accelerated fixed-point iteration. Thus, we present results using fixed-point iteration and Anderson-accelerated fixed-point iteration only.

The following sections present the discretizations and solvers needed to efficiently evaluate $G$ numerically.

3. Mesh and Finite Element Preliminaries

3.1. Description of the Mesh

Let $\mathcal{D} \in \mathbb{R}^{\text{dim}}$ with $\text{dim} = 2, 3$ be the domain of the problem. Consider the tessellation

$$\mathcal{D} = \bigcup_{K_e \in \mathcal{T}} K_e$$

with $K_e$ the $e^{th}$ element in the mesh $\mathcal{T}$. Each coordinate of the mesh is represented by a piecewise continuous polynomial. In other words, the mesh itself is a member of an $[\mathcal{H}^1(\mathcal{D})]^{\text{dim}}$ finite element space. This allows representation of curved surfaces and enforces continuity of the mesh coordinates along the interfaces between elements. Figure 1a depicts a mesh of two quadratic, quadrilateral elements where the mesh control points labeled 2, 7, and 12 are shared between the two elements to enforce continuity of the shared interior interface between them.

The mesh element $K_e$ is given as the image of the reference element $\hat{K}$ under an invertible, polynomial mapping $T_e : \hat{K} \rightarrow K_e$ where $T_e \in [\mathcal{P}_m(\hat{K})]^{\text{dim}}$ for simplicial elements (triangles and tetrahedra) or $T_e \in [\mathcal{Q}_m(\hat{K})]^{\text{dim}}$ for tensor product elements (quadrilaterals and hexahedra). Here, $\mathcal{P}_m(\hat{K})$ is the space of all polynomials of total degree at most $m$ in all variables and $\mathcal{Q}_m(\hat{K})$ the space of all polynomials of degree at most $m$ in each variable. For example, in two dimensions,

$$\mathcal{P}_1(\hat{K}) = \{1, \xi_1, \xi_2\}$$

(18)

while

$$\mathcal{Q}_1(\hat{K}) = \{1, \xi_1, \xi_2, \xi_1\xi_2\}.$$  

(19)

We do not consider the use of $\mathcal{P}_m(\hat{K})$ on tensor-product elements for either the mesh or the solution.

The reference element is the unit dim-simplex for simplicial elements (i.e. a triangle with coordinates (0,0), (1,0), and (0,1)) or the unit dim-cube $\hat{K} = [0, 1]^{\text{dim}}$ for tensor product elements. Figure 1b depicts a mesh transformation for a non-affine, linear, quadrilateral element. In the remainder of this document, we assume the use of tensor product elements however the derivations apply analogously to simplicial elements.

Let $\xi \in \hat{K}$ denote the reference coordinate. The Jacobian matrix of the mapping is

$$\mathbf{F}_e = \frac{\partial T_e}{\partial \xi} \in \mathbb{R}^{\text{dim} \times \text{dim}}.$$  

(20)

Further define $J_e = |\mathbf{F}_e|$ as the determinant of the Jacobian matrix. As an example, the transformation, Jacobian matrix, and determinant for the transformation depicted in Fig. 1b are

$$\mathbf{T} = \begin{bmatrix} h\xi_1 + \alpha \xi_2 (2\xi_1 - 1) \\ h \xi_2 \end{bmatrix}, \quad \mathbf{F} = \begin{bmatrix} 2\alpha \xi_2 + h & \alpha (2\xi_1 - 1) \\ 0 & h \end{bmatrix}, \quad J = 2\alpha \xi_2 h + h^2.$$  

(21)

The mesh transformations are used to perform integration in reference space using:

$$\int (\cdot) \, d\mathbf{x} = \int_{\hat{K}} (\cdot) \, J_e d\xi,$$

(22)
\[ \nabla_x = F^{-T} \nabla_\xi. \] (23)

Integration over surfaces is performed over the dim - 1 dimensional reference element using the transformed element of surface area. In this document, integration over the domain \( \mathcal{D} \) is implicitly performed using numerical quadrature and the relations in Eqs. 22 and 23. Finally, the characteristic mesh length, \( h \), is computed with

\[ h_e = \left( \int_K J_e \, d\xi \right)^{1/\text{dim}}, \] (24)

with \( h = \max_{K_e \in \mathcal{T}} h_e \).  

### 3.2. Finite Element Spaces

On each element, we will seek solutions to the transport and VEF drift-diffusion equations in the space of polynomials mapped from the reference element \( \hat{K} \) defined by

\[ \mathbb{Q}_p(K_e) = \{ u = \hat{u} \circ T_e^{-1} : \hat{u} \in \mathbb{Q}_p(\hat{K}) \}, \] (25)

where \( \hat{u} \) indicates a function defined on the reference element. The delineation between \( \mathbb{Q} \) and \( \mathbb{Q}_p \) is required when non-affine\(^1\) mesh transformations are used. In such a case, \( u = \hat{u} \circ T_e^{-1} \notin \mathbb{Q}_p(K_e) \) even if \( \hat{u} \in \mathbb{Q}_p(\hat{K}) \). That is, the solution can be non-polynomial due to the composition with the inverse of the element transformation. For example, the inverse of the transformation in Fig. 1b is

\[ T^{-1} = \begin{bmatrix} \frac{hx_1 + \alpha x_2}{h^2 + 2\alpha x_2} \\ \frac{x_2}{h} \end{bmatrix} \] (26)

which is non-polynomial in the first coordinate.

The degree-\( p \) DG finite element space is:

\[ Y_p = \{ u \in L^2(\mathcal{D}) : u|_{K_e} \in \mathbb{Q}_p(K_e), \ \forall K_e \in \mathcal{T} \} \] (27)

so that each function \( u \in Y_p \) is a piecewise polynomial mapped from the reference element with no continuity enforced between elements. Its vector-valued analog is

\[ W_p = \{ v \in [L^2(\mathcal{D})]^\text{dim} : v \in [\mathbb{Q}_p(K_e)]^\text{dim}, \ \forall K_e \in \mathcal{T} \}, \] (28)

\(^1\)Examples of non-affine transformations include mapping the reference square to a trapezoid or any high-order, curved element.
which simply uses the scalar DG space for each component of the vector. We will also need the discrete $H^1(\Omega)$, or continuous finite element space, defined as:

$$V_p = \{ u \in H^1(\Omega) : u|_{K_e} \in Q_p(K_e), \quad \forall K_e \in T \}. \quad (29)$$

Here, $u \in V_p$ is a piecewise continuous mapped polynomial.

A nodal basis for the element-local polynomial space is used. For a degree-$p$ element, let $\xi_i$ denote the $(p+1)$ Gauss-Lobatto or Gauss-Legendre points in the interval $[0,1]$. The $(p+1)^{\dim}$ points $\xi_i$ on the unit cube $[0,1]^{\dim}$ are given by the dim-fold Cartesian product of the one-dimensional points. Let $\ell_i$ denote the Lagrange interpolating polynomial satisfying $\ell_i(\xi_j) = \delta_{ij}$ where $\delta_{ij}$ is the Kronecker delta. The set of functions $\{\ell_i\}$ form a basis for the space $Q_p(K)$. The DG and $H^1(\Omega)$ finite element spaces are built element-by-element from this local basis.

Note that the Gauss-Lobatto points include the interval end points while the Gauss-Legendre points do not. Thus, using Gauss-Lobatto points yields both points on the interior and the boundary of the element while using Gauss-Legendre leads to points on the interior of the element only. These are referred to as closed and open bases, respectively. In the case of DG, no continuity between elements is enforced so it is acceptable to use either an open or closed basis. Both Gauss-Lobatto and Gauss-Legendre have the required properties to be accurate in the limit $p \to \infty$ so the choice of Gauss-Lobatto versus Gauss-Legendre is typically dictated by other aspects of the overall algorithm such as preconditioners. The basis formed from the Gauss-Legendre points has the beneficial property of diagonal mass matrices on affine meshes, while the basis formed from Gauss-Lobatto points typically leads to sparser global systems since closed bases couple fewer unknowns on interior faces. A closed basis is required for $H^1(\Omega)$ finite element spaces to enable the strong enforcement of continuity between elements.

3.3. Mathematical Notation

It is helpful to define the “broken” gradient, denoted $\nabla_h$, obtained by applying the gradient locally on each element. That is,

$$ (\nabla_h u)|_{K_e} = \nabla(u|_{K_e}), \quad \forall K_e \in T. \quad (30) $$

This distinction is important since for $u \in Y_p$, $\nabla u$ is not well-defined since $u$ may be discontinuous across element interfaces. However, $\nabla_h u$ is well-defined since $u$ is locally differentiable on each element.

We will use the following notation to describe the jump and average of a discontinuous function along an interior mesh face. Let $\Gamma$ be the set of all unique faces in the mesh and $\Gamma_0 = \Gamma \setminus \partial \Omega$ the set of unique interior faces. Additionally, define $\Gamma_b = \Gamma \cap \partial \Omega$ as the set of faces on the boundary so that $\Gamma = \Gamma_0 \cup \Gamma_b$. We
define $\hat{n}_K$ as the outward unit normal to element $K$. On an interior face $F \in \Gamma_0$ between elements $K_1$ and $K_2$, we use the convention that $\hat{n}$ is the the unit vector perpendicular to the shared face $K_1 \cap K_2$ pointing from $K_1$ to $K_2$ (see Fig. 2). On such an interior face, the jump, $[\cdot]$, and average, $\langle \cdot \rangle$, are defined as

$$[u] = u_1 - u_2, \quad \langle u \rangle = \frac{1}{2} (u_1 + u_2), \quad \text{on } F \in \Gamma_0,$$

where $u_i = u|_{\partial K_i}$ with analogous definitions for vectors.

Note that in contrast to the notation of [2], our jump operator does not change the rank of its argument: the jump of a scalar is a scalar and the jump of a vector is a vector. Consequently, our notation is not invariant under element renumbering, since flipping the ordering of the elements negates the value of the jump. However, the bilinear and linear forms presented in this paper always pair the jump with another normal-dependent term. The negation of the jump induced by swapping the element ordering is then balanced by flipping the orientation of the normal vector, and so the discretizations under consideration are in fact invariant with respect to the element ordering.

On the boundary of the mesh, we set the jump and average to

$$[u] = u, \quad \langle u \rangle = u, \quad \text{on } F \in \Gamma_b,$$

and likewise for vector-valued functions on the boundary. A straightforward computation shows that

$$\sum_{K \in T} \int_{\partial K} u \, v \cdot \hat{n}_K \, ds = \int_{\Gamma} [u \, v \cdot \hat{n}] \, ds = \int_{\Gamma} \langle u \rangle \langle v \cdot \hat{n} \rangle \, ds + \int_{\Gamma_0} \langle u \rangle \langle v \cdot \hat{n} \rangle \, ds.$$

We refer to this as the “jumps and averages identity”. The restriction of the integration to interior faces for the second term in the last equality is consistent with the notation of [2] and is used so that only one term contributes on the boundary of the mesh.

Finally, we refer to a function as “single-valued” on an interior face if its values obtained from approaching from each side of the face are identical so that

$$[u] = 0, \quad \langle u \rangle = u.$$

Note in particular that the jump and average operators are single-valued.

4. Transport Discretizations

In this work, we assume the transport equation is discretized with the Discrete Ordinates ($S_N$) angular model and an arbitrary-order Discontinuous Galerkin (DG) spatial discretization compatible with curved meshes (e.g. [39, 1]). In $S_N$, the transport equation is collocated at discrete angles, $\Omega_d$, and integration is numerically approximated using a suitable angular quadrature rule $\{\Omega_d, w_d\}_{d=1}^{N_\Omega}$ on the unit sphere. The VEF data are then

$$E(x) = \frac{\sum_{d=1}^{N_\Omega} w_d \, \Omega_d \otimes \Omega_d \, \psi_d(x)}{\sum_{d=1}^{N_\Omega} w_d \psi_d(x)},$$

$$E_b(x) = \frac{\sum_{d=1}^{N_\Omega} w_d \, |\Omega_d \cdot \hat{n}| \, \psi_d(x)}{\sum_{d=1}^{N_\Omega} w_d \psi_d(x)},$$

where $\psi_d(x) = \psi(x, \Omega_d)$ is the discrete angular flux in direction $\Omega_d$. With degree-$p$ DG in space, the angular flux in each discrete direction $\Omega_d$ is a member of $Y_p$. Through the standard finite element interpolation procedure, the Eddington tensor and boundary factor in Eq. 35 can be evaluated at any location in the mesh. Note that it is important to interpolate the numerator and denominator of the VEF data independently. That is, the boundary factor and each component of the Eddington tensor are represented as degree-$p$ improper rational polynomials on each element.
Defining
\[ P(x) = \sum_d w_d \Omega_d \otimes \Omega_d \psi_d(x) \] (36)
as the discrete second moment of the angular flux and using the quotient rule, the local divergence of the Eddington tensor
\[ \nabla_h \cdot E = \frac{(\nabla_h \cdot P)\phi - P \cdot \nabla_h \phi}{\phi^2} \] (37)
is well-defined assuming \( \phi > 0 \). Here, the divergence of a second-order tensor is the vector formed by taking the divergence of each of the columns of the tensor.

We restrict our attention to problems where \( \psi \geq \delta > 0 \) inside the domain, for some \( \delta \). This assumption is reasonable for our applications but may be violated in shielding or deep penetration problems. Application of a positivity-preserving negative flux fixup then ensures that \( \phi \) is bounded away from zero, so that \( E, E_h \), and \( \nabla_h \cdot E \) are all bounded. Thus, through \( S_N \) angular quadrature and finite element interpolation, and the local divergence of the Eddington tensor can be evaluated at any point in any element of the mesh. This completes the definition of the connection between the discrete transport equation and the VEF drift-diffusion equation. Note that since the angular flux is generally discontinuous across interior mesh interfaces, the Eddington tensor and its divergence also will be. Thus, we will carefully design the discretization of the VEF drift-diffusion equation to accommodate discontinuous data.

The VEF scalar flux connects with the transport equation in the scattering source. To support generality, we assume that the finite element space for the VEF scalar flux and the finite element space for the angular fluxes are different. The scattering source is then constructed using a mixed-space mass matrix that has test functions in the space for the angular flux and trial functions in the space for the VEF scalar flux.

5. Derivation of DG VEF

In this section, we adapt the derivation of a unified framework for DG methods designed for the Poisson equation from [2] to the VEF equations. This enables the use of any of the DG methods described there. Arnold et al. [2] derive a family of DG methods for:
\[ q = \nabla u, \] (38a)
\[ -\nabla \cdot q = f, \] (38b)
with Dirichlet boundary conditions. The present goal is to adapt their derivation to the VEF equations:
\[ \nabla \cdot (E_\phi) + \sigma_t J = Q_1, \] (39a)
\[ \nabla \cdot J + \sigma_a \phi = Q_0, \] (39b)
with the Robin style boundary conditions given in Eq. 6. We will see significant differences in the final equation since the Eddington tensor is inside the divergence. Additionally, the presence of a right-hand side in the first moment equation as well as non-unit coefficients introduce further complications. We will then derive analogues of the interior penalty (IP), second method of Bassi and Rebay (BR2), and Minimal Dissipation Local Discontinuous Galerkin (MDLDG) variants. Finally, we will show how to extract a continuous finite element method from this framework.

5.1. Adaption of the Unified Framework to VEF

We seek the VEF scalar flux in the degree-\( p \) DG finite element space \( Y_p \) and the current in the degree-\( p \) vector-valued DG finite element space \( W_p \). The weak form is then: find \( (\phi, J) \in Y_p \times W_p \) such that for all \( K \in \mathcal{T}: \)
\[ \int_{\partial K} v \cdot \mathbf{E}_\phi \hat{n}_K \, ds - \int_K \nabla v : \mathbf{E}_\phi \, dx + \int_K \sigma_t v \cdot J \, dx = \int_K v \cdot Q_1 \, dx, \quad \forall v \in [Q_p(K)]^{\dim}, \] (40a)
\[
\int_{\partial K} u \vec{J} \cdot \vec{n}_K \, ds - \int_K \nabla u \cdot \vec{J} \, dx = \int_K \sigma_a u \varphi \, dx = \int_K u Q_0 \, dx, \quad \forall u \in \mathbb{Q}_p(K),
\]
(40b)

where the numerical fluxes \( \vec{E} \varphi \) and \( \vec{J} \) are approximations of \( \vec{E} \varphi \) and \( \vec{J} \) on the boundaries of the elements in the mesh. We group the product \( \vec{E} \varphi \) as the numerical flux to mimic the integration by parts of a tensor times a vector. Here, the gradient of a vector is

\[
(\nabla v)_{ij} = \left( \frac{\partial v_j}{\partial x_i} \right) \in \mathbb{R}^{\text{dim} \times \text{dim}}
\]
(41)

and

\[
A : B = \sum_{i=1}^{\text{dim}} \sum_{j=1}^{\text{dim}} A_{ij} B_{ij}, \quad A, B \in \mathbb{R}^{\text{dim} \times \text{dim}}
\]
(42)

is the scalar contraction of two tensors. Note that if \( E = \frac{1}{3} I \) then

\[
\nabla v : E = \frac{1}{3} \nabla \cdot v
\]
(43)

and the symmetric weak form for radiation diffusion is recovered.

Summing the zeroth moment over all elements:

\[
\int_{\Gamma} [v] \left[ \left\{ \vec{J} \cdot \vec{n} \right\} \right] ds + \int_{\Gamma_0} [u] \left[ \left\{ \vec{J} \cdot \vec{n} \right\} \right] ds - \int_K \nabla_h u \cdot \vec{J} \, dx + \int_K \sigma_a u \varphi \, dx = \int_K u Q_0 \, dx,
\]
(44)

where the jumps and averages identity (Eq. 33) was used. We will now use the discrete first moment to determine a functional form for \( \vec{J} \). Integrating by parts locally over element \( K \), we have that

\[
\int_K \nabla \cdot \vec{E} \phi \, dx = \int_{\partial K} \vec{v} \cdot \vec{E} \phi \vec{n}_K \, ds - \int_K \vec{v} \cdot \nabla \cdot (\vec{E} \phi) \, dx.
\]
(45)

The first moment’s weak form on each element becomes:

\[
\int_{\partial K} \vec{v} \cdot \left( \vec{E} \phi \vec{n}_K - \vec{E} \phi \vec{n}_K \right) \, ds + \int_K \vec{v} \cdot \nabla \cdot (\vec{E} \phi) \, dx + \int_K \sigma_t \vec{v} \cdot \vec{J} \, dx = \int_K \vec{v} \cdot Q_1 \, dx, \quad \forall \vec{v} \in [\mathbb{Q}_p(K)]^\text{dim}.
\]
(46)

Summing over all elements and using the jumps and averages identity, the weak form for the first moment is:

\[
\int_{\Gamma} [v] \cdot \left[ \vec{E} \phi \vec{n} - \vec{E} \phi \vec{n} \right] \, ds + \int_{\Gamma_0} [v] \cdot \left[ \left\{ \vec{E} \phi \vec{n} - \vec{E} \phi \vec{n} \right\} \right] ds
\]

\[
+ \int_K \vec{v} \cdot \nabla_h \cdot (\vec{E} \phi) \, dx + \int_K \sigma_t \vec{v} \cdot \vec{J} \, dx = \int_K \vec{v} \cdot Q_1 \, dx, \quad \forall \vec{v} \in W_p,
\]
(47)

where \( \nabla_h \cdot (\vec{E} \phi) \) is evaluated as \( \nabla_h \cdot (\vec{E} \phi) = \vec{E} \nabla_h \varphi + (\nabla_h \cdot \vec{E}) \phi \), and the term \( \nabla_h \cdot \vec{E} \) is computed using Eq. 37.

We now wish to write all terms as volumetric integrals so that a functional form for the current can be found. To that end, define lifting operators \( \mathbf{r}(\tau) \in W_p \) and \( \ell(\chi) \in W_p \) such that

\[
\int_K \sigma_t \vec{v} \cdot \mathbf{r}(\tau) \, dx = - \int_{\Gamma} [\vec{v}] \cdot \tau \, ds, \quad \forall \vec{v} \in W_p,
\]
(48a)

\[
\int_K \sigma_t \vec{v} \cdot \ell(\chi) \, dx = - \int_{\Gamma_0} [\vec{v}] \cdot \chi \, ds, \quad \forall \vec{v} \in W_p,
\]
(48b)

where \( \tau \) and \( \chi \) are vector functions that are singled-valued on \( \Gamma_0 \). Note that the lifting operators are finite element grid functions just as the current is and that the left hand sides are simply the \( W_p \) total interaction
mass matrix. Since $W_p$ is piecewise discontinuous, the $W_p$ mass matrix is block-diagonal by element and thus the systems of equations corresponding to Eqs. 48a and 48b are amenable to efficient direct factorization (see Appendix Appendix A).

Setting $\tau = \left[ \mathbf{E} \mathbf{\phi} \mathbf{n} - \mathbf{E} \mathbf{\phi} \mathbf{n} \right]$ and $\chi = \left\{ \left[ \mathbf{E} \mathbf{\phi} \mathbf{n} - \mathbf{E} \mathbf{\phi} \mathbf{n} \right] \right\}$ and using the definitions of the lifting operators, Eq. 47 can be written entirely in terms of volumetric integrals as:

$$
\int \sigma_t \mathbf{v} \cdot \mathbf{J} \, d\mathbf{x} = \int \sigma_t \mathbf{v} \cdot \left[ \frac{1}{\sigma_t} (\mathbf{Q}_1 - \nabla_h \cdot (\mathbf{E} \mathbf{\phi})) + \mathbf{r} \left( \left[ \mathbf{E} \mathbf{\phi} \mathbf{n} - \mathbf{E} \mathbf{\phi} \mathbf{n} \right] \right) + \mathbf{t} \left( \left\{ \left[ \mathbf{E} \mathbf{\phi} \mathbf{n} - \mathbf{E} \mathbf{\phi} \mathbf{n} \right] \right\} \right) \right] \, d\mathbf{x} \quad (49)
$$

for all $\mathbf{v} \in W_p$. Subtracting the right hand side and setting the integrand to zero implies that

$$
\mathbf{J} = \frac{1}{\sigma_t} (\mathbf{Q}_1 - \nabla_h \cdot (\mathbf{E} \mathbf{\phi})) + \mathbf{r} \left( \left[ \mathbf{E} \mathbf{\phi} \mathbf{n} - \mathbf{E} \mathbf{\phi} \mathbf{n} \right] \right) + \mathbf{t} \left( \left\{ \left[ \mathbf{E} \mathbf{\phi} \mathbf{n} - \mathbf{E} \mathbf{\phi} \mathbf{n} \right] \right\} \right). \quad (50)
$$

Observe that the above represents the element-local strong form of the current, $\frac{1}{\sigma_t} (\mathbf{Q}_1 - \nabla_h \cdot (\mathbf{E} \mathbf{\phi}))$ found by analytically eliminating the current, with additional terms that capture the effect of the numerical fluxes. In other words, we have derived the discrete elimination of the current.

Using this discrete form for the current and the definitions of the lifting operators to convert from volumetric integrals back to surface integrals, the zeroth moment becomes:

$$
\int_{\Gamma} \left\{ \left[ \mathbf{J} \cdot \hat{n} \right] \right\} \, d\Gamma + \int_{\Gamma_0} \left\{ \left[ \mathbf{J} \cdot \hat{n} \right] \right\} \, d\Gamma + \int_{\Gamma} \left\{ \left[ \nabla_h \mathbf{u} \right] \right\} \cdot \left[ \mathbf{E} \mathbf{\phi} \mathbf{n} - \mathbf{E} \mathbf{\phi} \mathbf{n} \right] \, d\Gamma + \int_{\Gamma_0} \left\{ \left[ \nabla_h \mathbf{u} \right] \right\} \cdot \left[ \mathbf{E} \mathbf{\phi} \mathbf{n} - \mathbf{E} \mathbf{\phi} \mathbf{n} \right] \, d\Gamma = \int_{\Gamma} \mathbf{Q}_0 \, d\Gamma + \int_{\Gamma} \nabla_h \mathbf{u} \cdot \frac{\mathbf{Q}_1}{\sigma_t} \, d\Gamma, \quad \forall \mathbf{u} \in Y_p. \quad (51)
$$

On boundary faces, we apply the Miften-Larsen boundary conditions by setting

$$
\mathbf{J} \cdot \hat{n} = 2g + E_{b\mathbf{\phi}}, \quad \mathbf{E} \mathbf{\phi} \mathbf{n} = \mathbf{E} \mathbf{\phi} \mathbf{n}, \quad \text{on} \ F \in \Gamma_b. \quad (52)
$$

All the methods we consider use so-called conservative numerical fluxes such that

$$
\left[ \mathbf{J} \cdot \hat{n} \right] = 0, \quad \left\{ \left[ \mathbf{J} \cdot \hat{n} \right] \right\} = \mathbf{J} \cdot \hat{n}, \quad \text{on} \ F \in \Gamma_0, \quad (53a)
$$

$$
\left[ \mathbf{E} \mathbf{\phi} \mathbf{n} \right] = 0, \quad \left\{ \left[ \mathbf{E} \mathbf{\phi} \mathbf{n} \right] \right\} = \mathbf{E} \mathbf{\phi} \mathbf{n}, \quad \text{on} \ F \in \Gamma_0. \quad (53b)
$$

Using the boundary conditions and the assumption of conservative numerical fluxes, Eq. 51 becomes:

$$
\int_{\Gamma_b} E_b \mathbf{u} \mathbf{\phi} \, d\Gamma + \int_{\Gamma_0} \left\{ \left[ \mathbf{J} \cdot \hat{n} \right] \right\} \, d\Gamma - \int_{\Gamma_0} \left\{ \left[ \nabla_h \mathbf{u} \right] \right\} \cdot \left[ \mathbf{E} \mathbf{\phi} \mathbf{n} - \mathbf{E} \mathbf{\phi} \mathbf{n} \right] \, d\Gamma
$$

$$
+ \int_{\Gamma_0} \left\{ \left[ \nabla_h \mathbf{u} \right] \right\} \cdot \left[ \left\{ \mathbf{E} \mathbf{\phi} \mathbf{n} - \mathbf{E} \mathbf{\phi} \mathbf{n} \right] \right\} \, d\Gamma + \int \nabla_h \mathbf{u} \cdot \frac{1}{\sigma_t} \nabla_h \cdot (\mathbf{E} \mathbf{\phi}) \, d\Gamma + \int \sigma_a \mathbf{u} \mathbf{\phi} \, d\Gamma = \int_{\Gamma} \mathbf{Q}_0 \, d\Gamma + \int \nabla_h \mathbf{u} \cdot \frac{\mathbf{Q}_1}{\sigma_t} \, d\Gamma - 2 \int_{\Gamma_b} \mathbf{u} g \, d\Gamma, \quad \forall \mathbf{u} \in Y_p. \quad (54)
$$

Equation 54 defines a family of DG methods. That is, through the specification of the numerical flux for the current on interior faces, analogues of all the methods listed in [2] can be derived.
5.2. Specification of Numerical Fluxes

All the methods we consider use numerical fluxes of the form

\[ \mathbf{J} \cdot \mathbf{n} = \left\{ \left\{ \frac{1}{\sigma_t} (\mathbf{Q}_1 - \nabla_h \cdot (\mathbf{E} \varphi)) \cdot \hat{n} \right\} + \alpha(\varphi) \right\} + \theta(\varphi), \quad \text{on } \Gamma_0, \]

\[ \mathbf{E} \varphi \mathbf{n} = \| \mathbf{E} \varphi \mathbf{n} \| + \theta(\varphi), \quad \text{on } \Gamma_0, \]

where \( \alpha(\varphi) \) and \( \theta(\varphi) \) are single-valued functions whose purpose is to ensure a stable discretization. The IP, BR2, and LDG methods differ only in the choice of \( \alpha(\varphi) \) and \( \theta(\varphi) \). With these numerical fluxes, Eq. 54 becomes:

\[ \int_{\Gamma_b} E_h u \varphi \, ds + \int_{\Gamma_0} [u] \alpha(\varphi) \, ds - \int_{\Gamma_0} \left\{ \left\{ \frac{1}{\sigma_t} \nabla_h \cdot (\mathbf{E} \varphi) \cdot \hat{n} \right\} \left\{ \begin{array}{c} \mathbf{Q}_1 \\ \sigma_t \end{array} \right\} \right\} \, ds + \int_{\Gamma_0} \left\{ \left\{ \frac{\nabla_h u}{\sigma_t} \right\} \cdot \mathbf{E} \varphi \right\} \, ds \\
+ \int_{\Gamma_b} \left\{ \frac{\nabla_h u}{\sigma_t} \right\} \cdot \theta(\varphi) \, ds + \int \nabla_h u \cdot \frac{1}{\sigma_t} \nabla_h \cdot (\mathbf{E} \varphi) \, dx + \int \sigma_a u \varphi \, dx \\
\quad = \int u Q_0 \, dx + \int \nabla_h u \cdot \frac{Q_1}{\sigma_t} \, dx - \int [u] \left\{ \begin{array}{c} \mathbf{Q}_1 \cdot \hat{n} \\ \sigma_t \end{array} \right\} \, ds - 2 \int_{\Gamma_b} u g \, ds, \quad \forall u \in Y_p. \] \tag{56}

Recall that this form has already applied boundary conditions according to Eq. 52. In other words, the above corresponds to a DG scheme with the following numerical fluxes:

\[ \mathbf{J} \cdot \mathbf{n} = \left\{ \left\{ \frac{1}{\sigma_t} (\mathbf{Q}_1 - \nabla_h \cdot (\mathbf{E} \varphi)) \cdot \hat{n} \right\} + \alpha(\varphi) \right\}, \quad \text{on } \Gamma_0 \]

\[ \mathbf{E} \varphi \mathbf{n} = \left\{ \begin{array}{c} \| \mathbf{E} \varphi \mathbf{n} \| \right\} + \theta(\varphi), \quad \text{on } \Gamma_0 \]

5.2.1. Interior Penalty

An interior penalty (IP)-like method uses

\[ \alpha(\varphi) = \kappa \| \varphi \|, \quad \theta(\varphi) = 0, \]

where \( \kappa \) is the penalty parameter. IP methods require that \( \kappa \propto \sigma_t^{-1} p^2 / h \) in order to guarantee stability. The full IP weak form is then: find \( \varphi \in Y_p \) such that

\[ \int_{\Gamma_b} E_h u \varphi \, ds + \int_{\Gamma_0} \kappa [u] [\| \varphi \|] \, ds - \int_{\Gamma_0} [u] \left\{ \left\{ \frac{1}{\sigma_t} \nabla_h \cdot (\mathbf{E} \varphi) \cdot \hat{n} \right\} \right\} \, ds - \int_{\Gamma_0} \left\{ \left\{ \frac{\nabla_h u}{\sigma_t} \right\} \cdot \mathbf{E} \varphi \right\} \, ds \\
+ \int \nabla_h u \cdot \frac{1}{\sigma_t} \nabla_h \cdot (\mathbf{E} \varphi) \, dx + \int \sigma_a u \varphi \, dx \\
\quad = \int u Q_0 \, dx + \int \nabla_h u \cdot \frac{Q_1}{\sigma_t} \, dx - \int [u] \left\{ \begin{array}{c} \mathbf{Q}_1 \cdot \hat{n} \\ \sigma_t \end{array} \right\} \, ds - 2 \int_{\Gamma_b} u g \, ds, \quad \forall u \in Y_p. \] \tag{59}

5.2.2. BR2

The second method of Bassi and Rebay (BR2) uses an alternative penalty term. Let \( \mathbf{p}_f(\omega) \in W_p \) be a face-local lifting operator defined by

\[ \int_{\Gamma_0} \mathbf{v} \cdot \mathbf{p}_f(\omega) \, dx = - \int_{f} [\mathbf{v} \cdot \hat{n}] \omega \, ds, \quad \forall \mathbf{v} \in W_p, \quad \text{on } f \in \Gamma_0. \] \tag{60}

Here, \( \omega \) is a scalar function that is single-valued on the interior face \( f \). Note that the integration on the left hand side is over the entire domain while the right hand side is localized to a single interior face. This
means the right hand side, and thus $\rho_f(\omega)$, will be non-zero only for DOFs in elements that share the face $f$.

A BR2-like discretization sets

$$\alpha(\varphi) = -\eta \left\{ \rho_f([\varphi]) \cdot \hat{n} \right\}, \quad \text{on } f \in \Gamma_0, \quad \theta(\varphi) = 0,$$

so that the relevant term is

$$\int_{\Gamma_0} [u] \alpha(\varphi) = -\sum_{f \in \Gamma_0} \int_f [u] \left\{ \rho_f([u]) \cdot \hat{n} \right\} ds$$

$$= \sum_{f \in \Gamma_0} \int \eta \rho_f([u]) \cdot \rho_f([\varphi]) dx.$$ 

This BR2 numerical flux avoids the need to tune the penalty parameter while still allowing element-by-element assembly (see Appendix A).

The BR2 DG VEF discretization is then: find $\varphi \in Y_p$ such that

$$\int_{\Gamma_b} E_h u \varphi ds - \int_{\Gamma_0} [u] \left\{ \left[ \frac{1}{\sigma_t} \nabla_h \cdot (E \varphi) \cdot \hat{n} \right] \right\} ds - \int_{\Gamma_0} \left\{ \left[ \frac{\nabla_h u}{\sigma_t} \right] \cdot [E \varphi \hat{n}] \right\} ds$$

$$+ \sum_{f \in \Gamma_0} \int \eta \rho_f([u]) \cdot \rho_f([\varphi]) dx + \int \nabla_h u \cdot \frac{1}{\sigma_t} \nabla_h \cdot (E \varphi) dx + \int \sigma_a u \varphi dx$$

$$= \int u Q_0 dx + \int \nabla_h u \cdot Q^1_{\sigma_t} dx - \int_{\Gamma_0} [u] \left\{ \left[ \frac{Q^1 \cdot \hat{n}}{\sigma_t} \right] \right\} ds - 2 \int_{\Gamma_b} u g ds, \quad \forall u \in Y_p.$$

### 5.2.3. Local Discontinuous Galerkin

Finally, we consider the Local Discontinuous Galerkin (LDG) method. In general, LDG uses the following numerical fluxes:

$$\mathbf{J} \cdot \hat{n} = \left\{ \mathbf{J} \cdot \hat{n} \right\} + \beta \left\{ \mathbf{J} \cdot \hat{n} \right\} + \kappa [\varphi],$$

$$\mathbf{E} \varphi \hat{n} = \left\{ \mathbf{E} \varphi \hat{n} \right\} - \beta \left\{ \mathbf{E} \varphi \hat{n} \right\},$$

where $\mathbf{J}$ is defined as the discrete elimination of the current derived in Eq. 50. The scalar parameter $\beta$ can be defined as

$$\beta = \begin{cases} 1/2, & w \cdot \hat{n} > 0 \\ -1/2, & w \cdot \hat{n} < 0 \end{cases},$$

where $w$ is any constant, non-zero vector. This choice imposes an arbitrary upwinding on the current that is balanced by an opposing choice for the scalar flux. With this choice of $\beta$, the LDG method is stable for any $\kappa \geq 0$; if $\kappa \equiv 0$, the method is referred to as the Minimal Dissipation LDG (MDLDG) method [43].

Using the numerical flux for the scalar flux, the discrete current simplifies to

$$\mathbf{J} = \frac{1}{\sigma_t} \left( Q_1 - \nabla_h \cdot (E \varphi) \right) - r_0([E \varphi \hat{n}]) - \ell(\beta \left\{ \mathbf{E} \varphi \hat{n} \right\}),$$

where $r_0(\tau) \in W_p$ is another lifting operator defined by

$$\int_{\Gamma_0} \sigma_t v \cdot r_0(\tau) dx = -\int_{\Gamma_0} [v] \cdot \tau ds, \quad \forall v \in W_p,$$

that differs from $r(\tau)$ only in the region of integration on the right hand side. The LDG method is then equivalent to setting

$$\alpha(\varphi) = -\left\{ r_0([E \varphi \hat{n}]) \cdot \hat{n} \right\} + \ell(\beta \left\{ \mathbf{E} \varphi \hat{n} \right\}) \cdot \hat{n}$$

$$+ \beta \left[ \frac{1}{\sigma_t} \left( Q_1 - \nabla_h \cdot (E \varphi) \right) \cdot \hat{n} - r_0([E \varphi \hat{n}]) \cdot \hat{n} - \ell(\beta \left\{ \mathbf{E} \varphi \hat{n} \right\}) \cdot \hat{n} \right] + \kappa [\varphi],$$

$$\text{Eq. 68a}.$$
\( \theta(\varphi) = -\beta [E \varphi \hat{n}] \).

We then have that

\[
\int_{\Gamma_0} [u] \alpha(\varphi) \, ds = \int_{\Gamma_0} \beta [u] \left[ \frac{Q_1 \cdot \hat{n}}{\sigma_t} \right] \, ds - \int_{\Gamma_0} \beta [u] \left[ \frac{1}{\alpha_t} \nabla_h \cdot (E \varphi) \cdot \hat{n} \right] \, ds \\
+ \int (\rho_0([u]) + \lambda([u])) \cdot (r_0([E \varphi \hat{n}]) + \ell(\beta [E \varphi \hat{n}])) \, dx + \int_{\Gamma_0} \kappa [u] [\varphi] \, ds \tag{69}
\]

where \( \rho_0(\omega), \lambda(\nu) \in W_p \) such that

\[
\int \mathbf{v} \cdot \rho_0(\omega) \, dx = -\int_{\Gamma_0} \left\{ \mathbf{v} \cdot \hat{n} \right\} \omega \, ds, \quad \forall \mathbf{v} \in W_p, \tag{70}
\]

\[
\int \mathbf{v} \cdot \lambda(\nu) \, dx = -\int_{\Gamma_0} [\mathbf{v} \cdot \hat{n}] \nu \, ds, \quad \forall \mathbf{v} \in W_p, \tag{71}
\]

are analogues of \( r_0(\tau) \) and \( \ell(\chi) \), respectively, that do not include the total interaction cross section in the left hand side mass matrices and have scalar arguments. The LDG VEF discretization is then: find \( \varphi \in Y_p \) such that

\[
\int_{\Gamma_h} E_h u \varphi \, ds + \int_{\Gamma_0} \kappa [u] [\varphi] \, ds - \int_{\Gamma_0} [u] \left\{ \left[ \frac{1}{\sigma_t} \nabla_h \cdot (E \varphi) \cdot \hat{n} \right] \right\} \, ds - \int_{\Gamma_0} \left\{ \left[ \frac{\nabla_h u}{\sigma_t} \right] \cdot [E \varphi \hat{n}] \right\} \, ds \\
+ \int (\rho_0([u]) + \lambda([u])) \cdot (r_0([E \varphi \hat{n}]) + \ell(\beta [E \varphi \hat{n}])) \, dx \\
+ \int \nabla_h u : \frac{1}{\alpha_t} \nabla_h \cdot (E \varphi) \, dx + \int \sigma_a u \varphi \, dx \\
= \int u Q_0 \, dx + \int \nabla_h u : \frac{Q_1}{\sigma_t} \, dx - \int_{\Gamma_0} [u] \left\{ \left[ \frac{Q_1 \cdot \hat{n}}{\sigma_t} \right] + \beta \left[ \frac{Q_1 \cdot \hat{n}}{\sigma_t} \right] \right\} \, ds - 2 \int_{\Gamma_h} u g \, ds, \quad \forall u \in Y_p. \tag{72}
\]

The advantage of LDG is that the penalty parameter does not need to scale with the mesh size. However, the LDG stabilization term has a non-compact stencil that connects neighbors of neighbors, leading to less sparsity than the IP or BR2 methods.

5.3. Continuous Finite Element Discretization of VEF

We now show how a continuous finite element (CG) discretization of the VEF drift-diffusion equation can be extracted from the DG framework presented above. An approximate inversion of this operator is one stage of the subspace correction preconditioner described in Section 6 that is used to efficiently solve the IP and BR2 VEF discretizations. This CG operator is also a VEF method itself and represents an extension to multiple dimensions, arbitrary-order, and curved meshes of the algorithm in [13]. A CG VEF method has fewer unknowns than an analogous DG method and requires simpler methods to solve the resulting linear system. We will show that this CG discretization has similar accuracy to DG and does not degrade convergence of the fixed-point iteration even in the asymptotic thick diffusion limit. However, it is unclear if using a continuous finite element space would negatively impact robustness and stability in the larger radiation-hydrodynamics multiphysics setting.

Let \( u, \varphi \in V_p \), the degree-\( p \) continuous finite element space, then

\[
[u] = 0, \quad [\varphi] = 0, \quad \text{on} \; \mathcal{F} \in \Gamma_0. \tag{73}
\]

However, since the Eddington tensor is still discontinuous, we have that

\[
[E \varphi \hat{n}] = [E \hat{n}] \varphi. \tag{74}
\]
Note that, for \( u \in V_p, \nabla u \in W_p \). In other words, while \( u \in V_p \) is continuous \( \nabla u \) is not. Thus, by starting from the DG VEF discretization and assembling onto \( V_p \), we arrive at a CG VEF discretization of the form: find \( \varphi \in V_p \) such that

\[
\int_{\Gamma_b} E_b u \varphi \, ds - \int_{\Gamma_0} \left\{ \frac{\nabla u}{\sigma_t} \right\} \cdot [E \hat{\varphi}] \, ds + \int \nabla u \cdot \frac{1}{\sigma_t} \nabla_h \cdot (E \varphi) \, dx + \int \sigma_u u \varphi \, dx = \int u Q_0 \, dx + \int \nabla u \cdot \frac{Q_1}{\sigma_t} \, dx - 2 \int_{\Gamma_b} u g \, ds, \quad \forall u \in V_p. \tag{75}
\]

Observe that in the thick diffusion limit, where \( E = \frac{1}{3} I \) and \( E_b = 1/2 \), a CG discretization of radiation diffusion with Marshak boundary conditions arises since \([E \hat{\varphi}] = 0\) and \( \frac{1}{\sigma_t} \nabla_h \cdot (E \varphi) = \frac{1}{\sigma_t} \nabla \varphi \).

### 6. Subspace Correction Preconditioners

In this section, we develop effective and efficient preconditioners for the linear systems resulting from the DG discretizations of the VEF equations developed in Section 5. These preconditioners are built using the additive Schwarz or parallel subspace correction framework \([44, 45]\). We will first discuss the preconditioning of symmetric positive-definite DG discretizations of diffusion equations, and then extend the results to the non-symmetric VEF discretizations. We begin by reviewing some preliminary results from the domain decomposition literature \([46]\).

**Remark 1.** In what follows, we will be interested in proving estimates that are independent of discretization parameters such as mesh size \( h \), polynomial degree \( p \), and penalty parameter \( \kappa \). For simplicity of notation, we will write \( a \lesssim b \) to mean \( a \leq C b \), for some constant \( C \), independent of \( h, p, \) and \( \kappa \). Similarly, \( a \gtrsim b \) is used to mean \( b \gtrsim a \), and \( a \approx b \) means that both \( a \lesssim b \) and \( b \lesssim a \).

We consider a decomposition of the DG finite element space \( Y_p \) as the sum of subspaces

\[
Y_p = Y_1 + Y_2 + \cdots + Y_J. \tag{76}
\]

Let \( \mathcal{A}(u, v) \) denote a symmetric positive definite bilinear form, and let \( A \) denote the corresponding operator, i.e.

\[
\mathcal{A}(u, v) = (Au, v), \tag{77}
\]

where \((\cdot, \cdot)\) is the standard \( L^2(D) \) inner product. For example, we can take \( \mathcal{A}(u, v) \) to be one of the standard DG discretizations of the diffusion equation as presented in \([2]\). Let \( A_j \) denote the restriction of \( A \) to the subspace \( Y_j \), and let \( P_j \) be the elliptic projections onto \( Y_j \). That is,

\[
\mathcal{A}(P_j u, v_j) = \mathcal{A}(u, v_j) \quad \text{for all } v_j \in Y_j. \tag{78}
\]

Similarly, define the \( L^2 \) projections onto \( Y_j \) by

\[
(Q_j u, v_j) = (u, v_j) \quad \text{for all } v_j \in Y_j. \tag{79}
\]

It can be seen that

\[
A_j P_j = Q_j A,
\]

and so \( P_j = A_j^{-1} Q_j A \). Inverting the local problems \( A_j \) exactly may be computationally infeasible, and so we can replace \( A_j^{-1} \) with an approximate inverse \( \tilde{A}_j^{-1} \) such that \( \tilde{A}_j^{-1} A_j \) is uniformly well-conditioned. Then, we make use of the operators \( T_j = \tilde{A}_j^{-1} Q_j A \). The **preconditioned operator** \( T \) is defined as the sum of the subspace operators, \( T = \sum_{j=1}^J T_j \). The corresponding preconditioner is given by \( \sum_{j=1}^J T_j^{-1} Q_j \). Under certain conditions on the subspaces \( Y_j \), the preconditioned system \( T = \sum_{j=1}^J T_j^{-1} Q_j A \) is well-conditioned.
6.1. Decomposition into conforming and interface subspaces

At this point, we consider the particular decomposition of \( Y_p \) into the sum of two subspaces (cf. [47]),

\[
Y_p = Y_B + V_p,
\]

where we recall that \( V_p \subset Y_p \) consists of functions that are globally continuous, i.e. \( V_p = Y_p \cap H^1(D) \). \( Y_B \) consists of functions that vanish at all element-interior Gauss-Lobatto points (but which may take arbitrary values at element-boundary Gauss-Lobatto points). This decomposition is closely related to the idea of preconditioning discontinuous Galerkin discretizations with a related continuous Galerkin discretization [48, 49, 50]. It is easy to see that an arbitrary function \( w \in Y_p \) has a (non-unique) decomposition as \( w = w_B + v, w_B \in Y_B, v \in V_p \). Adopting the above notation, let \( P_B \) and \( P_V \) denote the elliptic projections onto \( Y_B \) and \( V_p \) respectively.

We recall some results concerning this space decomposition from [47, 51]. Let \( \mathcal{A} \) denote here the standard interior penalty DG discretization of the diffusion equation.

**Proposition 1** (Cf. [47], Theorem 1). The space decomposition \( Y_p = Y_B + V_p \) is stable, i.e. for any \( w \in Y_p \), there exist a decomposition \( w = w_B + v, w_B \in Y_B, v \in V_p \) such that

\[
\mathcal{A}(w_B, w_B) + \mathcal{A}(v, v) \lesssim \mathcal{A}(w, w).
\]

As a consequence of Lions’ lemma [52], we have

\[
\mathcal{A}(w, w) \lesssim \mathcal{A}(P_B w, w) + \mathcal{A}(P_V w, w).
\]

An upper bound on \( \mathcal{A}(P_V h, v_h) \), where \( P = P_B + P_V \) is obtained by noting that the operators \( P_B \) and \( P_V \) are projections.

**Corollary 1.** The preconditioned operator \( P = P_B + P_V \) is uniformly well-conditioned.

Notice that the operator \( \mathcal{A} \) restricted to the continuous space \( V_p \) corresponds to a standard \( H^1 \) discretization of the diffusion equation. As a result, the local solver \( \tilde{A}_V^{-1} \) can be replaced with any uniform preconditioner \( \tilde{A}_V^{-1} \) for diffusion problems to obtain the approximate operator \( T_V \). For instance, we can take \( \tilde{A}_V^{-1} \) to be one V-cycle of hypre’s BoomerAMG [53].

It remains to find an approximate solver for the operator \( A_B \). Suppose the mesh \( T \) is conforming, and the space \( Y_p \) has constant polynomial degree. Let \( \tilde{A}_B^{-1} \) be the simple point Jacobi preconditioner applied to \( A_B \). Then, we have the following result from [47].

**Theorem 1.** Let \( T_B = \tilde{A}_B^{-1} Q_B A \) and let \( T_V = \tilde{A}_V^{-1} Q_V A \), where \( \tilde{A}_B^{-1} \) is the point Jacobi preconditioner for \( A_B \), and \( \tilde{A}_V^{-1} \) represents one V-cycle of BoomerAMG (or any other uniform preconditioner for the \( H^1 \)-conforming discretization of diffusion). Then, the preconditioned operator \( T = T_B + T_V \) is uniformly well-conditioned.

**Remark 2.** When the mesh \( T \) is nonconforming (e.g. as the result of adaptive mesh refinement), or when the DG finite element space \( Y_p \) has variable polynomial degrees, then a more sophisticated subspace decomposition is required [3]. In this case, the boundary subspace \( Y_B \) is decomposed into a collection of smaller subspaces defined on each non-conforming edge. Each of these small subspaces is solved independently, giving rise to a block Jacobi-type method. In the case that the mesh is conforming and the polynomial degree is constant, this construction reduces to the point Jacobi approximate solver described above.

6.2. Symmetric VEF discretizations

We extend the analysis of the above preconditioners to the family of DG discretizations of the VEF equations given by Eq. 56. We first treat the simple case where \( E = \frac{1}{4} I \). In this case, the system defined by Eq. 56 is symmetric and positive-definite. These results can also be extended to the more general case of constant Eddington tensor; in this case, the results will depend on the spectrum of \( E \). Let \( \mathcal{B}(u, v) \) denote the
bilinear form defined by Eq. 56. We consider the subspace correction preconditioner defined above, and seek to extend Theorem 1 to this modified system. In order to do this, we must first show that the decomposition $Y_p = Y_B + V_p$ is stable with respect to the modified bilinear form $B$. To do this, it suffices to show that the norm induced by $B$ is equivalent to the norm induced by $A$. We first note that the standard interior penalty DG discretization of the definite Helmholtz operator $\sigma_a u - \nabla \cdot (\sigma_t^{-1} \nabla u)$ satisfies the following bounds (cf. [47, 54])

$$A(u, v) \lesssim |||u||| |||v|||,$$

$$A(u, u) \gtrsim ||u||^2,$$

where the mesh-dependent DG norm $||| \cdot |||$ is defined by

$$||u||^2 = ||\sigma_a u||_0^2 + ||\sigma_t^{-1/2} \nabla_h u||_2^2 + \frac{p^2}{h} ||\sigma_t^{-1/2} [u]||_0^2.$$

We first consider the interior penalty version of the VEF discretization, given by Eq. 59. It is straightforward to see that $B$ satisfies the same inequalities,

$$B(u, v) \lesssim ||u|| ||v||,$$

$$B(u, u) \gtrsim ||u||^2.$$

The extension to BR2 and LDG discretizations follows from estimates of the lifting operators $\rho_g$, $r$, and $\ell$, which are considered in [55, 54, 51].

As a consequence of this equivalence in norms, we expect the parallel subspace preconditioner described above to result in a uniformly well-conditioned operator, independent of mesh size $h$, polynomial degree $p$ (as well as the size of the interior penalty stabilization penalty parameter $\kappa$).

### 6.3. Non-symmetric VEF discretizations

The case of more general Eddington tensor $E$ is more difficult to treat because the resulting bilinear form $B$ is no longer symmetric. We analyze the convergence of the preconditioned GMRES iterative method, with the preconditioner defined by the parallel subspace correction procedure described above. The rate of convergence of the GMRES method applied to a non-symmetric, but positive definite operator is controlled by the ratio of the minimal eigenvalue of the symmetric part of the operator to the norm of the operator [56]. We wish to show that this ratio remains independent of the discretization parameters, and therefore that the number of GMRES iterations required to converge remains uniformly bounded. To do this, recalling the literature on additive Schwarz methods for nonsymmetric problems (cf. [57, 58]), we must show that the nonsymmetric part of the operator is small in some sense.

In order to simplify the analysis, we consider a slightly modified VEF discretization that results from iteratively lagging certain nonsymmetric terms. In particular, we write $\nabla_h \cdot (E \varphi) = E \nabla_h \varphi + (\nabla_h \cdot E) \varphi$, and iteratively lag the second term on the right-hand side, replacing $(\nabla_h \cdot E) \varphi$ with $(\nabla_h \cdot E) \tilde{\varphi}$, where $\tilde{\varphi}$ is given from the previous iteration. The iteratively lagged version of Eq. 59 then gives rise to

$$B(u, \varphi) = \int_{\Gamma_n} E_h u \varphi \, ds + \int_{\Gamma_n} \kappa [u] [\varphi] \, ds - \int_{\Gamma_n} [u] \left\{ \frac{1}{\sigma_t} E \nabla_h \varphi \cdot \hat{n} \right\} \, ds$$

$$- \int_{\Gamma_n} \left\{ \frac{\nabla_h u}{\sigma_t} \right\} \cdot [E \varphi  \hat{n}] \, ds + \int \nabla_h u \cdot \frac{1}{\sigma_t} E \nabla_h \varphi \, dx + \int \sigma_a u \varphi \, dx \quad (83)$$

We decompose $B$ into its symmetric and skew-symmetric parts, $B(u, \varphi) = S(u, \varphi) + N(u, \varphi)$, where

$$S(u, \varphi) = \frac{1}{2} (B(u, \varphi) + B(\varphi, u)),$$

$$N(u, \varphi) = \frac{1}{2} (B(u, \varphi) - B(\varphi, u)).$$
Cf. Theorem 1.3 from [57], preconditioned GMRES will converge uniformly with respect to the discretization parameters if there exists a constant $0 \leq \delta < 1$ such that

$$|N(u, Pu)| \leq \delta B(u, Pu),$$

(84)

where $P = P_B + P_V$ is the preconditioned operator. We see that the skew-symmetric part of Eq. 83 is given by

$$N(u, \varphi) = \frac{1}{2} \left( -\int_{\Gamma_0} [u] \left\{ \left\{ \frac{1}{\sigma_t} E\nabla_h \varphi \cdot \hat{n} \right\} \right\} \mathrm{d}s + \int_{\Gamma_0} [\varphi] \left\{ \left\{ \frac{1}{\sigma_t} E\nabla_h u \cdot \hat{n} \right\} \right\} \mathrm{d}s$$

$$- \int_{\Gamma_0} \left\{ \left\{ \frac{\nabla_h u}{\sigma_t} \right\} \cdot [E\varphi]\hat{n} \right\} \mathrm{d}s + \int_{\Gamma_0} \left\{ \left\{ \frac{\nabla_h \varphi}{\sigma_t} \right\} \cdot [Eu\hat{n}] \right\} \mathrm{d}s \right).$$

(85)

Applying the identity $\|ac\| - \|a\|\|b\| = \frac{1}{2} \|c\| (a_1 b_2 + a_2 b_1)$ to the above expression yields the following boundedness property

$$|N(u, \varphi)| \lesssim \|E\| \|u\| \|\varphi\|,$$

where $\|E\|$ represents an upper bound on the jump of $E$ over all element interfaces in the mesh. Using that $Y_p = Y_B + V_p$ is a stable decomposition, we have

$$B(u, Pu) = B(u, P_B u) + B(u, P_V u) = B(P_B u, P_B u) + B(P_B u, P_V u)$$

$$= S(P_B u, P_B u) + S(P_B u, P_V u) \gtrsim \|u\|^2.$$

Furthermore, since $P_B$ and $P_V$ are projections,

$$\|Pu\| = \|P_B u + P_V u\| \leq 2\|u\|.$$

Combining the above estimates, we obtain

$$|N(u, Pu)| \lesssim \|E\| \|u\| \|Pu\| \lesssim \|E\| B(u, Pu).$$

Therefore, in order to obtain the bound (84) with $0 \leq \delta < 1$, according to the size of the jumps $\|E\|$, we may choose $\kappa$ sufficiently large in the symmetric penalty term

$$\int_{\Gamma_0} \kappa \|u\| \|\varphi\| \mathrm{d}s.$$

Having chosen $\kappa$ to satisfy this bound, preconditioned GMRES applied to this system will converge uniformly, independent of the discretization parameters.

**Remark 3.** While the GMRES convergence estimates shown in this section apply in the case of the modified (iteratively lagged) VEF discretization with sufficiently large penalty parameter, in practice we observe uniform convergence for the non-lagged VEF discretizations, without additional conditions on the size of the penalty parameter $\kappa$. This behavior is typical of domain decomposition algorithms applied to nonsymmetric and indefinite problems, for which the theoretical convergence estimates tend to be pessimistic [46].

**Remark 4 (AMG convergence).** The practical subspace correction preconditioner is obtained by replacing $A_V^{-1}$ (the inverse of the continuous discretization, which is infeasible to compute for large problems) with a tractable approximation $\tilde{A}_V^{-1}$, such as one V-cycle of algebraic multigrid, cf. Theorem 1. This procedure relies on $\tilde{A}_V^{-1}$ well approximating $A_V^{-1}$ (i.e. spectrally equivalent in the symmetric case). AMG performance may suffer on highly non-symmetric problems, and so in the following sections, we consider also choosing $\tilde{A}_V^{-1}$ to be one V-cycle of AMG built with a symmetrized version of the continuous operator $A_V$. 19
7. Results

The VEF algorithms described in this paper were implemented using the MFEM [59, 60] finite element framework. The stabilized bi-conjugate gradient (BiCGStab) and Jacobi solvers from MFEM were used to solve the VEF discretizations along with BoomerAMG, the AMG solver from the sparse linear algebra library hypre [53]. KINSOL, from the Sundials package [61], provided the Anderson-accelerated fixed-point solver. When iterative solver results are not presented, the parallel implementation of the sparse direct solver SuperLU [62] was used. We use the high-order DG $S_N$ transport solver from [1].

Unless otherwise specified, we set the penalty parameter to $\kappa = \left\{ \frac{(p + 1)^2}{\sigma_t h_e} \right\}$ and the BR2 stabilization parameter to $\eta = 4$. We use the MDLDG method, the variant of the LDG method where $\kappa \equiv 0$ and set the upwinding vector $w$ to be a unit vector at a $45^\circ$ angle from the $x$-axis. The VEF discretizations all use the element-local basis defined using the Gauss-Lobatto points to enable the use of the subspace correction preconditioner where required. The transport discretization is always solved with the same finite element order as the VEF scalar flux. However, we use the Gauss-Legendre basis for the transport discretization.

7.1. Method of Manufactured Solutions

The accuracy of the methods are ascertained with the Method of Manufactured Solutions (MMS). The solution is set to $\psi = \frac{1}{4\pi} \left( \sin(\pi x) \sin(\pi y) + \Omega x_1 \Omega y \sin(2\pi x) \sin(2\pi y) + \Omega^2 x \sin\left( \frac{3\pi(x + \delta)}{1 + 2\delta} \right) \sin\left( \frac{3\pi(y + \delta)}{1 + 2\delta} \right) + \gamma \right)$ (87)

where the parameters $\delta = 0.1$ and $\gamma = 0.5$ control the amount of spatially varying, quadratically anisotropic inflow and uniform, isotropic inflow respectively. The computational domain is $D = [0, 1]^2$. With this solution, the Eddington tensor varies in space and has non-zero off-diagonal components. Trigonometric functions are used so that the solution cannot be exactly represented by polynomials. The scalar flux is then

$\phi = \sin(\pi x) \sin(\pi y) + \frac{1}{3} \sin\left( \frac{3\pi(x + \delta)}{1 + 2\delta} \right) \sin\left( \frac{3\pi(y + \delta)}{1 + 2\delta} \right) + \gamma.$

These MMS angular and scalar flux solution functions are substituted into the transport equation to solve for the MMS source function.

The accuracy of the VEF discretizations can be investigated in isolation by computing the VEF data from the MMS angular flux and setting the sources $Q_0$ and $Q_1$ to the moments of the transport MMS source. This is accomplished by computing the VEF data from the MMS angular flux projected onto a finite element space of equal order to the VEF finite element space. An open, Gauss-Legendre basis is used for the angular flux so that the Eddington tensor will have discontinuities of magnitude $O(h^{p+1})$ on interior mesh faces. The VEF data and source moments are computed using level symmetric $S_4$ angular quadrature. The VEF equations are then solved as if $E$, $E_b$, $Q_0$, and $Q_1$ are provided data.

We use refinements of a third-order curved mesh created by distorting an orthogonal mesh according to the velocity field of the Taylor Green vortex. This mesh distortion is generated by advecting the mesh control points with $x = \int_0^T \mathbf{v} \, dt$, (89)

where the final time $T = 0.3\pi$ and

$\mathbf{v} = \begin{bmatrix} \sin(x_1) \cos(x_2) \\ -\cos(x_1) \sin(x_2) \end{bmatrix}$

(90)

is the analytic solution of the Taylor Green vortex. The time integration is calculated with 300 forward Euler time steps. An example mesh is shown in Fig. 3a.
Table 1: MMS error for each method as a function of the maximum characteristic mesh size, \( h \). The standard deviation of the four error values in each row is also provided. Third-order polynomial basis functions were used. The order of accuracy and error constant were computed with logarithmic regression.

| \( h \)     | IP        | BR2       | MDLDG     | CG        | Deviation     |
|------------|-----------|-----------|-----------|-----------|---------------|
| 8.345 \times 10^{-2} | 2.678 \times 10^{-4} | 2.676 \times 10^{-4} | 2.598 \times 10^{-4} | 2.688 \times 10^{-4} | 3.622 \times 10^{-6} |
| 5.564 \times 10^{-2} | 5.163 \times 10^{-5} | 5.158 \times 10^{-5} | 4.837 \times 10^{-5} | 5.169 \times 10^{-5} | 1.415 \times 10^{-6} |
| 4.173 \times 10^{-2} | 1.631 \times 10^{-5} | 1.630 \times 10^{-5} | 1.505 \times 10^{-5} | 1.632 \times 10^{-5} | 5.463 \times 10^{-7} |
| 3.338 \times 10^{-2} | 6.684 \times 10^{-6} | 6.680 \times 10^{-6} | 6.115 \times 10^{-6} | 6.686 \times 10^{-6} | 2.459 \times 10^{-7} |
| Order      | 4.028     | 4.028     | 4.092     | 4.032     |               |
| Constant   | 5.885     | 5.882     | 6.678     | 5.966     |               |

Figure 3b shows the \( L^2(D) \) error between the VEF solution and the exact MMS scalar flux solution as the mesh is refined for the IP, BR2, MDLDG, and CG VEF discretizations when quadratic basis functions are used. Here, \( h \) is the maximum value of the characteristic element length in the mesh. All methods have nearly identical error behavior and converge with third-order accuracy as expected. This experiment is repeated with \( p = 3 \) in Table 1. Logarithmic regression is used to compute the exponent and constant of the error function \( E = C h^\bar{p} \) with \( C \) the constant and \( \bar{p} \) the method’s experimentally-observed order of accuracy. The standard deviation of the four error values for each value of \( h \) is also provided to quantify the variance in the error behavior. Accuracy of \( O(h^{\bar{p}+1}) \) is observed and the four variants are shown to have variance below the discretization error.

7.2. Thick Diffusion Limit

Next, we investigate the iterative convergence properties of the VEF methods in the regime known as the asymptotic thick diffusion limit [14]. The material data is set to

\[
\sigma_t = \frac{1}{\epsilon}, \quad \sigma_a = \epsilon, \quad \sigma_s = \frac{1}{\epsilon} - \epsilon, \quad q = \epsilon
\]

with \( \epsilon \in (0, 1] \) and the thick diffusion limit corresponding to \( \epsilon \to 0 \). A coarse mesh that does not adequately resolve the mean free path is used to stress the convergence of the VEF algorithm. This is a numerically challenging, but common in practice, regime where robust performance is crucial.
Table 2: Number of iterations to convergence in the thick diffusion limit on a coarse, orthogonal mesh.

| $\epsilon$ | IP | BR2 | MDLDG | CG |
|------------|----|-----|-------|----|
| $10^{-1}$  | 8  | 8   | 8     | 8  |
| $10^{-2}$  | 6  | 6   | 6     | 6  |
| $10^{-3}$  | 4  | 4   | 4     | 4  |
| $10^{-4}$  | 3  | 3   | 3     | 3  |

We first demonstrate robust convergence on an $8 \times 8$ linear mesh with $D = [0, 1]^2$. Convergence was identical for linear, quadratic, and cubic basis functions so we present results for $p = 2$ only. Level symmetric $S_4$ angular quadrature is used. Fixed-point iteration without Anderson acceleration is used to solve the coupled transport-VEF system.

Table 2 shows the number of iterations required to converge to a fixed-point tolerance of $10^{-6}$ as $\epsilon \to 0$. All four VEF variants converge robustly and in an identical number of iterations for each value of $\epsilon$. All methods converged to the non-trivial diffusion limit solution. Lineouts of the 2D solutions are shown in Fig. 4 to demonstrate that the non-trivial, diffusion solution is obtained by four methods. Note that even the continuous finite element discretization paired with the discontinuous finite element transport discretization will be robust in the thick diffusion limit.

This experiment is repeated on the triple point mesh shown in Fig. 5. This mesh was generated by running a purely Lagrangian hydrodynamics simulation on a third-order mesh. The mesh contains concave/reentrant interior faces meaning the matrix corresponding to the transport discretization cannot be re-ordered to be strictly lower block triangular. The pseudo-optimally reordered sweep from [1], which lags the incoming angular flux on reentrant faces, is used to enable an element-by-element transport solve. Since the incoming fluxes on reentrant faces are lagged, the angular flux on these faces is not linearly eliminated. In other words, the presence of reentrant faces means that the transport equation is not fully inverted at every fixed-point iteration. In addition, the mesh elements in the “swirl” at the center are severely distorted and thus have poor approximation ability. In practice, the mesh would be remapped before this level of distortion were present. Due to this severe distortion, stability of the IP VEF discretization required scaling the penalty parameter according to

$$\kappa_e = C \frac{(p + 1)^2}{\sigma \chi_e^2},$$

where $C = \max_{K_e \in T} C_e$ with $C_e$ the condition number of the Jacobian matrix for element $K_e$. For the triple point mesh, $C = 169$.

Table 3 shows the number of fixed-point iterations without Anderson acceleration required to converge to a tolerance of $10^{-6}$ for the four VEF variants as $\epsilon \to 0$. Fixed-point convergence is shown when one, two, and three partial transport inversions are applied per fixed-point iteration. While one sweep per fixed-point iteration required more iterations than the equivalent orthogonal-mesh problem, especially for large values of $\epsilon$, the three sweep per fixed-point iteration option had similar convergence properties to its orthogonal-mesh counterpart. This suggests the iterative slow-down can be attributed to the inexact sweep.

Table 4 shows the diffusion scaling on the triple point mesh for the IP VEF method with Anderson acceleration. An Anderson space of size $a$ is used where $a = 0$ is equivalent to fixed-point iteration. We compare convergence when the Anderson space is built from the scalar flux only and when it is built from the scalar and angular fluxes. These variants are referred to as “low memory” and “augmented”, respectively. Note that to simplify the implementation, the augmented Anderson space is built from the entire angular flux and not just the subset of angular flux unknowns corresponding to reentrant faces. The augmented variant saw improvement for $\epsilon = 10^{-4}$ but otherwise converged equivalently to fixed-point iteration. The low memory option was not improved with Anderson acceleration and actually took 1-3 more iterations to converge. Since convergence is primarily hindered by the inexact transport inversion, it is expected that Anderson cannot improve convergence when the Anderson space is not augmented with the angular flux.
Figure 4: Lineouts of the 2D thick diffusion limit solutions taken at $y = \frac{1}{2}$ for the (a) IP, (b) BR2, (c) MDLDG, and (d) CG methods.

Figure 5: A depiction of the triple point mesh used to stress test the VEF algorithms on a severely distorted, third-order mesh. The mesh was generated with a purely Lagrangian hydrodynamics simulation.
Table 3: Number of fixed-point iterations required for convergence on the triple point mesh as $\epsilon \to 0$. On the triple point mesh, reentrant faces mean the transport equation is not fully inverted at each iteration. Each method is tested with 1, 2, and 3 partial transport inversions per fixed-point iteration.

| $\epsilon$ | IP   | BR2  | MDLDG | CG   |
|------------|------|------|-------|------|
| $10^{-1}$  | 19   | 11   | 10    | 23   |
| $10^{-2}$  | 11   | 8    | 7     | 19   |
| $10^{-3}$  | 8    | 5    | 4     | 9    |
| $10^{-4}$  | 6    | 4    | 1     | 6    |

Table 4: The number of Anderson-accelerated fixed-point iterations, with Anderson space of size $a$, required for convergence on the triple point mesh. The interior penalty VEF method is used. The low memory variant builds the Anderson space from the VEF scalar flux only while the augmented version builds the Anderson space from the VEF scalar flux and the angular flux.

| $\epsilon$ | Low Memory $a = 0$ | Augmented $a = 5$ |
|------------|--------------------|-------------------|
| $10^{-1}$  | 19                 | 19                |
| $10^{-2}$  | 11                 | 11                |
| $10^{-3}$  | 8                  | 8                 |
| $10^{-4}$  | 6                  | 6                 |

7.3. Linearized Crooked Pipe

We now demonstrate the efficacy of the methods on a more realistic, multi-material problem. A common benchmark is the crooked pipe problem. The geometry and materials are shown in Fig. 6. The problem consists of two materials, the wall and the pipe, which have an 1000x difference in total interaction cross section. We mock the time-dependent benchmark as a steady-state problem by adding artificial absorption and fixed-source terms corresponding to backward Euler time integration. We use a large time step such that $c\Delta t = 10^3$ with an initial condition $\psi_0 = 10^{-4}$ for all $(x, \Omega) \in D \times S^2$. The absorption and source terms are then

\begin{align}
\sigma_a &= \frac{1}{c\Delta t} = 10^{-3} \frac{1}{\text{cm}}, \\
q &= \frac{1}{c\Delta t}\psi_0 = 10^{-1} \frac{1}{\text{cm}^3 \cdot \text{s} \cdot \text{str}}.
\end{align}

The boundary conditions are

\begin{align}
f = \begin{cases} 
\frac{1}{\pi}, & x = 0 \text{ and } y \in [-1/2, 1/2], \\
0, & \text{otherwise}
\end{cases}
\end{align}

so that radiation enters the pipe at the left side of the domain. We use a Level Symmetric $S_{12}$ angular quadrature set. The zero and scale [63] negative flux fixup, a sweep compatible method that zeros out negativity and rescales so that particle balance is preserved, is used inside the transport inversion to ensure positivity.

A VEF solution to the crooked pipe using the IP method with $p = 2$ is shown in Fig. 7 where a non-uniform mesh is used to adequately resolve the interface between the optically thin pipe and the optically thick wall. Here, we see that VEF does capture the “shadow” induced by the radiation turning the corner around the inner wall. A radiation diffusion solution would non-physically show illumination on the back side of the wall.

The outer fixed-point and inner linear iterative efficiency is demonstrated by refining in $h$ and $p$. Note that to simplify the refinement process we use a uniform mesh. The outer solver is Anderson-accelerated.
Figure 6: Geometry, material data, and boundary conditions for the linearized crooked pipe problem.

Figure 7: VEF scalar flux solution to the linearized crooked pipe problem to show that VEF does capture the transport solution. The mesh is refined at the interface between thick and thin to adequately resolve the material interface. The IP VEF method with $p = 2$ was used.
Table 5: The number of Anderson-accelerated fixed-point iterations until convergence to a tolerance of $10^{-6}$ for the IP, BR2, MDLDG, and CG discretizations of VEF on the linearized crooked pipe problem refined in $h$ and $p$. An Anderson space of size two is used.

| $N_o$ | IP  | BR2 | MDLDG | CG  |
|-------|-----|-----|-------|-----|
|       | 112 | 10  | 10    | 14  | 10  |
| $p = 1$ | 448 | 11  | 11    | 16  | 12  |
|       | 1792 | 13  | 13    | 16  | 13  |
|       | 7168 | 14  | 14    | 17  | 14  |
| $p = 2$ | 112 | 13  | 13    | 16  | 13  |
|       | 448 | 15  | 15    | 16  | 15  |
|       | 1792 | 16  | 16    | 17  | 16  |
|       | 7168 | 17  | 17    | 19  | 17  |
| $p = 3$ | 112 | 15  | 15    | 17  | 15  |
|       | 448 | 16  | 16    | 18  | 16  |
|       | 1792 | 17  | 17    | 19  | 17  |
|       | 7168 | 18  | 18    | 19  | 18  |

fixed-point iteration with two Anderson vectors. Anderson acceleration is not required for convergence on this problem but does provide more uniform convergence in $h$. Since the mesh is orthogonal, the transport equation is fully inverted at each outer iteration. This allows use of the low memory variant so that the storage cost of Anderson acceleration is two scalar flux-sized vectors. The outer tolerance is $10^{-6}$. The inner BiCGStab tolerance is $10^{-8}$. The uniform subspace correction (USC) preconditioner with one Jacobi iteration and one AMG V-cycle per application is used for the IP and BR2 discretizations. The CG and MDLDG discretizations use one V-cycle of AMG as a preconditioner. The previous outer iteration is used as an initial guess to BiCGStab so that the initial guess becomes progressively better as the outer iteration converges.

Table 5 shows the number of outer Anderson-accelerated fixed-point iterations as the mesh is refined and the polynomial order is increased. At each refinement and polynomial order, the IP, BR2, and CG methods converge equivalently with MDLDG converging 1-4 iterations slower. Table 6 shows the maximum, minimum, and average number of inner BiCGStab iterations performed at each outer iteration. All methods are scalable in $h$ and $p$ with CG requiring the fewest iterations followed by MDLDG and then IP and BR2.

7.4. Weak Scaling

Finally, we show that the IP VEF system with $p = 2$ can be solved efficiently in parallel on larger problems. The parallel partitioning is such that there are $\sim 9000$ VEF scalar flux unknowns per processor. The results were generated on 32 nodes of the rztopaz machine at LLNL which has two 18-core Intel Xeon E5-2695 CPUs per node.

First, we investigate weak scaling on a mock problem where the VEF data are provided as inputs to the problem (as opposed to being solved for through fixed-point iteration). This allows the VEF system to be solved in isolation from the transport equation. We use the materials, geometry, and boundary conditions from the crooked pipe problem shown in Fig. 6 but set the Eddington tensor and boundary factor to

$$
E = \begin{cases} 
\begin{bmatrix} 9/11 & 0 \\
0 & 1/11 \\
1/3 & 0 \\
0 & 1/3 
\end{bmatrix}, & x \in \text{pipe} \\
\end{cases}, 
$$

(95a)

$$
E_b = \begin{cases} 
9/10, & x \in \partial(\text{pipe}) \\
1/2, & x \in \partial(\text{wall}) 
\end{cases}. 
$$

(95b)
Table 6: The maximum, minimum, and average number of inner BiCGStab iterations until convergence to an inner tolerance of $10^{-8}$ across all the outer iterations for each of the VEF methods. The previous outer iterate is used as the initial guess for the inner solver so that the number of inner iterations decreases as the outer iteration converges.

| $p$ | IP | BR2 | MDLDG | CG |
|-----|----|-----|-------|----|
|     | $N_e$ | Max | Min | Avg. | Max | Min | Avg. | Max | Min | Avg. | Max | Min | Avg. |
| 1   | 112  | 16  | 6   | 11.50 | 16  | 6   | 11.10 | 10  | 3   | 6.50  | 6   | 3   | 4.80  |
|     | 448  | 17  | 7   | 12.00 | 16  | 7   | 11.27 | 11  | 3   | 6.88  | 7   | 2   | 4.75  |
|     | 1792 | 18  | 4   | 11.23 | 18  | 5   | 11.54 | 11  | 3   | 7.00  | 7   | 2   | 4.85  |
|     | 7168 | 18  | 6   | 11.50 | 18  | 6   | 11.50 | 12  | 3   | 7.24  | 8   | 2   | 5.00  |
| 2   | 112  | 16  | 5   | 10.69 | 16  | 5   | 10.54 | 11  | 4   | 7.75  | 9   | 3   | 6.23  |
|     | 448  | 17  | 5   | 11.20 | 17  | 5   | 10.87 | 12  | 5   | 8.62  | 10  | 3   | 6.07  |
|     | 1792 | 17  | 4   | 11.12 | 16  | 4   | 10.88 | 14  | 5   | 9.35  | 10  | 3   | 6.38  |
|     | 7168 | 17  | 5   | 11.18 | 17  | 5   | 11.06 | 14  | 4   | 9.21  | 11  | 3   | 6.41  |
| 3   | 112  | 19  | 5   | 12.47 | 17  | 5   | 12.20 | 22  | 5   | 11.53 | 11  | 3   | 7.53  |
|     | 448  | 22  | 7   | 14.00 | 19  | 6   | 13.31 | 17  | 6   | 11.83 | 12  | 5   | 8.56  |
|     | 1792 | 22  | 6   | 14.06 | 20  | 6   | 13.65 | 19  | 6   | 12.37 | 13  | 4   | 8.71  |
|     | 7168 | 22  | 7   | 14.39 | 23  | 6   | 14.39 | 19  | 6   | 12.89 | 13  | 4   | 9.11  |

This corresponds to a linearly-transported/linearly-diffusive angular flux in the wall and an extremely forward peaked solution

$$\psi = \Omega^s$$

in the pipe. The motivation for this choice is that the solvers are predicted to struggle when the Eddington tensor is discontinuous. We stress that this setup does not correspond to a physically realistic problem.

Table 7 shows the number of BiCGStab iterations to convergence for the USC-preconditioned IP VEF system on this mock problem. The columns of the table parameterize the solver used for the continuous stage of the USC preconditioner. The standard USC preconditioner used in the previous results did not converge on this mock problem. However, when a sparse direct solver is used instead of AMG, uniform convergence is recovered. This suggests that AMG is failing to adequately solve the continuous operator (see Remark 4).

Note that AMG is effective on the standard continuous finite element discretization of diffusion. It is then expected that AMG will be effective in approximating the inverse of a symmetrized CG operator. By lagging the terms

$$-\int_{\Gamma_b} \left\{ \frac{\nabla u}{\sigma_t} \right\} : [E\varphi\hat{n}] \, ds + \int \nabla u \cdot \frac{1}{\sigma_t} (\nabla_h \cdot E) \varphi \, dx$$

in the CG VEF discretization (Eq. 75), a symmetric operator more amenable to accurate inversion via AMG is found. The symmetrized operator is then:

$$\int_{\Gamma_h} E_h u \varphi \, ds + \int \nabla_h u \cdot \frac{1}{\sigma_t} E \nabla_h \varphi \, dx + \int \sigma_a u \varphi \, dx,$$

with $u, v \in V_p$. This is a CG discretization of

$$-\nabla \cdot \frac{1}{\sigma_t} E \nabla \varphi + \sigma_a \varphi$$

which corresponds to the VEF drift-diffusion equation where the advective term $(\nabla \cdot E) \varphi$ is lagged and moved to the right hand side.

The “AMG-S” column of Table 7 shows the convergence on the mock problem when one AMG V-cycle is applied to the symmetrized CG operator instead of the non-symmetric CG operator. The method converges
Table 7: Weak scaling the IP VEF method with $p = 2$ on a non-physically difficult problem with mock VEF data. The preconditioner is parameterized by the method used for the approximate inverse of the CG operator. The standard USC preconditioner with AMG on the CG operator did not converge due to the large discontinuity in the VEF data. Convergence is recovered by applying AMG to a symmetrized version of the CG operator.

| Processors | DOF  | AMG  | Direct | AMG-S | AMG-S3 |
|------------|------|------|--------|-------|--------|
| 36         | 326,592 | 22   | 25     | 20    |        |
| 72         | 630,000  | 24   | 27     | 21    |        |
| 144        | 1,306,368 | 20   | 25     | 22    |        |
| 288        | 2,570,652 | 22   | 25     | 25    |        |
| 576        | 5,225,472 | 22   | 28     | 23    |        |
| 864        | 7,805,952 | 22   | 26     | 23    |        |
| 1,152      | 10,384,668 | 21   | 29     | 27    |        |

Table 8: Weak scaling the IP VEF method with $p = 2$ on the first iteration of the linearized crooked pipe problem. A parallel block Jacobi sweep is used to generate the VEF data needed to form the VEF system. On this physically realistic problem, both the standard USC and USC with symmetrized CG operator converge uniformly. The iterative efficiency is compared to solving IP radiation diffusion.

| Processors | DOF  | USC  | USC-S | Diffusion |
|------------|------|------|-------|------------|
| 36         | 326,592 | 20   | 20    | 15         |
| 72         | 630,000  | 19   | 20    | 14         |
| 144        | 1,306,368 | 20   | 18    | 13         |
| 288        | 2,570,652 | 22   | 18    | 15         |
| 576        | 5,225,472 | 22   | 20    | 15         |
| 864        | 7,805,952 | 22   | 21    | 15         |
| 1,152      | 10,384,668 | 22   | 21    | 16         |

and is roughly uniform in iteration counts as the mesh is refined. The “AMG-S3” column corresponds to the use of three iterations of an inner Richardson iteration to approximate the inverse of the non-symmetric CG operator. The Richardson iteration is preconditioned using one V-cycle of AMG on the symmetrized CG operator. In other words, this option approximates the inverse of the non-symmetric CG operator with three approximate inversions of the symmetrized CG operator. For this option, iterative efficiency generally fell between that of the sparse direct solver and using only AMG on the symmetrized CG operator. Inner iterations do reduce the number of total iterations to convergence but, since three V-cycles are performed per preconditioner application, not to the degree that fewer V-cycles are performed.

Next, we show weak scaling of the IP VEF linear solve on the first outer iteration of the linearized crooked pipe problem from Section 7.3 with $p = 2$. One parallel block Jacobi transport sweep is performed to provide angular fluxes to compute the VEF data. The VEF system is then solved using BiCGStab. Table 8 shows the number of iterations to convergence to a tolerance of $10^{-6}$ for BiCGStab preconditioned with USC and USC where the continuous operator is symmetrized. In addition, the number of iterations to solve the corresponding IP diffusion problem (by setting $E = \frac{1}{4}I$ and $E_b = 1/2$) with the USC preconditioner are shown. These results indicate that on a physically realistic problem the standard USC preconditioner and USC preconditioner with the symmetrized CG operator are both effective. Compared to IP diffusion, IP VEF only required 5-7 more iterations. Since no problems where the USC preconditioner with symmetrized CG operator failed to converge were found, the preconditioner with symmetrized CG operator may be more robust. Note that this discrepancy was not observed on physically realistic problems.
8. Conclusions

We have developed a family of high-order discretizations of the Variable Eddington Factor (VEF) equations that are compatible with curved meshes and have efficient preconditioned iterative solvers. When combined with a high-order Discontinuous Galerkin (DG) discretization of Discrete Ordinates ($S_N$) transport, the resulting VEF methods are efficient in both outer fixed-point iterations and inner linear iterations on a challenging proxy problem from thermal radiative transfer (TRT). We adapted the unified framework for DG methods for elliptic problems presented in [2] to the VEF equations to derive analogues of the interior penalty (IP), second method of Bassi and Rebay (BR2), Minimal Dissipation Local Discontinuous Galerkin (MDLDG), and continuous finite element (CG) methods. The uniform subspace correction preconditioner developed by Pazner and Kolev [3] was shown to be effective leading to iteration counts independent of the mesh size, polynomial order, and penalty parameter. GMRES convergence estimates for the preconditioned system were derived for the nonsymmetric VEF system of equations under relatively mild assumptions. Numerical results demonstrate that these preconditioners remain effective on a large range of practical problems. The MDLDG and CG discretizations were effectively preconditioned by Algebraic Multigrid (AMG).

The VEF methods were verified to converge with $O(h^{p+1})$ on refinements of a third-order mesh using a quadratically-anisotropic manufactured solution. They were also tested in the thick diffusion limit both on an orthogonal mesh and a severely distorted third-order mesh generated with a Lagrangian hydrodynamics code. In both cases, all the VEF methods preserved the thick diffusion limit and converged robustly. Convergence on the triple point mesh indicates that these methods are robust to extreme mesh distortions and inexact transport inversions arising from reentrant faces. In addition, the CG VEF method converged equivalently to the DG methods.

The methods were also tested on the linearized crooked pipe problem. This problem has a 1000x difference in total cross section and was designed to emulate a time-dependent TRT calculation. Using the stabilized bi-conjugate gradient method (BiCGStab), all VEF methods were efficiently solvable independent from the mesh size, polynomial order, and, if present, penalty parameter. Using a small Anderson space of size two, all VEF methods were solved in a uniform number of outer Anderson-accelerated fixed-point iterations as well. Convergence was nearly identical for all the VEF methods including the CG discretization of VEF. This indicates that a CG VEF method, which has fewer unknowns than an analogous DG method and requires simpler preconditioning strategies, is an effective transport scheme.

Finally, the IP VEF method was solved in parallel in a weak scaling study. A non-physically difficult mock problem with discontinuous VEF data was found to cause non-convergence of the linear solver. Uniform convergence was recovered by preconditioning the non-symmetric CG operator used in the subspace correction preconditioner with one V-cycle of AMG applied to a symmetrized version of the CG operator. The non-symmetric and symmetric subspace correction preconditioners were then applied to the more realistic crooked pipe problem where both preconditioners led to uniform convergence out to 10 million VEF scalar flux unknowns and 1152 processors. Compared to an equivalent radiation diffusion discretization, VEF only required 5-7 more linear iterations.

The primary takeaway from this work is that all of the VEF methods presented here are strong candidates for implementation in a TRT algorithm. All the methods were robust to the thick diffusion limit, inexact sweeps from reentrant faces, and strongly heterogeneous materials. The methods are primarily differentiated by the ease of their implementation of which the CG VEF method is the simplest. Future work includes extending these algorithms to thermal radiative transfer and verifying their efficacy on a full TRT problem such as the Marshak wave.

9. Acknowledgements

This work was performed under the auspices of the U.S. Department of Energy by Lawrence Livermore National Laboratory under Contract DE-AC52-07NA27344 (LLNL-JRNL-829396). S.O. was supported by the U.S. Department of Energy, Office of Science, Office of Advanced Scientific Computing Research, and the
References

[1] T. S. Haut, P. G. Maginot, V. Z. Tomov, B. S. Southworth, T. A. Brunner, and T. S. Bailey, “An efficient sweep-based solver for the $S_N$ equations on high-order meshes,” Nuclear Science and Engineering, 2019.
[2] D. N. Arnold, F. Brezzi, B. Cockburn, and L. D. Marini, “Unified analysis of discontinuous Galerkin methods for elliptic problems,” SIAM Journal on Numerical Analysis, vol. 39, no. 5, pp. 1749–1779, 2002.
[3] W. Pazner and T. Kolev, “Uniform subspace correction preconditioners for discontinuous Galerkin methods with hp-refinement,” Communications on Applied Mathematics and Computation, Jul. 2021.
[4] D. Mihalas, Stellar Atmospheres. W. H. Freeman and Co, 1978.
[5] V. Ya. Gol’din, “A quasi-diffusion method of solving the kinetic equation,” USSR Comp. Math. and Math. Physics, vol. 4, pp. 136–149, 1964.
[6] E. Aristova and D. Baydin, “Implementation of the quasidiffusion method for calculating the critical parameters of a fast neutron reactor in 3D hexagonal geometry,” Mathematical Models and Computer Simulations, vol. 5, pp. 145–155, 2013.
[7] A. Tamang and D. Y. Anistratov, “A multilevel projective method for solving the space-time multigroup neutron kinetics equations coupled with the heat transfer equation,” Nuclear Science and Engineering, vol. 177, no. 1, pp. 1–18, 2014.
[8] D. Y. Anistratov, E. N. Aristova, and V. Y. Gol’din, “A nonlinear method for solving the problems of radiation transfer in medium,” Matematicheskoe modelirovanie, vol. 8, no. 12, pp. 3–28, 1996.
[9] D. Y. Anistratov and V. Y. Gol’din, “Comparison of difference schemes for the quasidiffusion method for solving the transport equation,” Problems of Atomic Science and Engineering: Method and Codes for Numerical Solution Mathematical Physics Problems, vol. 2, pp. 17–23, 1986.
[10] E. Aristova and V. Gol’din, “Computation of anisotropy scattering of solar radiation in atmosphere (monoenergetic case),” Journal of Quantitative Spectroscopy and Radiative Transfer, vol. 67, no. 2, pp. 139–157, 2000.
[11] R. Alcouffe, “Diffusion synthetic acceleration methods for the diamond-differenced discrete-ordinates equations,” Nuclear Science and Engineering, vol. 64, pp. 344–355, 1977.
[12] D. Y. Anistratov and V. Y. Gol’din, “Nonlinear methods for solving particle transport problems,” Transport Theory and Statistical Physics, vol. 22, no. 2-3, pp. 125–163, 1993.
[13] J. Warsa and D. Anistratov, “Two-level transport methods with independent discretization,” Journal of Computational and Theoretical Transport, vol. 47, no. 4-6, pp. 424–450, 2018.
[14] E. W. Larsen, J. Morel, and J. Warren F. Miller, “Asymptotic solutions of numerical transport problems in optically thick, diffusive regimes,” Journal of Computational Physics, vol. 69, pp. 283–324, 1987.
[15] P. Ghassemi and D. Y. Anistratov, “Multilevel quasidiffusion method with mixed-order time discretization for multigroup thermal radiative transfer problems,” Journal of Computational Physics, vol. 409, p. 109315, 2020.
[16] B. Yee, S. Olivier, B. Southworth, M. Holec, and T. Haut, “A new scheme for solving high-order DG discretizations of thermal radiative transfer using the variable Eddington factor method,” in Proceedings of the International Conference on Mathematics and Computational Methods Applied to Nuclear Science and Engineering (M&C 2021), 2021.
[17] D. Y. Anistratov and J. M. Coale, “Implicit methods with reduced memory for thermal radiative transfer,” 2021.
[18] Y.-F. Jiang, J. M. Stone, and S. W. Davis, “A Godunov method for multidimensional radiation magnetohydrodynamics based on a variable Eddington tensor,” The Astrophysical Journal Supplement Series, vol. 199, no. 1, p. 14, 2012.
[19] N. Y. Gnedin and T. Abel, “Multi-dimensional cosmological radiative transfer with a variable Eddington tensor formalism,” New Astronomy, vol. 6, no. 7, pp. 437–455, 2001.
[20] M. Gehmeyr and D. Mihalas, “Adaptive grid radiation hydrodynamics with TTITAN,” Physica D: Nonlinear Phenomena, vol. 77, no. 1, pp. 320–341, 1994.
[21] S. W. Davis, J. M. Stone, and Y.-F. Jiang, “A radiation transfer solver for Athena using short characteristics,” The Astrophysical Journal Supplement Series, vol. 199, no. 1, p. 9, Feb 2012.
[22] S. S. Olivier and J. E. Morel, “Variable Eddington factor method for the $S_N$ equations with lumped discontinuous Galerkin spatial discretization coupled to a drift-diffusion acceleration equation with mixed finite-element discretization,” Journal of Computational and Theoretical Transport, vol. 46, no. 6-7, pp. 480–496, 2017.
[23] J. Lou, J. E. Morel, and N. Gentile, “A variable Eddington factor method for the 1-D grey radiative transfer equations with discontinuous Galerkin and mixed finite-element spatial differencing,” Journal of Computational Physics, vol. 393, pp. 258–277, 2019.
Appendix A. Implementation of Lifting Operators

Consider the face-local lifting operator \( \rho_f(\omega) \) used in the BR2 stabilization term defined in Eq. 60 with \( \omega = [u] \) which satisfies

\[
\int v \cdot \rho_f([u]) \, dx = - \int_{\Gamma_0} \| \nabla \cdot \hat{n} \| [u] \, ds, \quad \forall v \in W_p, \quad \text{on } \Gamma_0. \tag{A.1}
\]

Let \( y \) represent the vector of DOFs corresponding to a \( Y_p \) or \( W_p \) grid function \( y \). Let \( v, w \in W_p \) and define

\[
v^T M_w = \int v \cdot w \, dx \tag{A.2}
\]

as the \( W_p \) mass matrix. Further, define

\[
y^T A_f y = - \int_{\Gamma_0} \| \nabla \cdot \hat{n} \| [u] \, ds, \quad \text{on } \Gamma_0, \tag{A.3}
\]

for \( u \in Y_p \). Equation A.1 is then equivalent to

\[
M \rho_f([u]) = A_f y \iff \rho_f([u]) = M^{-1} A_f y. \tag{A.4}
\]

Since the \( W_p \) mass matrix is block diagonal by element, its inverse can be computed and stored without fill-in by simply inverting each block individually. The BR2 stabilization term can then be written as

\[
\sum_{f \in \Gamma_0} \int \rho_f([u]) \cdot \rho_f([\varphi]) \, dx = \sum_{f \in \Gamma_0} \rho_f([u])^T M_{\rho_f}([\varphi]) = \sum_{f \in \Gamma_0} y^T A_f^T M^{-1} M A_f \varphi \tag{A.5}
\]

\[
= \sum_{f \in \Gamma_0} y^T A_f^T M^{-1} A_f \varphi
\]
since $M$ is symmetric. Again, since $M^{-1}$ is block diagonal by element and the products $A_f \tilde{\varphi}$ and $u^T A_f^T \tilde{\varphi}$ are non-zero only on DOFs that share the face $f$, each argument of the sum only contributes to the DOFs that share the face $f$. Due to this, the matrix $\sum_{f \in \Gamma_0} A_f^T M^{-1} A_f$ can be assembled face by face.

Next, consider one part of the LDG stabilization term:

$$\int \rho_0([u]) \cdot r_0([E_f \tilde{n}]) \, dx. \quad (A.6)$$

Let,

$$u^T B \tilde{\varphi} = - \int_{\Gamma_0} \{ v \} \cdot [E_f \tilde{n}] \, ds, \quad (A.7)$$

and further define the total interaction $W_p$ mass matrix as

$$\varepsilon_T w = \int_{\Gamma_0} \sigma_t \cdot \varepsilon \cdot \varepsilon \, dx, \quad (A.8)$$

so that $r_0([E_f \tilde{n}]) = M^{-1}_t B \tilde{\varphi}$. In addition, define

$$u^T A u = - \int_{\Gamma_0} \{ v \} \cdot \hat{n} \cdot \{ u \} \, ds, \quad (A.9)$$

such that $A = \sum_{f \in \Gamma_0} A_f$. The LDG stabilization term under consideration is then

$$\int \rho_0([u]) \cdot r_0([E_f \tilde{n}]) \, dx = \rho_0([u])^T M_{\epsilon_0}([E_f \tilde{n}])$$

$$= u^T A^T M^{-T} M M_t^{-1} B \tilde{\varphi}$$

$$= u^T A^T M_t^{-1} B \tilde{\varphi}. \quad (A.10)$$

Note that since the matrices $A$ and $B$ are not face-local, this term cannot be assembled locally. The LDG stabilization term is instead formed through matrix multiplication as $A^T M_t^{-1} B$. 

33