Mesh Refinement for Anisotropic Diffusion in Magnetized Plasmas

Christopher J. Vogl*a, Ilon Josepha, Milan Holeca

aLawrence Livermore National Laboratory, 7000 East Ave, Livermore, 94550, CA, U.S.

Abstract

Highly accurate simulation of plasma transport is needed to drive the successful design and operation of magnetically confined fusion reactors. Unfortunately, the extreme anisotropy present in magnetized plasmas results in thin boundary layers that are expensive to resolve. This work investigates how various mesh refinement strategies might reduce that expense to allow for more efficient simulation by comparing standard variable refinement approaches that use a field quantity to an adaptive approach that uses an error estimator. It is first verified that higher order discretization only realizes the proper rate of convergence once the mesh resolves the thin boundary layer, therefore motivating the focusing of refinement on the boundary layer. For three two-dimensional test cases that contain characteristic features of tokamak magnetic fields, an exponential refinement strategy based on the magnetic flux function, which is the standard refinement approach in the field, is compared to an adaptive strategy utilizing the established Zieniewiecz and Zhu error estimator. The adaptive mesh refinement strategy consistently achieves the same accuracy using orders of magnitude less degrees of freedom than either exponential or uniform refinement. This result makes the adaptive refinement strategy more efficient than the exponential refinement strategy while also being more generalizable to problems with complex magnetic geometries. Scaling laws are derived that quantify the improvement in cost of the adaptive refinement strategy over other refinement approaches in 2D and 3D.

Keywords:
anisotropic diffusion, adaptive mesh refinement, magnetic confinement fusion, boundary layers, finite element method, high-order methods

1. Introduction

The design of magnetically confined fusion reactors relies critically on the ability to accurately model plasma transport [1, 2]. To this end, the magnetized confinement fusion community has developed highly sophisticated plasma transport modeling tools such as UEDGE [3] and SOLPS-ITER [4], which typically consider transport processes, including atomic and molecular physics, in a fixed background magnetic field. Due to the interest in understanding the transport effects of perturbations to the magnetic field, general multi-fluid magnetohydrodynamics solvers, such as NIMROD [5, 6], M3D-C1 [7], and JOREK [8] have been developed that also have the ability to model transport processes in an evolving
magnetic field. These codes are often based on finite volume and finite element approaches in order to precisely satisfy conservation laws for particles, momentum, and energy.

The physics of strongly magnetized plasmas is highly anisotropic with respect to the direction of the magnetic field. To lowest order, a charged particle must travel in a helical orbit around a magnetic field line, which implies that plasma transport is fast along field lines, but slow across the field lines [9]. The fastest transport process is the parallel, so-called Spitzer-Härm [10] thermal conduction coefficient $\kappa_\parallel$, due to the lightest particles, the electrons. In contrast, the perpendicular thermal conduction coefficient $\kappa_\perp$, due to collisions is small enough that it is typically dominated by turbulent transport processes. The ratio between the parallel and thermal conduction coefficients, $\kappa_\parallel/\kappa_\perp$, can be as large as $10^9 - 10^{12}$ at the edge of a tokamak and gets higher still as the core temperature of 10-15 keV is approached. In the core, fluid transport is no longer accurate and kinetic transport processes such as neoclassical transport [11] must be considered.

At the edge of a magnetized plasma, the boundary conditions conspire to generate a narrow boundary layer, traditionally called the "scrape-off layer" (SOL). The width of this layer relative to the minor radius of the device scales as the square root of the inverse anisotropy ratio, $\sqrt{\kappa_\perp/\kappa_\parallel}$. Thus, even if turbulent perpendicular thermal transport reduces the anisotropy ratio to $10^6 - 10^8$, this still requires resolving a boundary layer width that is on the order of $10^{-4} - 10^{-3}$ of the domain size. This is very challenging for 3D simulations and can still be challenging in 2D, unless care is taken to reduce the complexity of the problem.

Within the finite element method framework, this work investigates how various approaches to refinement might capture the boundary layer dynamics while reducing computational effort away from the boundary layer (i.e., reduction in complexity of the problem). The focus is more on mesh refinement ($h$-refinement) than on using higher-order polynomial spaces ($p$-refinement), because it is shown herein that the mesh itself must resolve the boundary layer before the expected higher-order accuracy with $p$-refinement is realized. Standard variable refinement strategies using the magnetic flux function [2, 12] and an adaptive strategy using an error estimation are leveraged with the goal of achieving the same accuracy as a uniform refinement approach but with substantially fewer degrees of freedom. In a test case where the magnetic field is aligned with the mesh, two refinement strategies are investigated: one that focuses the refinement across the layer using rectangular elements with an aspect ratio equal to the boundary layer width and one that allows for exponential growth in mesh spacing away from the layer. Both strategies show substantial improvement over uniform refinement in efficiency, although the aspect-ratio strategy does not generalize to cases where the magnetic field is no longer aligned with the mesh. For three non-aligned, two-dimensional test cases that contain characteristic features of tokamak magnetic fields, the exponential refinement strategy based on the magnetic flux function, which is the standard refinement approach in the field, is compared to an adaptive strategy utilizing the established Zienkiewicz and Zhu error estimator. The adaptive mesh refinement (AMR) strategy requires orders of magnitude less degrees of freedom than either the exponential or uniform strategies. Thus, it is found that the AMR strategy is substantially more efficient than the standard strategies for the problems investigated. While those problems use constant thermal conductivity and magnetic fields that are simple approximations of those expected in a realistic tokamak, the expectation is that the ability of AMR to overcome the challenges found in this study will generalize to more complex nonlinear models of thermal transport and more realistic models...
of the magnetic field geometry.

This work also includes a discussion generalizing the results to larger magnetic anisotropy ratios and to three-dimensional problems. It is shown that, relative to uniform refinement, isotropic AMR reduces the number of degrees of freedom (dofs) by the ratio of the layer width to domain size. Still, this requires a number of dofs on the order of this ratio in 2D or the square of this ratio in 3D. In principle, it is shown anisotropic AMR might only require a number of dofs that grows with the log of this ratio in any dimensions. It is also shown that the number of conjugate gradient iterations scales as a power of the ratio when preconditioned with algebraic multigrid, whereas the number of iterations is independent of the ratio when preconditioned with incomplete LU. All the results herein motivate the replacement of current strategies with anisotropic AMR using the Zienkiewicz and Zhu error estimator with scalable incomplete LU preconditioning strategies to reduce the total computational work in production tokamak simulation codes.

The presentation of this work begins with background information in Section 2 on steady state plasma transport in magnetically confined fusion devices, followed by the mathematic modeling of that steady state plasma transport in a variety of magnetic fields that are meant to approximate different aspects of the complex magnetic fields in a tokamak reactor, including planar and spatially varying magnetic fields. Section 3 describes the finite element discretization used, including arbitrary order polynomial function spaces, as well as the mesh refinement and associate error estimation techniques used. The resulting numerical results are presented in Section 4, including a quantitative measurement of computational efficiency across methods. Section 5 discusses scaling laws for degrees of freedom and condition number as means to comment on the computational cost when the anisotropy ratio is increased beyond the values in Section 4, as well as when three-dimensional problems are considered. The paper concludes with a summary of the results and a brief discussion of future research directions in Section 6.

2. Steady State Plasma Transport

Plasma transport in magnetically confined fusion devices is a highly developed subject. The fluid equations in a strongly magnetized plasma are often referred to as the Braginskii equations after the influential review article [9]. A concise introduction to these equations is given in [13].

2.1. Magnetic Confinement

First, it is important to understand the magnetic field geometry used to confine the plasma fuel. In a magnetic field, charged particles follow a helical gyro-motion that ensures they are tightly bound to field lines. Bending the field lines into the shape of a torus ensures that the charged particles are trapped (to lowest order in an adiabatic expansion). In order to control the slow drift of charged particles across field lines, as well as plasma turbulence and magnetohydrodynamic instabilities, there must be both a toroidal magnetic field, \( B_t \), pointing the long way around the torus and a poloidal magnetic field, \( B_p \), pointing the short way around the torus [2]. In order to have a simple yet concrete picture in mind, it is useful to consider the tokamak configuration, shown in Fig. 1 which is idealized as having a perfectly axisymmetric magnetic field. The major radius of the torus will be denoted \( R \) and
The tokamak is a toroidal magnetic confinement device with major radius \( R \) and minor radius \( a \). The magnetic field has a component, \( B_t \), in the toroidal direction going the long way around the torus and a component, \( B_p \), in the poloidal direction going the short way around the torus. A typical magnetic field line trajectory is shown in red. The separatrix (white dashed lines) separates the region of closed field lines in the core from the region of open field lines in the scrape-off layer (SOL) that intersect the divertor target plates. The variation of the electron temperature, \( T_e \), is illustrated on the cross-section of the torus to the right.

The minor radius of the torus will be denoted \( a \). The rate at which field lines twist toroidally versus poloidally is called the “safety factor” and is approximately \( q_{mag} \sim B_t a / B_p R \). In a typical tokamak, the ratio of \( B_t / B_p \sim 10 \) and \( R/a \sim 3 \), so that the edge safety factor is of order \( q_{mag} \sim 3^{-4} \).

This work will focus on the important problem of predicting the temperature, \( T(x) \), and heat flux, \( q(x) \), within the plasma volume. Due to the Lawson criterion, there is a range of optimal temperatures for each choice of fusion fuel [1]. The first generation of fusion fuels that are typically under consideration is deuterium (D) and tritium (T) precisely because they have the lowest ignition temperature, which is in the range of 10-15 keV, as well as the highest fusion reaction rates. Yet, the heat fluxes delivered to the surrounding walls and material surfaces must be less than 10 MW/m² in order to maintain structural integrity. In a fusion reactor, this generally requires the regions of high plasma density near the walls to be less than 1 eV – a factor of \( 10^4 \) difference in temperature. The goal of this work is to develop meshing strategies that accurately capture the variation of \( T_e \) over the domain, including the core, the boundary layer outside the separatrix, and along the divertor target plates.
2.2. Heat Flux

For this application area, it is very important to be able to predict both the heat flux and the temperature with high accuracy. The temperature in the interior of the plasma is important for understanding the total stored energy, but it is the heat flux near the material surfaces that is important for ensuring the reactor stays within prescribed operational limits. The heat flux is related to the temperature gradient via the relation

\[ q(x) = -\kappa(x) \cdot \nabla T(x) \] (1)

where \( \kappa(x) \) is the thermal conductivity tensor. The thermal conductivity tensor has the form

\[ \kappa(x) = \kappa_\parallel(x) \hat{b}(x) \otimes \hat{b}(x) + \kappa_\wedge(x) \hat{b}(x) \times + \kappa_\perp(I - \hat{b}(x) \otimes \hat{b}(x)) \] (2)

where \( \hat{b} = B/B \) is a unit vector pointing along the magnetic field direction and \( B = |B| \).

According to the Braginskii equations, the components of the heat flux tensor satisfy the nonlinear scaling laws

\[ \kappa_\parallel(x) \propto T^{5/2}(x)/m^{1/2} \]
\[ \kappa_\perp(x) \propto n^2(x)m^{1/2}/B^2(x)T^{1/2}(x) \]
\[ \kappa_\wedge(x) \propto n(x)T(x)/B(x) \] (3)

where \( m \) is the mass of the charged particles and \( n(x) \) is the particle number density. The ratio of parallel to perpendicular thermal conductivity is extremely large and scales as \( \kappa_\parallel/\kappa_\perp \sim B^2/m \), which implies that it increases as the magnetic field is increased and as the mass is reduced. Thus, this ratio is approximately 3600\times smaller for electrons than for deuterium. Because the focus of this work is on the use of advanced numerical methods for modeling plasma transport, we will eliminate the complexity associated with nonlinear variations with magnetic field, density, and temperature by treating \( \kappa_\parallel \) and \( \kappa_\perp \) as constant in space. More generally, the heat flux also depends on electric current, but this dependence is neglected here for simplicity.

The \( \kappa_\wedge \) term in the heat flux is known as the magnetic drift term and understanding its effect is part of the subject of “neoclassical transport” theory [14, 15]. In a toroidal magnetic field, \( \kappa_\wedge \) generates an effective enhancement of radial perpendicular transport that is approximately \( q_{\text{mag}}^2 \kappa_\perp \). This represents roughly a factor of 10-16 enhancement over the \( \kappa_\perp \). For simplicity, in this work, the drift term \( \kappa_\wedge \) will be neglected.

In today’s fusion experiments, the actual magnitude of perpendicular thermal transport is often enhanced by orders of magnitude over the Braginskii/neoclassical levels. This disagreement with collisional transport theory is attributed to the effect of turbulent processes which convect heat and particles from the heat source towards the edge of the plasma. Plasma transport researchers that model the dynamic equilibrium state of a turbulent magnetized plasma experiment typically adjust the perpendicular thermal conduction coefficient \( \kappa_\perp \) until experimental profiles can be matched [3, 4]. In today’s experiments, this reduces the effective anisotropy ratio to the “measured” range of \( 10^4 - 10^7 \).
2.3. General Test Problem

Consider solving the linear, steady-state anisotropic diffusion problem

\[-\nabla \cdot \left[ \kappa(\hat{b}(x)) \cdot \nabla T(x) \right] = s(x), \quad x \in \Omega,\]

\[\text{BC}(T)(x) = 0, \quad x \in \partial \Omega,\]

\[\kappa(\hat{b}(x)) = \kappa_{||} \hat{b}(x) \otimes \hat{b}(x) + \kappa_{\perp} (I - \hat{b}(x) \otimes \hat{b}(x)),\]  

(4)

where \(T\) is temperature, \(\hat{b}\) is the magnetic field unit vector, \(s\) is a source, and \(\text{BC}\) is an operator on \(T\), and potentially derivatives thereof, specifying the boundary conditions. Choices of \(\hat{b}\), \(s\), and \(\text{BC}\) can create boundary and internal layers of width on the order of \(\sqrt{\kappa_{\perp}/\kappa_{||}}\), motivating an investigation into meshing strategies.

2.4. Boundary Layer Width

The anisotropic thermal conduction tensor in (4) tends to generate narrow boundary layers at the edge of the plasma. Consider, for example, the following 2D choices with \(\hat{b}(x,y) = \hat{x}\) (the unit vector in the \(x\) direction):

\[- \left( \kappa_{||} T_{xx}(x,y) + \kappa_{\perp} T_{yy}(x,y) \right) = \kappa_{||} \sin(x), \quad x \in (0,\pi), \quad y \in (0,\infty),\]

\[T(0,y) = T(\pi,y) = T(x,0) = 0, \quad \lim_{y \to \infty} T_y(x,y) = 0.\]  

(5)

The solution to this problem is

\[T(x,y) = \left(1 - \exp \left[ -y \sqrt{\kappa_{||}/\kappa_{\perp}} \right] \right) \sin(x),\]  

(6)

where the boundary layer at \(y = 0\) has a width proportional to \(\sqrt{\kappa_{\perp}/\kappa_{||}}\). Such a layer is not unique to the simplistic geometry in (5): similar layers are indeed expected within physical tokamaks. In a source-free region, the parallel and perpendicular heat fluxes must balance one another via \(\nabla \cdot q = 0\). The parallel spatial scale is set by the connection length to the wall, \(R_{||} \sim q_{\text{mag}} R\), while the radial boundary layer width, \(w\), is determined by the balance between parallel and perpendicular flows. Assuming that the field lines are approximately straight leads to an approximately separable solution of the form \(T(x) \propto \exp (-y/w + i\ell/R_{||})\) where \(y\) is the radial direction and \(\ell\) is the length along a field line. Thus, the balance can only hold when \(w/R_{||} = (\kappa_{\perp}/\kappa_{||})^{1/2}\). Assume that the simulation domain length scale, \(L\), is on the order of the parallel connection radius, \(L \sim R_{||}\). Then, in order to span all spatial scales from \(L\) to \(w\), uniform refinement requires the number of elements

\[N_{\text{uniform}} = (L/w)^d = (\kappa_{||}/\kappa_{\perp})^{d/2}.\]  

(7)

Thus, considering the extremely high anisotropy ratios observed for strongly magnetized fusion plasmas, addressing the boundary layer with uniform refinement is not possible with present-day computing resources.
2.5. Constant magnetic field problem

Consider Ω as the \((0, \pi) \times (0, 1)\), the magnetic field \(\mathbf{b}(x, y) = \hat{x}\), the source \(s(x, y) = \kappa_\parallel \sin(x)\), and \(\mathcal{BC}\) such that homogeneous Dirichlet conditions are enforced in \(x\) and \(y\):

\[
-\left(\kappa_\parallel T_{xx}(x, y) + \kappa_\perp T_{yy}(x, y)\right) = \kappa_\parallel \sin(x), \quad x \in (0, \pi), \quad y \in (0, 1),
\]

\[
T(0, y) = T(\pi, y) = T(x, 0) = T(x, 1) = 0.
\]

Note the exact solution to (8) is

\[
T(x, y) = \left(1 - \frac{\exp[-y\sqrt{\kappa_\parallel/\kappa_\perp}] + \exp[-(1 - y)\sqrt{\kappa_\parallel/\kappa_\perp}]}{1 + \exp[-\sqrt{\kappa_\parallel/\kappa_\perp}]}\right) \sin(x).
\]

Defining the boundary layer width as where the corresponding exponential goes from value 1 to \(\exp[-1]\), the boundary layers at \(y = 0\) and \(y = 1\) have width \(\sqrt{\kappa_\perp/\kappa_\parallel}\). For scenarios where it may be more natural to think of forcing on \(\partial \Omega\) rather than in \(\Omega\), consider that

\[
\tilde{T}(x, y) = T(x, y) - \sin(x)
\]

satisfies

\[
-\left(\kappa_\parallel \tilde{T}_{xx}(x, y) + \kappa_\perp \tilde{T}_{yy}(x, y)\right) = 0, \quad x \in (0, \pi), \quad y \in (0, 1),
\]

\[
\tilde{T}(0, y) = \tilde{T}(\pi, y) = 0, \quad \tilde{T}(x, 0) = \tilde{T}(x, 1) = -\sin(x).
\]

2.6. Poloidally varying magnetic field problems

While the constant magnetic field in (8) leads to an exact solution, it is indeed very simplistic when compared to fields expected in a magnetically confined fusion reactor. Recall that the magnetic field \(\mathbf{B}\) in such a device must be topologically toroidal and that the total magnetic field is the combination of the poloidal and toroidal fields: \(\mathbf{B}(x) = \mathbf{B}_p(x) + \mathbf{B}_t(x)\). Let us introduce the cylindrical coordinates \(\mathbf{x} = R \cos(\zeta) \hat{x} + R \sin(\zeta) \hat{y} + \hat{z}\), where \(\zeta\) is the toroidal angle. The toroidal field induced by external coils is much larger than that generated by plasma currents, so the toroidal field is approximately \(\mathbf{B}_t \simeq B_{\text{tor}} \hat{\zeta}/R\) where \(B_{\text{tor}}\) is constant. The poloidal field is then defined via \(\mathbf{B}_p = \nabla \times \psi(x) \nabla \zeta\) where \(\psi(x)\) is the poloidal magnetic function; i.e. it measures the magnetic flux through a poloidal section of the torus. The ideal MHD equilibrium conditions [2], which require nested toroidal flux surfaces, imply that the field lines must lie in surfaces of constant \(\psi(x)\).

In order to further simplify the problem to focus on the key issues, we will assume a large aspect ratio toroidal geometry, where the minor radius is much smaller than the major radius \(R\); i.e. so that the major radius can be assumed to be approximately constant. In this case, one can assume that the toroidal field is constant \(\mathbf{B}_t = \hat{z}\) and that the poloidal field is given by \(\mathbf{B}_p = \nabla \times [A_z(x, y) \hat{z}]\), where \(\psi \simeq A_z R\). Furthermore, we will assume that variations along the toroidal direction now given by \(\hat{z}\) vanish, so that \(\mathbf{b}(x, y) \cdot \nabla = \mathbf{B}_p(x, y)/|\mathbf{B}(x, y)| \cdot \nabla\). Three such magnetic fields, illustrated in Fig. 2 are now defined along with the associated test problem to be investigated.

2.6.1. Single-null magnetic field problem

The first poloidally varying magnetic field we shall study is the “two-wire model” of a single null diverted tokamak, with the magnetic flux function

\[
A_z(x) = \log(||x - x_1|| \cdot ||x - x_2||).
\]
Figure 2: Contour plots showing magnetic flux surfaces for poloidally varying problems, specifically (a) $-\exp(A_z)$ for single-null diverted tokamak, (b) $-A_z$ for double-null diverted tokamak, and (c) $-A_z$ for magnetic island.

Note the point $x_s = \frac{1}{2}(x_1 + x_2)$ represents the null point where the poloidal magnetic field vanishes. The level curve through $x_s$, called the separatrix, partitions $\Omega$ into level curves that form closed field lines in the core that do not intersect the boundary and open field lines at the edge that do reach the boundary. Generally, the level curves consist of points $x$ where the product of distances from $x$ to $x_1$ and to $x_2$ remains constant. Considering $\Omega$ to be the unit square, the specific problem to be investigated is

$$-\nabla \cdot \left[ \kappa \left( \hat{b}(x,y) \right) \cdot \nabla T(x,y) \right] = e^{-\frac{1}{2} \left( (\frac{x-1/2}{1/8})^2 + (\frac{y-1/2}{1/8})^2 \right)}, \quad x \in (0,1), \quad y \in (0,1),$$

$$T(0,y) = T(1,y) = T(x,0) = T(x,1) = 0,$$

Placing a heat source, $s(x,y)$, inside the separatrix, traps the heat in the middle of the domain, because it can only slowly diffuse outward at the much smaller rate set by $\kappa_\perp$. Note that the magnetic field geometry drives an internal temperature layer of width on the order of $\sqrt{\kappa_\perp / \kappa_\parallel}$ near the separatrix.

2.6.2. Double null magnetic field problem

The second poloidally varying magnetic field we shall study models the field in a double null diverted tokamak, with the magnetic flux function

$$A_z(x,y) = \frac{1}{2}(x - x_0)^2 + \frac{1}{2} \left( \frac{1}{4} \sin \left[ 2\pi (y - y_0) \right] \right)^2.$$  \hspace{1cm} (12)

Note the points $x_s^\pm = (x_0, y_0 \pm \frac{1}{4})$ act as the two poloidal field nulls. The separatrix curve that passes through those points partitions $\Omega$ in a similar fashion to the single null in (10). Again considering $\Omega$ to be the unit square, the specific problem to be investigated is

$$-\nabla \cdot \left[ \kappa \left( \hat{b}(x,y) \right) \cdot \nabla T(x,y) \right] = e^{-\frac{1}{2} \left( (\frac{x-1/2}{1/8})^2 + (\frac{y-1/2}{1/8})^2 \right)}, \quad x \in (0,1), \quad y \in (0,1),$$

$$T(0,y) = T(1,y) = T(x,0) = T(x,1) = 0,$$

and $A_z(x,y)$ defined by (12) with $(x_0, y_0) = (\frac{1}{2}, \frac{1}{2})$. As with (11), placing a source inside the separatrix focuses the heat generated by $s(x,y)$ in the middle of the domain. The difference
between this problem \([13]\) and \([11]\) is that there are two nulls near which the poloidal magnetic field vanishes. Again, internal boundary layers of width on the order of \(\sqrt{\kappa_\perp/\kappa_\parallel}\) are generated near the separatrix.

2.6.3. Magnetic island problem

The third poloidally varying magnetic field we will study models the way that a perturbation to the magnetic field can cause a topological change in the flux surfaces that produces a “magnetic island.” The magnetic flux function to model such a phenomenon is a modified version of \([12]\), where the perturbation vanishes at the radial boundaries, set by \(x = 0\) and \(x = L\):

\[
A_z(x, y) = \frac{1}{2} (x - x_0)^2 + \frac{1}{2} \left(1 - \frac{x - x_0}{L/2}\right) \left(1 + \frac{x - x_0}{L/2}\right) \sin \left(2\pi (y - y_0)\right) \right)^2.
\]

The magnetic island width is defined by the separatrix which passes through the points \(x_s^\pm = (x_0, y_0 \pm \frac{1}{2})\), formed on the background magnetic flux function \(A_z(x, y) = \frac{1}{2} (x - x_0)^2\). Again considering \(\Omega\) to be the unit square, the specific problem to be investigated is

\[
-\nabla \cdot \left[\kappa(\hat{b}(x, y)) \cdot \nabla T(x, y)\right] = 0, \quad x \in (0, 1), \quad y \in (0, 1),
\]

\[
T(0, y) = 1, \quad T(1, y) = 0,
\]

\[
T(x, 0) = T(x, 1), \quad T_y(x, 0) = T_y(x, 1),
\]

and \(A_z(x, y)\) defined by \([14]\) with \((x_0, y_0) = (\frac{1}{2}, \frac{1}{2})\) and \(L = 1\). For this problem, there is an effective heat source at the \(x = 0\) boundary, which maintains the temperature differential, and heat must pass through the magnetic island to reach the effective sink at \(x = 1\).

3. Numerical Methods

The general steady state heat problem \([4]\) is discretized in space using the finite element method. The corresponding weak/variational form of \([4]\) is that

\[
\int_{\Omega} \left[\kappa(\hat{b}) \cdot \nabla T\right] \cdot \nabla \psi \, d\nu - \int_{\partial \Omega} \psi \left[\kappa(\hat{b}) \cdot \nabla T\right] \cdot \hat{n} \, dS = \int_{\Omega} s \psi \, d\nu \quad \forall \psi(x) \in H^1(\Omega),
\]

\[
T(x) \in \mathcal{V}(\Omega) = \{f(x) \in H^1(\Omega): \mathcal{BC}(f)(x) \text{ for } x \in \partial \Omega = 0\}
\]

where \(H^1(\Omega)\) is a Hilbert space. Denote \(\Omega_h\) as a collection of elements (mesh) that discretizes \(\Omega\). The space of polynomials of order \(p\) that are piecewise defined on \(\Omega_h\) is denoted \(H^1_p(\Omega_h)\). Denote \(\mathcal{V}_h^p(\Omega_h) = \{f(x) \in H^1_p(\Omega_h): \mathcal{BC}(f)(x) = 0 \text{ for } x \in \partial \Omega\}\) and note that \(\mathcal{V}_h^p(\Omega_h) \subset \mathcal{V}(\Omega)\).

The exact solution to \([16]\) is thus approximated by \(T(x) \in \mathcal{V}_h^p(\Omega)\): \(T(x) = \sum_{i=1}^{n} T_i \psi_i(x)\) where \(\{\psi_1(x), \ldots, \psi_n(x)\}\) is a set of basis functions for \(\mathcal{V}_h^p(\Omega_h)\). Because \(\mathcal{V}_h^p(\Omega_h)\) is also a subspace of \(H^1(\Omega)\), the governing equations, \(i = 1, \ldots, n\), for the approximate solution \(T(x)\) are chosen to be

\[
\sum_{j=1}^{n} \left[\int_{\Omega} \nabla \psi_i \cdot \left[\kappa(\hat{b}) \cdot \nabla \psi_j\right] \, d\nu - \int_{\partial \Omega} \psi_i \left[\kappa(\hat{b}) \cdot \nabla \psi_j\right] \cdot \hat{n} \, dS\right] T_j = \int_{\Omega} s \psi_i \, d\nu.
\]

The finite element problem \([17]\) is assembled and solved using the MFEM software library \([16]\). MFEM provides a variety of arbitrary-order function spaces for unstructured meshes, including \(H^1_p(\Omega_h)\). The conjugate gradient method implemented in MFEM is used to solve the system, preconditioned with BoomerAMG using the provided interface to hypre \([17]\).
3.1. Error Estimation

The finite element method provides a variety of advantages that include error estimation techniques. One such estimation technique for the steady state problem (4) that utilizes various function spaces commonly used by the finite element method is from the work of Zienkiewicz and Zhu [18, 19, 20], commonly referred to as ZZ estimation. The key observation is that the flux $-\kappa \nabla T$ in (4) should be continuous across element edges/faces to make the weak divergence of the flux on an element $E$, evaluated as $\int_{\partial E} \kappa \nabla T \cdot \hat{n} dS$, well-defined. In other words, the flux $-\kappa \nabla T$ should be in $H_{\text{div}}(\Omega_h)$. The temperature field $T$ is typically chosen to be in $H^1_p$, which results in $\nabla T \in H_{\text{curl}}(\Omega_h)$.

Zienkiewicz and Zhu surmised that the difference $\tilde{e}_\sigma$ between $\kappa \nabla T$ and its projection to $H_{\text{div}}(\Omega_h)$ will provide an estimate for the difference $e_\sigma$ between the numerical and analytic fluxes. Specifically for $T \in H^1_p$, the value of $\tilde{e}_\sigma$ is defined as $\kappa \nabla T - \sum_{i=1}^{\tilde{n}} \sigma_i \psi_i$, where $\{\psi_1, \ldots, \psi_{\tilde{n}}\}$ is a set of basis functions for $H_{\text{div}}^{p-1}(\Omega_h)$ and $\{\sigma_1, \ldots, \sigma_{\tilde{n}}\}$ is determined by the projection of $\kappa \nabla T$ on $H_{\text{div}}^{p-1}(\Omega_h)$. The MFEM implementation of the ZZ error estimate uses an $L_2$ projection defined by $\{\sigma_1, \ldots, \sigma_{\tilde{n}}\}$ satisfying $\int_\Omega \kappa \nabla T \cdot \psi_i d\mathcal{V} = \sum_{j=1}^{\tilde{n}} \sigma_j \psi_j \cdot \psi_i$ for $i = 1, \ldots, \tilde{n}$. The error estimate is then computed per element $E$ as

$$|\tilde{e}_\sigma|_E = \int_E \|\tilde{e}_\sigma\|_2 d\mathcal{V} \quad (18)$$

It is worth noting if $BC$ enforces Dirichlet or Neumann conditions, then the energy norm of the error in $T$ is equivalent to $-\int_\Omega e_\sigma \cdot \kappa^{-1} e_\sigma d\mathcal{V}$, and therefore can be approximated by $\tilde{e}_\sigma$.

3.2. Mesh Refinement

MFEM supports non-conforming refinement of unstructured meshes given a refinement strategy that identifies elements for refinement or for de-refinement/coarsening. The refinement strategy can depend on the current solution (adaptive refinement) or it can depend on some $a$ priori known variables (variable refinement). This work utilizes one adaptive strategy and one variable strategy for identifying elements for refinement. The adaptive refinement strategy marks elements where the current ZZ error estimate for flux error (18) is greater than $\frac{3}{4 N} \sum_{i=1}^N |\tilde{e}_\sigma|_{E_i}$, $\{E_1, \ldots, E_N\}$ being the collection of elements that define $\Omega_h$. As such, this strategy will be referred to as ZZ refinement.

The variable strategy marks elements that either (i) contain the separatrix curve or (ii) have not reached a target size relative to that of the elements containing the separatrix. For (ii), the ratio of the target size and the size of elements containing the separatrix is given by an exponential function that grows proportional to some measure of distance of the element from the separatrix. As such, this strategy will be referred to as exponential refinement. The exponential growth is motivated by the exact solution (6) and will be derived from approximation theory in Section 4. Note that the variable refinement strategy does require the location of the separatrix to be known $a$ priori, making it the less general of the two strategies.

4. Numerical Results

The finite-element approximation (17) to the steady-state anisotropic diffusion problem (4) is now solved with various magnetic fields to investigate the ability of mesh refinement to
improve the computational efficiency for large $\kappa_\parallel/\kappa_\perp$ values. First, the implementation of the discretization (17) is verified on the constant magnetic field problem (8) with convergence testing against the analytic solution (9). The efficiency of using a mesh whose elements have an aspect ratio equal to the magnetic anisotropy and of using exponential refinement is investigated for the field-aligned mesh in (8). Finally, for the non-field-aligned single null magnetic field (11), double null magnetic field (13), and magnetic island (15) problems, the efficiency of both ZZ and exponential refinement is compared to that of uniform refinement.

4.1. Verification

The implementation of the discretization (17) is first verified by a convergence test on the constant magnetic field problem (8). The mesh $\Omega_h$ partitions $\Omega$ into $m^2$ uniform, rectangular elements (i.e. that have edges of length $h_x = \pi/m$ and $h_y = 1/m$). The resulting linear systems are solved with conjugate gradient preconditioned by hypre’s BoomerAMG across 6 MPI ranks. The convergence test uses an error norm that focuses on the discrete $L^2$ error in and around the boundary layer at $y = 0$:

$$\|T - T^*\| = \left( \frac{1}{15000} \sum_{i=0}^{100} \sum_{j=0}^{150} (T(x_i, y_j) - T^*(x_i, y_j))^2 \right)^{1/2}, \quad x_i = \frac{i}{100} \pi, \quad y_j = \frac{j}{50} \sqrt{\kappa_\perp/\kappa_\parallel},$$

where $T^*(x, y)$ is the exact solution to (17). The test is conducted with piecewise linear ($p = 1$) and piecewise cubic ($p = 3$) function spaces for $m = 10, 20, \ldots, 2560$ and $\kappa_\parallel/\kappa_\perp = 10^2, 10^3, \ldots, 10^6$. Figure 3 verifies that both the linear and cubic approximations of (9) reach the theoretically predicted convergence order once $h_x = h_y$ is within the asymptotic regime. Note that the asymptotic regimes in Figure 3 are consistent with requiring the mesh resolve the boundary layers, that $h_y \leq \sqrt{\kappa_\perp/\kappa_\parallel}$, regardless of whether $p = 1$ or $p = 3$. The solution profiles in Figure 3 also indicate that the mesh itself must resolve the boundary layer to avoid overshoot behavior in the numerical solution.

Approximation theory can explain the over/under-shooting behavior and why one might expect the requirement that the mesh must resolve the boundary layer before the appropriate convergence rate is obtained. Consider the result from Ciarlet and Raviart [21] that provides the bound

$$\|u - \Pi_p u\|_{W^{k,l}(E)} \leq Ch^{p+1-k} \left[ \int_E \sum_{a+b=p+1} \left| \frac{\partial^{p+1} u}{\partial x^a \partial y^b} \right|^l d\mathcal{V} \right]^{1/l},$$

where $\|v\|_{W^{k,l}(E)} := \left[ \int_E \sum_{a+b=c} \left| \frac{\partial^c v}{\partial x^a \partial y^b} \right|^l d\mathcal{V} \right]^{1/l}$ for any element $E$ in $\Omega_h$, where $u$ is a function in Sobolev space $W^{p+1,l}(\Omega)$, $\Pi_p u$ is the piecewise defined polynomial of order $p$ that interpolates $u$. Because the exact solutions (6) and (9) are in $H^{p+1}(\Omega)$, the bound applies with $l = 2$, $u = T$, and $\Pi_p T \in V^p_h(\Omega_h)$. Because the solution (9) lends itself to more compact analysis with results that can generalize to (9).
that has similar boundary layer structure, consider $\kappa//\kappa_\perp \gg 1$ for $T$ from (6):

$$\left[ \int_E \sum_{a+b=p+1} \left( \frac{\partial^{p+1} T}{\partial x^a \partial y^b} \right)^2 dV \right]^{1/2} \sim \left( \sqrt{\frac{\kappa//}{\kappa_\perp}} \right)^{p+1} \left[ \int_E \exp[-y \sqrt{\kappa///\kappa_\perp} \sin(x)]^2 \right]^{1/2},$$

for some $\tilde{y}$ in element $E$ per the integral Mean Value Theorem. Thus, for the discrete $L^2$ error (i.e., the $W^{0,2}(\Omega_h)$ error) to asymptotically decrease at the appropriate rate, one must have $h \sqrt{\kappa///\kappa_\perp} < 1$.

4.2. Mesh refinement for constant magnetic field problem

As seen above, the requirement of the mesh resolving the boundary layer to (i) avoid overshoot behavior in the solution and (ii) capitalize on convergence rates of higher-order polynomial spaces becomes very expensive for uniform meshes. Consider that piecewise polynomial function spaces of order $p$ require $p+1$ degrees of freedom per element in each direction. Continuous function spaces, such as $H_p^1$, share degrees of freedom across elements so that $(p/h_x + 1)(p/h_y + 1)$ degrees of freedom are required for $\Omega_h$ containing $(1/h_x)(1/h_y)$ elements. Thus, resolving the boundary layer of width $\sqrt{\kappa_\perp/\kappa//}$, denoted here as $\bar{w}$, requires on the order of $\kappa///\kappa_\perp$ degrees of freedom. For field-aligned problems, one might consider
using meshes that are more refined across the boundary layer than along the layer. Such meshes are studied for the constant magnetic field problem \(^8\) using \(\kappa||/\kappa_{\perp} = 10^4\) and a piecewise cubic function space. To start, instead of the rectangles in the uniform mesh, rectangles along the y direction with \(m_y/m_x = 100/3\) (i.e., \(h_x/h_y = 100\pi/3 \approx 100\)). A second approach builds on the first by allowing \(h_y\) to vary with \(y\). The form of \(h_y(y)\) is motivated by an observation using the approximation theory bound \(^{19}\) for \(\bar{T}(y) = T(x^*, y)\) with \(x^* \in (0, \pi)\) and \(T(x, y)\) from \(^6\). Consider the one-dimensional element \(E = (y^E, y^{E+h_y})\) where

\[
\left[ \int_{y^E}^{y^{E+h_y}} \left| \frac{d^{p+1}T}{dy^{p+1}} \right|^2 dy \right]^{1/2} = \left( \frac{1}{\bar{w}} \right)^{p+1} \left[ \int_{y^E}^{y^{E+h_y}} \left| \exp[-y/\bar{w}] \sin(x^*) \right|^2 dy \right]^{1/2} \leq \left( \frac{1}{\bar{w}} \right)^{p+1} \left[ \int_{y^E}^{y^{E+h_y}} \left| \exp[-y/\bar{w}] \right|^2 dy \right]^{1/2} = \left( \frac{1}{\bar{w}} \right)^{p+1} \exp[-y^E/\bar{w}] \left[ 1 - \exp[-2h_y^E/\bar{w}] \int_{y^E}^{y^{E+h_y}} dy \right]^{1/2} \leq \pi^{1/2} \sum_{E \in \Omega_h} \left[ \frac{h_y^E}{\bar{w}} \right]^{p+1} \exp[-y^E/\bar{w}] \left[ 1 - \exp[-2h_y^E/\bar{w}] \int_{y^E}^{y^{E+h_y}} dy \right]^{1/2} \Rightarrow \| \bar{T} - \Pi_p \bar{T} \|_{L^2(E)} \leq C \left( \frac{h_y^E}{\bar{w}} \right)^{p+1} \exp[-y^E/\bar{w}] \left[ 1 - \exp[-2h_y^E/\bar{w}] \int_{y^E}^{y^{E+h_y}} dy \right]^{1/2} \Rightarrow \| \bar{T} - \Pi_p \bar{T} \|_{L^2(\Omega_h)}^2 \leq \sum_{E \in \Omega_h} \left[ C \left( \frac{h_y^E}{\bar{w}} \right)^{p+1} \exp[-y^E/\bar{w}] \right]^{2} \frac{1 - \exp[-2h_y^E/\bar{w}]}{2h_y^E/\bar{w}} \int_{y^E}^{y^{E+h_y}} dy \right.}

where \(\| \bar{T} - \Pi_p \bar{T} \|_{L^2(\Omega_h)}^2 = \sum_{E \in \Omega_h} \| \bar{T} - \Pi_p \bar{T} \|_{L^2(E)}^2\) was used. Consider a given desired element size at the boundary layer, denoted by \(h^E_y\), defined relative to boundary layer width: \(h^E_y = h_y \bar{w}\). Note that if one chooses \(h^E_y\) so that

\[
\left[ \left( \frac{h_y^E}{\bar{w}} \right)^{p+1} \exp[-y^E/\bar{w}] \right]^{2} \frac{1 - \exp[-2h_y^E/\bar{w}]}{2h_y^E/\bar{w}} = [h_y^{p+1}]^2 \Rightarrow \| \bar{T} - \Pi_p \bar{T} \|_{L^2(\Omega_h)} \leq \| \bar{T} - \Pi_p \bar{T} \|_{L^2(\Omega_h)} \leq L_y C h_y^{p+1},
\]

where \(L_y\) denotes the length of the domain in the \(y\) direction. Thus, scaling the elements according to \(^{20}\) results in the approximation bound retaining the desired decay behavior as \(h_y\) goes to zero while allowing the mesh elements to grow exponentially in size away from the layer. Note that a similar scaling law will result by solving the optimization problem that minimizes the error bound over a set of element sizes. That all said, strictly enforcing

13
Figure 4: Efficiency diagram for the constant magnetic field problem (8) using uniform \(m_x = m_y\) meshing \((h_x = h_y)\), aspect ratio \(m_y = 100/3m_x\) \((h_x = 100h_y \pi/3)\), and exponential meshing (24) strategies (black dashed lines indicate theoretical power law). Note the aspect ratio and exponential refined meshes require orders of magnitude less degrees of freedom to achieve the same accuracy than is needed by the uniform meshing strategy.

(20) requires solving the following transcendental equation for \(h_y^E\):

\[
(h_y^E)^{p+1}\exp\left[-y^E/\tilde{w}\right]\left[1 - \exp\left[-2h_y^E/\tilde{w}\right]/2h_y^E/\tilde{w}\right]^{1/2} = (h_y^s)^{p+1}. \tag{21}
\]

Note that for \(h_y^E/\tilde{w} \ll 1\), enforcing (21) to leading order becomes

\[
(h_y^E)^{p+1}\exp\left[-y^E/\tilde{w}\right] = (h_y^s)^{p+1} \Rightarrow h_y^E = h_y^s\exp[y^E\sqrt{\kappa_\parallel/\kappa_\perp}/(p + 1)], \tag{22}
\]

whereas enforcing (21) to leading order for \(h_y^E/\tilde{w} \gg 1\) becomes

\[
(h_y^E)^{p+1}\exp\left[-y^E/\tilde{w}\right][2h_y^E/\tilde{w}]^{-1/2} = (h_y^s)^{p+1} \Rightarrow

h_y^E = (h_y^s)^{p+1} \left[2\sqrt{\kappa_\parallel/\kappa_\perp}\right]^{p+1/2}\exp[y^E\sqrt{\kappa_\parallel/\kappa_\perp}/(p + 1/2)]. \tag{23}
\]

While the use of (21) directly or an approach that switches from (22) to (23) away from the layer would yield more efficiency, (22) is used throughout the mesh because it is easier to implement than (21) and is more stringent (i.e., has slower exponential growth) away from the layer than (23). The exponentially refined mesh \(\Omega_h\) that enforces (22) is constructed by first partitioning \(\Omega\) into \(m_x\) rectangles (i.e. \(h_x = \pi/m_x\) and \(h_y = 1\)), and then refining all rectangles \(E\) in the \(y\) direction into two smaller rectangles until

\[
h_y^E \leq h_y^s \exp[y^E\sqrt{\kappa_\parallel/\kappa_\perp}/(p + 1)] \quad \text{and} \quad h_y^E \leq h_y^s \exp[(1 - (y^E + h^E))\sqrt{\kappa_\parallel/\kappa_\perp}/(p + 1)] \tag{24}
\]

The computational efficiency, measured as the error norm in Section 4.1 versus number of degrees of freedom (dofs), is obtained for uniform, aspect ratio, and exponential refined meshes with \(\kappa_\parallel/\kappa_\perp = 10^4\). Both the aspect ratio and exponential refined meshes use \(m_x =\)
Figure 5: Efficiency diagram for threshold refinement using “true” flux error (reference solution) compared to using the ZZ estimator. The results show that the efficiency of AMR with threshold refinement using the ZZ estimator is similar to that using the reference flux error, and thus, the ZZ estimator is a sufficient error estimator for threshold refinement of our problem.

3, 6, 12, 24, 48, 96 and \( h_y^s = 1/(10^2 \cdot m_{yz}/3) \), with the aspect ratio meshes then using \( h_y^E = h_y^s \) while the exponential refined meshes use (24). Note that one expects a theoretical power law dependence of error on the number of dofs \( N \). Consider that \( N \) is proportional to the number of elements in the domain, which is proportional to \( 1/(h_y^s)^2 \) and thus \( h_y^s \sim 1/\sqrt{N} \). Thus, one might expect the error in temperature to be proportional to \( (h_y^s)^p \) to \( 1/N^{(p+1)/2} \) using the approximation theory bound (19) with \( k = 0 \) and \( l = 1 \). Figure 4 shows the efficiency results for piecewise linear \( (p = 1) \) and piecewise cubic \( (p = 3) \) function spaces. Both the aspect ratio and exponential refined meshes obtain the respective theoretical power law with far fewer dofs than required by the uniform meshes. Generally speaking, the aspect ratio meshes can attain the same accuracy solution as the uniform meshes with around an order of magnitude less dofs. Furthermore, the exponential refined meshes can obtain a solution of a given accuracy using an order of magnitude less dofs than the aspect ratio meshes and two to three orders of magnitude less dofs than the uniform meshes.

4.3. Mesh refinement for poloidally varying magnetic field problems

Note that while the aspect ratio and exponential refined meshes are very efficient for a constant magnetic field that is aligned with the mesh, these approaches do not immediately generalize to situations where either the mesh is not aligned with the magnetic field or when the magnetic field is dynamic in time. In particular, the aspect ratio mesh requires the mesh is aligned the magnetic field. The exponential refinement approach is more general in that it requires that the magnetic field is known \textit{a priori} but not that the mesh is aligned to the field. The caveat is that \( y^E \) and \( 1 - (y^E + h^E) \) in (24) must be replaced with some other measurement of the influence of the separatrix. The most general approach is adaptive mesh refinement (AMR), as it can be used to focus the refinement where it is needed regardless of whether the magnetic field is known \textit{a priori}.

Recall from Section 3.2 that the ZZ refinement strategy combines a threshold strategy paired with the ZZ estimate of the flux error (18). To first evaluate the performance of the ZZ error estimator independent of the performance of the threshold strategy, a reference
solution is computed on a highly refined mesh for the single null problem (11) to provide the "true" flux error on each element to the element identification routine (in lieu of an analytic solution). The reference solutions for \( \kappa_\parallel / \kappa_\perp \) values of 10^2, 10^4, and 10^6 are computed on uniform meshes consisting of 250^2, 1000^2 and 2000^2 elements, respectively. Note that the reference mesh sizes are chosen so that (i) the reference mesh has elements that are smaller than the boundary layer and (ii) the results herein are qualitatively the same as when 125^2, 500^2, and 1000^2 are used. The final piece of implementation detail is that the mesh refinement is halted when any element on the refined mesh becomes as small as the uniform elements in the reference solution. Figure 5 compares the efficiency of the AMR approach using the ZZ error estimate to the efficiency of the AMR approach using the reference flux error. Note that the efficiency of the AMR using the ZZ error estimate is qualitatively similar, and sometimes even better, than that of the AMR using the reference error. Figure 6 shows that the final meshes (and \( T \) field) are qualitatively the same for \( \kappa_\parallel / \kappa_\perp = 10^4 \) (i.e., the same regions are highly-refined).

With the ZZ error estimator verified as a sufficient estimator for threshold AMR refinement, the efficiency of the approach is now measured against exponential and uniform refinement for the single null magnetic field (11), double null magnetic field (13), and magnetic island (15) problems. Two different error metrics are considered for measuring the efficiency. The first is the error in temperature, defined as \( \int_\Omega |T - T^*| dV \). The second is the error in flux, defined as before: \( \int_\Omega \| \kappa \nabla T - \kappa \nabla T^* \|_2 dV \). Note that the errors are measured using the reference solution as \( T^* \) and quadrature rules to compute integrals. Recall that one might expect the error in temperature to be proportional to \( h^{p+1} \sim 1/N^{(p+1)/2} \) using the approximation theory bound (19). While a similar prediction for the flux error from (19) is not as direct, one still expects the error to be proportional to \( h^{p} \sim 1/N^{p/2} \) following the same decrease in order of \( h \) using \( k = 1 \) in (19).

For the single null magnetic field problem (11), the exponential refinement strategy is defined by first marking all elements that contain the separatrix that is implicitly defined by
\[ A_z(x) = A_z(x_s), \text{ and then using the newly refined element size } h^* \text{ to mark all other elements where} \]

\[ h^E > h^* \exp[|A_z(x_i) - A_z(x_s)|\sqrt{\kappa_{||}/\kappa_{\perp}/(p + 1)}] \quad (25) \]

with \( x_i \) being any of the quadrature points for the element \( E \). Figure 7 shows the ZZ and exponential refinement solutions with \( \kappa_{||}/\kappa_{\perp} \) values of \( 10^2 \), \( 10^4 \), and \( 10^6 \). Note that both solutions show temperature increasing proportional to the increase in \( \kappa_{||}/\kappa_{\perp} \) and show refinement in the vicinity of the separatrix, although the vicinity is typically larger with exponential refinement. Another key difference is that the ZZ refinement approach refines more than the exponential refinement approach in the interior of the separatrix where the source drives the maximum temperature. Note this is consistent with the (25) being independent of the source in (11).

Figure 8 investigates the efficiency of the ZZ and exponential refinement approaches, as well as that of uniform refinement for comparison. The efficiency results show a clear advantage for the ZZ refinement approach over that of uniform refinement, characterized by the ability to attain the theoretical power law dependence of error on dofs (\( 1/N^2 \) and \( 1/N^{3/2} \), respectively) with significantly fewer dofs. Specifically, the power law dependence is attained with at least one or two orders of magnitude less dofs for \( \kappa_{||}/\kappa_{\perp} \) values of \( 10^2 \) and \( 10^4 \), respectively. While neither solution attains the power law dependence for \( \kappa_{||}/\kappa_{\perp} = 10^6 \) with the allotted number of dofs, the results indicate even more advantage at the higher \( \kappa_{||}/\kappa_{\perp} \) ratio for the ZZ refinement solution. For the exponential refinement approach, a similar advantage in efficiency over uniform refinement is seen for anisotropy ratios large
enough to limit the vicinity of the separatrix that is refined (i.e., $\kappa/\kappa_\perp$ values of $10^4$ and $10^6$). The efficiency advantage becomes degraded, however, once the dominant source of error becomes resolving the high temperatures inside the separatrix, which is something the continued refinement in the vicinity of the separatrix curve does not address. While the point at which the advantage is degraded likely depends on the choice of $|A_z(x_i) - A_z(x_s)|$ in (25), there will still be a degradation point due to the independence of (25) on the source in the single null magnetic field problem (11).

For the double null magnetic field problem (13), the exponential refinement strategy is defined by again first marking all elements containing the separatrix defined by $A_z(x) = A_z(x_s)$, and then using the newly refined element size $h^E$ to mark all other elements where

$$h^E > h^s \exp \left[ \frac{A_z(x_i) - A_z(x_s)}{A_z(x_0) - A_z(x_s)} \sqrt{\frac{\kappa/\kappa_\perp}{p+1}} \right]$$

(26)

with $x_i$ being any of the quadrature points for the element $E$. Figure [9] shows the ZZ and exponential refinement solutions for the double null magnetic field problem (13) with $\kappa/\kappa_\perp$ values of $10^2$, $10^4$, and $10^6$. Like the single null magnetic field problem, both solutions show the temperature increasing proportional to $\kappa/\kappa_\perp$ and refinement in the vicinity of the separatrix. Also like the single null magnetic field problem, the ZZ refinement approach refines the interior of the separatrix where the source is driving the temperature whereas the exponential refinement approach is focused only on the vicinity of separatrix. While one might adjust the vicinity by choosing a different scaling and/or dependence of (26) on $A_z(x_i)$, the independence of (26) on the source or boundary in the double null magnetic field problem (13) means the approach can miss solution features that require additional resolution.

Figure [10] shows the corresponding efficiency results for the double null magnetic field problem. For $\kappa/\kappa_\perp = 10^2$, there is no significant advantage for the ZZ or exponential refinement approaches over the uniform refinement approaches as all approaches attain the theoretical power law dependence around the same small number of dofs. Note that the
ZZ refinement solution for $\kappa_\parallel/\kappa_\perp = 10^2$ in Figure 9 looks close to a uniform mesh, which is consistent with the ZZ refinement efficiency more or less matching that of uniform refinement. Note also that the exponential refinement approach again sees degraded efficiency once the dominant error comes from a region in the domain that is not included in the vicinity of the separatrix. For $\kappa_\parallel/\kappa_\perp = 10^4$, the ZZ and uniform refinement approaches both attain a power law dependence that is stronger than the theoretical predictions for both error metrics, with the power of $1/N$ exceeding $2$ and $3/2$, respectively. That said, the ZZ refinement approach does attain the same power law dependence shown by the uniform refinement approach using only around half the dofs. The exponential refinement approach does not exhibit a sustained power law dependence and is the least efficient of the three approaches, likely due to not sufficiently refining the regions away from the separatrix. As with the single null problem, no approach clearly attains the power law dependence for $\kappa_\parallel/\kappa_\perp = 10^6$ with the allotted number of dofs, although the results indicate a strong advantage of the ZZ refinement approach at the higher $\kappa_\parallel/\kappa_\perp$ ratio.

For the magnetic island problem \cite{15}, the exponential refinement strategy is the same as for the double null magnetic field problem, i.e., using \cite{26}. Figure 11 shows the ZZ and exponential refinement solutions for $\kappa_\parallel/\kappa_\perp$ values of $10^2$, $10^4$, and $10^6$. Unlike the single null and double null magnetic field problems, the magnetic island problem does not have a source inside the separatrix. As such, higher values of $\kappa_\parallel/\kappa_\perp$ do not result in higher temperatures but, instead, in a more constant temperature problem inside the separatrix. Again, the key difference between the ZZ and exponential refinement approaches is that the former focuses refinement both along the separatrix and in regions of large temperature gradients.
Figure 10: Efficiency diagrams for the double null magnetic field problem with various $\kappa_\parallel/\kappa_\perp \in \{10^2, 10^4, 10^6\}$, using $L_1$ temperature error (left) and sum of element flux errors defined in (18) (right) compared to theoretical power law (dotted black lines) and empirical power law (dotted blue lines) when theoretical power law is exceeded. Results show that ZZ refinement starts to require substantially fewer dofs to attain the power law exhibited by uniform refinement once $\kappa_\parallel/\kappa_\perp$ is large enough, indicating that ZZ refinement is more efficient for large $\kappa_\parallel/\kappa_\perp$.

whereas the latter focuses on the separatrix due to (26) being unable to account for source or boundary effects.

Figure 11 shows the corresponding efficiency results for the magnetic island problem. Like the double null magnetic field problem, the ZZ and uniform refinement approaches are about as efficient as the other for $\kappa_\parallel/\kappa_\perp$, due to ZZ refinement being essentially uniform (see Figure 11), with the ZZ refinement approach gaining a substantial advantage for larger $\kappa_\parallel/\kappa_\perp$. As with all the poloidally varying magnetic field problems, the advantage can again be explained as the ZZ refinement approach requiring substantially less dofs to obtain the same power law dependence as the uniform refinement approach eventually does. The exponential refinement approach again shows a degradation in efficiency, likely due to the lack of refinement outside the separatrix where there is a large temperature gradient. As with all the poloidally varying magnetic field problems, the exponential refinement lacks a consistent efficiency advantage due to its inability to refine solution features that result from sources and/or boundary conditions.

5. Discussion of Computational Efficiency

Section 4 demonstrated how adaptive mesh refinement can substantially improve the computational efficiency in solving anisotropic diffusion problems, such as the linear steady-state problem (4). Those results, however, were limited to a handful of idealized poloidally varying magnetic field problems. As such, some discussion is warranted on how much improvement in computation efficiency one can expect for more general magnetic fields, particularly with large anisotropy ratios. First, theoretical scaling laws for the degrees of freedom requirements with mesh refinement and general magnetic fields are presented. Then, the dependence of the matrix condition number on the anisotropy ratio is discussed with a focus on the implications on the cost of solving the underlying discrete anisotropic diffusion problem. The
Figure 11: ZZ refinement (top row) and exponential refinement (bottom row) solutions for the magnetic island problem with various $\kappa_\parallel/\kappa_\perp$ values: $10^2$ (left), $10^4$ (center), $10^6$ (right).

Figure 12: Efficiency diagrams for the magnetic island problem with $\kappa_\parallel/\kappa_\perp \in \{10^2, 10^4, 10^6\}$, using $L_1$ temperature error (left) and flux error (right) compared to theoretical power law (dotted black lines) and empirical power law (dotted blue lines) once it exceeds the theoretical power law. Results show that, once $\kappa_\parallel/\kappa_\perp$ is large enough, ZZ refinement requires substantially fewer dofs to attain the power law exhibited by uniform refinement, indicating that ZZ refinement is more efficient for large $\kappa_\parallel/\kappa_\perp$. 
resulting scaling laws are then combined to get an approximation of the total cost $C$ under uniform, isotropic, and anisotropic adaptive mesh refinement.

5.1. Scaling of dofs for constant magnetic field

To develop mesh refinement scaling laws for general magnetic fields, first consider the constant magnetic field $\hat{b} = \hat{x}$ that causes a boundary layer at $y = 0$ with width $w$. MFEM supports a variety of mesh refinement approaches, including to uniformly refine a quadrilateral/hexahedral element (i) isotropically into four/eight quadrilateral/hexahedral elements or (ii) anisotropically in two/four quadrilateral/hexahedral elements.

Consider applying either of these uniform refinement approach to a single element mesh, where the elements containing the boundary layer are refined along the $y$ direction. If the refinement process is repeated $J$ times, the resulting mesh will have nodes at $y = 0$ and at $y_j = L/2^j$ for $j = 0, \ldots, J$. Choose $J$ to be the smallest value such that the boundary layer is resolved by at least one element, or equivalently that $y_J \leq w$ and $y_{J-1} > w$. Note that $J$ is the smallest integer that is greater than $\log(L/w)/\log(2)$. The quantity of interest is how many elements $N$ result from $J \approx \log(L/w)/\log(2)$ iterations of refinement, which can be written as the sum, $N = N_1 + \ldots + N_J$, where $N_j$ is the number of elements from the $j$-th round of refinement (i.e., elements that have nodes at $y_j$ and $y_{j-1}$).

First, consider anisotropic refinement in two dimensions (2D), i.e., where a quadrilateral element is uniformly refined into two smaller quadrilaterals. After the first refinement, $j = 1$, the mesh now has two elements, one of which contains the boundary and one which does not contain the boundary layer. Note that the element that does not contain the boundary layer will not be further refined: $N_1 = 1$. The element that contains the boundary layer is refined into two elements: one that contains the boundary layer and one that does not. Note that the new element that does not contain the boundary layer will not be further refined: $N_2 = 1$. Generalizing, one can see that $N_j = 1$ for $j < J$. At the last round of refinement, $j = J$, all the new elements are no longer refined: $N_J = 2$. Thus, the number of elements for the two-dimensional isotropic refinement mesh is

$$N_{\text{aniso-amr}} = J + 1 \approx \frac{\log(L/w)}{\log(2)} + 1 \sim \log \sqrt{\frac{\kappa_\parallel}{\kappa_\perp}}$$

for $\kappa_\parallel/\kappa_\perp \gg 1$. Note that this result also applies to three dimensions (3D).

Now, consider isotropic refinement in two dimensions (2D), i.e., where a quadrilateral element is uniformly refined into four smaller quadrilaterals. After the first refinement level, $j = 1$, the mesh now has four elements, two of which contain the boundary and two which do not contain the boundary layer. Note that the two elements that do not contain the boundary layer will not be further refined: $N_1 = 2$. The two elements that contain the boundary layer are each refined into four elements: two that contain the boundary layer and two that do not. Note that the new elements that do not contain the boundary layer will not be further refined: $N_2 = 2 \cdot N_1 = 4$. The four elements that contain the boundary layer are each refined into four elements: two that contain the boundary layer and two that do not. Again note that the new elements that do not contain the boundary layer will not be further refined: $N_3 = 2 \cdot N_2$. Generalizing, one can see that $N_j = 2 \cdot N_{j-1}$ for $j < J$. At the final round of refinement, $j = J$, all the elements are no longer refined: $N_J = 4 \cdot N_{J-1}$.
Thus, the number of elements for the two-dimensional isotropic refinement mesh is

\[ N_{iso-amr}^{2D} = \sum_{j=1}^{J} 2^j + 2^J = 3 \cdot 2^J - 2 \approx 3(L/w) - 2 \sim \sqrt{\kappa_\parallel / \kappa_\perp} \]  

(28)

for \( \kappa_\parallel / \kappa_\perp \gg 1 \). Note that in three dimensions, where each hexahedron is refined into eight smaller hexahedrons, one has \( N_1 = 4, N_j = 4 \cdot N_{j-1} \) for \( j < J \), and \( N_J = 8 \cdot N_{J-1} \). Thus, the number of elements for the three-dimensional isotropic refinement mesh is

\[ N_{iso-amr}^{3D} = \sum_{j=1}^{J} 4^j + 4^J = \frac{7}{3} (2^J)^2 - \frac{4}{3} \approx \frac{7}{3} (L/w)^2 - \frac{4}{3} \sim \kappa_\parallel / \kappa_\perp. \]  

(29)

Thus, one can clearly see the advantage of using anisotropic refinement with field aligned meshes for large \( \kappa_\parallel / \kappa_\perp \), especially in three-dimensional simulations.

5.2. Scaling of dofs for general geometry

The expression for the interpolation error (19) implies that in order to restrict the maximum error, it is beneficial to exponentially pack the mesh near the boundary layer. For a simple Cartesian mesh, the mesh points would lie at the point \( y_j = L e^{-j/n_r} \), where \( L \) is the overall mesh size and the parameter \( n_r \) controls the rate at which cells are refined. For field aligned meshes, the previous section found, \( n_r = 1 / \log(2) \), but this rate could potentially differ for more general geometric situations. The number of elements in the radial direction required to reach the boundary layer width \( w \) is

\[ J = n_r \log (L/w) \sim n_r \log \sqrt{\kappa_\parallel / \kappa_\perp}. \]  

(30)

The lesson learned from our analysis is that, for isotropic refinement one also needs to pack at the smallest radial distance \( w \) in the directions tangent to flux surfaces. In this case, refining a factor of 2 at each level yields \( 2^{d-1} \) more elements per level for a total of

\[ \sim 2^J (d-1) = (L/w)^{(d-1)} \]  

elements. This leads to the estimate

\[ N_{iso-amr} \sim (L/w)^{d-1} \sim (\sqrt{\kappa_\parallel / \kappa_\perp})^{(d-1)} \]

\[ N_{iso-amr}/N_{uniform} \sim w/L \sim (\kappa_\perp / \kappa_\parallel)^{1/2} \]  

(31)

which generalizes the results of the previous section.

In principle, anisotropic refinement can perform much better than isotropic, if the elements can become field aligned at the smallest scales, because one does not need to resolve such small length scales tangent to the boundary layer. Assume that this still requires \( n_s = L/h_s \) points per dimension tangent to boundary layer, where \( h_s \) is the distance needed to resolve the surface. This yields the estimate

\[ N_{aniso-amr} \sim (L/h_s)^{d-1} n_r \log(L/w) \sim n_r h_s^{d-1} \log \sqrt{\kappa_\parallel / \kappa_\perp} \]

\[ N_{aniso-amr}/N_{uniform} \sim (w^d / h_s^{d-1} L) n_r \log(L/w) \sim n_r h_s^{2d-1} (\kappa_\perp / \kappa_\parallel)^{d/2} \log \sqrt{\kappa_\parallel / \kappa_\perp}. \]  

(32)
In 2D, \( n_p = L/h_p \) where \( h_p \) is the smallest distance in poloidal direction. This is set by the needs of accuracy and is related to the poloidal curvature of the surface, \( a \), rather than the boundary layer width. Thus, the estimate is simply

\[
N_{\text{aniso-amr}} \sim (L/h_p)n_r \log(L/w) \sim n_r n_p \log \sqrt{\kappa_\parallel/\kappa_\perp}.
\]

(33)

In 3D, the number of points required in the toroidal direction is \( n_t = L/h_t \) where \( h_t \) is the smallest distance required in the toroidal direction. Again, this is set by the need to resolve the toroidal curvature, \( R \), rather than by the boundary layer width. Thus, the result is

\[
N_{\text{aniso-amr}} \sim (L^2/h_p h_t)n_r \log(L/w) \sim n_r n_p n_t \log \sqrt{\kappa_\parallel/\kappa_\perp}.
\]

(34)

To compare against the theoretical scaling laws, Figure 13 shows the number of dofs, which are proportional to the number of elements, for the single null problem. The minimum element size \( h_{\text{min}} \) is noted as the meshes for \( \kappa_\parallel/\kappa_\perp \in \{10^2, 10^4, 10^6\} \) are refined according to the uniform and ZZ refinement strategies. Note that when the number of dofs at each iteration of refinement are compared to \( 1/h_{\text{min}} \), the uniform and ZZ refinement follow power laws of \((\kappa_\parallel/\kappa_\perp)^2\) and \((\kappa_\parallel/\kappa_\perp)\), respectively. The number of dofs for both uniform and ZZ refinement once \( h_{\text{min}} \) becomes smaller than \( \sqrt{\kappa_\perp/\kappa_\parallel} \) is also noted for each value of \( \kappa_\parallel/\kappa_\perp \in \{10^2, 10^4\} \). Note that the ratio of those dofs scales as \( \sqrt{\kappa_\parallel/\kappa_\perp} \), as predicted by (33). Specifically, the ZZ refinement mesh resolves the layer with a factor of almost 2 fewer dofs for \( \kappa_\parallel/\kappa_\perp = 10^2 \) and a factor of almost 20 fewer dofs for \( \kappa_\parallel/\kappa_\perp = 10^4 \). Extrapolating these results to an anisotropy ratio of \( 10^6 \), one would predict a reduction in dofs by a factor of around 200.

5.3. Scaling of dofs for filamentary structures

Although this work is focused on steady-state transport, in the future it would be useful to extend the results to the consideration of dynamic plasma turbulence and MHD activity.
It is well known that it is advantageous for the instabilities responsible for driving plasma turbulence and MHD preferentially form elongated filamentary structures along field lines. This is perhaps best exemplified by edge localized modes that are active in the pedestal of a tokamak operating in the high-performance (H-mode) confinement regime.

Isotropic AMR is an excellent strategy for capturing low-dimensional structures in higher dimensional volumes. In order to resolve a small region with perpendicular width \( w \) at the corner of a \( d \)-dimensional hypercube, one requires \( J \) iterations with \( N \sim 2^d \) elements for each iteration for a total number of elements \( \sim 2^d J \). If the point is in the interior of the domain, this requires \( \sim 2^d J \sim 2^d \log (L/w) \) elements.

For a filamentary structure, there is only one direction along the filament. Thus, isotropic AMR in the perpendicular directions is a good strategy for resolving the core of the filament. For an anisotropic AMR procedure that can align with the filament at the smallest scales, then one can assume that the number of elements required along the field line is \( n_\parallel = L/h_\parallel \), where \( h_\parallel \) is determined by the parallel field line curvature. This yields

\[
N_{\text{aniso-amr}} \sim (L/h_\parallel) 2^{(d-1)} \log(L/w) \sim n_\parallel 2^{(d-1)} \log \sqrt{\kappa_\parallel/\kappa_\perp}.
\]

For isotropic AMR, the direction along the filament must also be resolved to width \( w \). Each iteration layer has \( \sim 2^{(d-1)} \) perpendicular regions with \( \sim 3 \times 2^J \) elements required to cover the parallel direction, for a total of \( 3 \cdot 2^{(d-1)} (L/w) \) elements. Hence, the result is

\[
N_{\text{iso-amr}} \sim 3 \cdot 2^{(d-1)} (L/w) \sim 3 \cdot 2^{(d-1)} \sqrt{\kappa_\parallel/\kappa_\perp}.
\]

Thus, isotropic AMR performs relatively well in this scenario.

5.4. Solver Iterations

With an iterative approach to solving the underlying sparse linear system that represents the discrete steady state problem (17), the total cost of the solution (i.e., number of flops) is on the order of the product of the number of iterations required and the total number of degrees of freedom (dofs). Thus, while the number of dofs is a significant factor in the cost of solving the problem, the dependence of the number of required iterations on the condition number of the matrix in (17) is another important factor. For example, the convergence rate of the conjugate gradient method depends on the reciprocal of the condition number. To discuss the matrix condition number in (17), consider that the eigenfunction problem for the constant magnetic field problem (8) has solutions of \( \exp[i2q_x x] \exp[i2\pi q_y y] \), for non-zero integers \( q_x \) and \( q_y \), with corresponding eigenvalues \( \lambda_{q_x,q_y} = \kappa_\parallel (2q_x)^2 + \kappa_\perp (2\pi q_y)^2 \). While the eigenfunctions and eigenvalues of the finite-dimensional operator in (17) will not be the same as those for the operator in (8), the difference between the two should be proportional to the mesh spacing. Thus, let \( q_{x,h} \) and \( q_{y,h} \) represent the largest frequencies supported by \( \Omega_h \) so that the matrix condition number in (17) is approximated by

\[
\frac{\max_{q_x,q_y} |\lambda_{q_x,q_y}|}{\min_{q_x,q_y} |\lambda_{q_x,q_y}|} = \frac{\kappa_\parallel (2q_{x,h})^2 + \kappa_\perp (2\pi q_{y,h})^2}{\kappa_\parallel^2 + \kappa_\perp (2\pi)^2} = \frac{(\kappa_\parallel/\kappa_\perp)(2q_{x,h})^2 + (2\pi q_{y,h})^2}{(\kappa_\parallel/\kappa_\perp)^2 + (2\pi)^2}.
\]
Consider when $\Omega_h$ consists of $m \times m$ uniform elements, resulting in $q_{x,h} = q_{y,h}$. Denote that largest supported frequency by $q_h$ and note that (37) is now

$$\frac{\max_{q_x,q_y} |\lambda_{q_x,q_y}|}{\min_{q_x,q_y} |\lambda_{q_x,q_y}|} = q_h^2$$

(38)

Recall that Section 4 found that the boundary layer of width $\sqrt{\kappa_\perp/\kappa_\parallel}$ must be resolved by the elements in $\Omega_h$ to reach the asymptotic regime, which would result in $q_h = \sqrt{\kappa_\parallel/\kappa_\perp}$. Therefore, it is important to note that as larger values of $\kappa_\parallel/\kappa_\perp$ are considered, the number of conjugate gradient iterations required to obtain the solution on a uniform mesh is expected to grow proportional to $\sqrt{\kappa_\parallel/\kappa_\perp}$ (i.e., the square-root of the reciprocal of the matrix condition number). To test this $\sqrt{\kappa_\parallel/\kappa_\perp}$ scaling, the iteration counts for conjugate gradient preconditioned with BoomerAMG are collected for (8). Figure 14 shows that the iteration counts do indeed exhibit a power law dependence on anisotropy ratio bounded above by the $\sqrt{\kappa_\parallel/\kappa_\perp}$ scaling law for both linear ($p = 1$) and cubic ($p = 3$) function spaces.

When the mesh is aligned with the magnetic field, the ratio and exponential refinement approaches investigated in Section 4 can potentially avoid iteration count growth with anisotropy ratio. Consider a mesh where $m_x = \alpha m_y$ and $m_y = \sqrt{\kappa_\parallel/\kappa_\perp}$. The largest supported frequencies on $\Omega_h$ are now $q_{x,h} = \alpha \sqrt{\kappa_\parallel/\kappa_\perp}$ and $q_{y,h} = \sqrt{\kappa_\parallel/\kappa_\perp}$. Thus, the matrix condition number is now

$$\frac{\max_{q_x,q_y} |\lambda_{q_x,q_y}|}{\min_{q_x,q_y} |\lambda_{q_x,q_y}|} = \frac{\alpha^2 (\kappa_\parallel/\kappa_\perp)^2 2^2 + (\kappa_\parallel/\kappa_\perp)(2\pi)^2}{(\kappa_\parallel/\kappa_\perp)^2 + (2\pi)^2},$$

(39)

which is $O(1)$ if $\alpha \sim \sqrt{\kappa_\perp/\kappa_\parallel}$ for $\kappa_\parallel/\kappa_\perp \gg 1$. One would then expect the iteration count on such a ratio refinement mesh, including the $m_y/m_x = 100/3$ mesh for $\kappa_\parallel/\kappa_\perp = 10^4$, to be almost independent of growth in anisotropy ratio. To test the independence, the iteration counts for conjugate gradient preconditioned with BoomerAMG are again collected for (8). Figure 15 shows that indeed the iteration counts for both linear ($p = 1$) and cubic ($p = 3$) function spaces do indeed appear to be independent of $\kappa_\parallel/\kappa_\perp$. 

26
The condition number estimates in (38) and (39) can be generalized beyond field aligned meshes for constant magnetic fields. Generally speaking, the condition number of the discrete parallel and perpendicular Laplacian operators scale as \(\kappa_{\parallel}/\kappa_{\perp}\), respectively, where \(d\) is the dimension, \((L_{\parallel}, h_{\parallel})\) are the longest and shortest wavelengths present in the parallel direction, and \((L_{\perp}, h_{\perp})\) are the longest and shortest wavelengths present in the perpendicular direction. Assume the mesh resolves the boundary layer with width \(w\) (i.e., \(h_{\perp} = w\)) and that the perpendicular directions interact with the domain boundary so that \(L_{\perp}\) is approximately the domain length (i.e., \(L_{\perp} \approx L\)). The condition number of the combined anisotropic Laplacian for such a problem is approximately

\[
\frac{\kappa_{\parallel}(1/h_{\parallel})^2 + (d-1)\kappa_{\perp}(1/w)^2}{\kappa_{\parallel}(1/L_{\parallel})^2 + (d-1)\kappa_{\perp}(1/L)^2}.
\]

Consider first problems where the parallel directions also interact with the boundary so that \(L_{\parallel} = L\). The condition number estimate (40) is now

\[
(L/w)^2 \frac{\kappa_{\parallel}(w/h_{\parallel})^2 + (d-1)\kappa_{\perp}}{\kappa_{\parallel} + (d-1)\kappa_{\perp}} = \frac{\kappa_{\parallel} \kappa_{\perp}(w/h_{\parallel})^2 + (d-1)\kappa_{\perp}}{\kappa_{\parallel} + (d-1)\kappa_{\perp}}.
\]

Note if a uniform mesh is used where \(h_{\parallel} \approx h_{\perp} = w\), the condition number scales approximately as \(\kappa_{\parallel}/\kappa_{\perp}\), which is consistent with (38). Note if a ratio refined mesh is used where \(h_{\parallel} \approx \sqrt{\kappa_{\parallel}/\kappa_{\perp}} w\), the condition number scales approximately independent of \(\kappa_{\parallel}/\kappa_{\perp}\), which is consistent with (39).

Consider now problems where the parallel directions do not interact with the boundary so that \(L_{\parallel} = \infty\). The condition number estimate (40) is now

\[
(L/w)^2 \frac{\kappa_{\parallel}(w/h_{\parallel})^2 + (d-1)\kappa_{\perp}}{(d-1)\kappa_{\perp}} = \frac{\kappa_{\parallel} \kappa_{\perp}(w/h_{\parallel})^2 + (d-1)\kappa_{\perp}}{(d-1)\kappa_{\perp}}.
\]

Note that if a uniform mesh is now used where \(h_{\parallel} \approx w\), the condition number scales approximately as \((\kappa_{\parallel}/\kappa_{\perp})^2\). Thankfully, it is observed that the single null (11), double-null (12),
Figure 16: Dependence of iteration count on anisotropy ratio $\kappa_\parallel/\kappa_\perp$ for uniform refinement meshes when ILU preconditioning is used for the constant magnetic field problem \cite{Anon}. Results show a substantial decrease in iteration count over BoomerAMG preconditioning, with the counts actually decreasing with larger anisotropy ratio.

and magnetic island \cite{Anon} problems all exhibit iteration counts that scale as $\sqrt{\kappa_\parallel/\kappa_\perp}$. This behavior is seen for both uniform refinement and ZZ refinement meshes, which is welcome news as even though ratio refinement theoretically can restore the condition number scaling to $\kappa_\parallel/\kappa_\perp$ when the parallel directions do not interact with the boundary, a ratio refined mesh approach for poloidally varying magnetic fields may be quite nontrivial. Even once one has the field-aligned mesh, the dofs themselves may not be perfectly field-aligned even in an element with edges aligned.

A successful solver approach for poloidally varying magnetic fields will therefore likely require both meshing and preconditioning strategies. As an example, ILU preconditioning results in iteration counts that do not grow with increasing magnetic anisotropy even on uniform meshes. In certain cases, including when applied to the constant magnetic field problem \cite{Anon}, the iteration count actually decreases with increasing magnetic anisotropy (see Figure 16). The downsides are that ILU preconditioning is not generally scalable across multiple MPI ranks and that the same benefit did not appear to extend to non field-aligned meshes in our initial testing.

\subsection*{5.5. Computational Cost}
To summarize, one can estimate the total cost to solution by multiplying the scaling laws for dofs by the empirically measured scaling law for iteration count. For uniformly refined meshes, the cost would be

$$C_{\text{uniform}} \sim (L/w)^{d+1} \sim (\sqrt{\kappa_\parallel/\kappa_\perp})^{d+1}. \quad (41)$$

For isotropic adaptive refined meshes, the cost is improved by the factor $L/w \sim \sqrt{\kappa_\parallel/\kappa_\perp}$

$$C_{\text{iso-amr}} \sim (L/w)^d \sim (\sqrt{\kappa_\parallel/\kappa_\perp})^d. \quad (42)$$

For anisotropic adaptive refined meshes, the cost could potentially be as low as

$$C_{\text{aniso-amr}} \sim (L/w) \log(L/w) \sim \sqrt{\kappa_\parallel/\kappa_\perp} \log \sqrt{\kappa_\parallel/\kappa_\perp}. \quad (43)$$
6. Conclusion

This work investigated the efficiency of various mesh refinement strategies for boundary layers formed by the extreme magnetic anisotropy found in tokamak reactors. While verifying the implementation of a finite element discretization in MFEM against an analytic solution for a test problem, we noted the theoretical convergence rate is not attained for any polynomial order until the mesh elements containing the boundary layer are small enough to capture the layer (see Figure 3). We found such behavior consistent with established approximation theory that indicates the error in the polynomial finite element solution will depend on the polynomial order and mesh size in such a way that the mesh size must be small enough in the layer to counter the increasing magnitude of higher-order derivatives of the exact solution. The same approximation theory result indicates that substantially less mesh refinement is needed in the direction along the layer relative to across the layer. Furthermore, the approximation theory result indicates less mesh refinement is needed for elements far away from the layer than for elements within the layer. As such, we utilized two variable mesh strategies for meshes that are aligned with the magnetic field: one that used uniform rectangular elements with an aspect ratio of the layer width, and a second that chooses the aspect ratio of the rectangular elements to capture the exponential nature of the approximation theory bound. We found that the uniform aspect ratio mesh attains the same accuracy as the uniform rectangular mesh using an order of magnitude less degrees of freedom (dofs) and that the exponential refined rectangular mesh attains another order of magnitude reduction in the number of dofs needed (see Figure 4).

The variable mesh refinement approaches do not directly generalize to magnetic fields that do not align with the mesh, such as the spatially (and temporally) varying fields in a tokamak reactor. As an example, the exponential refinement approach in Figure 4 can be used for magnetic fields with a given potential function after the $y_E$ input is replaced with a function of the potential value at $y_E$. Note that the use of such a function based on the magnetic flux function is the current standard refinement approach in the field. For a more general refinement approach, we explored an adaptive mesh refinement (AMR) approach that refines elements that exhibit an estimated error larger than a specified threshold, with the estimated error obtained by the Zienkiewicz and Zhu approach. After empirically verifying that the Zienkiewicz and Zhu error estimation leads to similar refinement behavior as error estimation using a highly refined reference solution, we measured the efficiency of the AMR and exponential refinement approaches against uniform refinement for three magnetic fields that model the spatial variation of poloidal fields expected in a tokamak. For the single-null tokamak magnetic field geometry, we found the AMR approach attained the same accuracy as the uniform mesh using at least an order of magnitude less dofs (see Figure 8), with the efficiency gain increasing with stronger anisotropy. The exponential refinement approach also sees increased computational efficiency for higher anisotropy ratios while solution error along the separatrix remains dominant. Once the solution error resolving the interior of the separatrix becomes dominant, the solution accuracy stagnates resulting in the loss of the efficiency advantage. For the double null magnetic field that models the poloidal field in a double-null tokamak reactor, we found the AMR approach had at least an order of magnitude gain in efficiency once the anisotropy was strong enough whereas the exponential refinement approach did not exhibit a substantial improvement in efficiency (see Figure 10).
The magnetic island field that models a perturbation to the poloidal field showed similar results as with the double null magnetic field: significant efficiency gains for the AMR approach and no significant efficiency gain for the exponential refinement approach (see Figure 12). For all three fields, the efficiency advantage for the AMR approach is characterized by the AMR solution requiring less dofs than either the uniform or exponential solutions before attaining the power law dependence of error on dofs exhibited by the uniform solution. Given that the AMR approach is both more efficient for these fields at higher anisotropy ratios than the exponential strategy and does not require locating the separatrices and defining a distance function based on the magnetic flux function, we recommend the AMR approach be considered instead of the exponential strategy when simulating more complex magnetic geometries.

While we explored anisotropy ratios up to $10^6$, the anisotropy in a tokamak reactor is expected to be much stronger. We found that extending our experiments to anisotropy ratios beyond $10^6$ quickly becomes infeasible because of the increasing number of dofs required to resolve the layer compounded by the number of iterations required by the conjugate gradient solver used with algebraic multigrid preconditioning. We derived scaling laws that showed the number of dofs for isotropic adaptive mesh refinement is less than uniform refinement by a factor of the square-root of the anisotropy ratio (31), whereas the number of dofs for anisotropic adaptive mesh refinement is even smaller with only a logarithmic dependence on the square-root of the anisotropy ratio for both two-dimensional (33) and three-dimensional (34) problems. For the field-aligned test problem used to verify the finite element discretization, we derived a theoretical estimate that the condition number associated with uniform refinement grows as the anisotropy ratio (38) and empirically verified the resulting prediction that the conjugate gradient iteration count, with algebraic multigrid preconditioner, grows as the square-root of the anisotropy ratio result (see Figure 14). We similarly derived a theoretical estimate that the condition number associated with ratio refinement, where the rectangular elements have an aspect ratio equal to the layer width, is independent of the anisotropy ratio (39) and empirically verified that the iteration count is independent of anisotropy ratio (see Figure 15). Such independence of iteration count and anisotropy ratio is crucial to generalize beyond ratio refined meshes, such as those produced by AMR, that can resolve spatially and temporally varying magnetic fields. We note that ILU preconditioning does indeed result in independence of and substantial reduction in iteration count on a uniform mesh for the field-aligned test problem (see Figure 16); however, ILU preconditioning does not readily scale across multiple MPI ranks to tackle the larger number of dofs required by more realistic problems. Thus, in the future, it would be desirable to explore other promising solver techniques, e.g. that use line smoothing or geometric multigrid along with the direct solution of the resulting coarsened problems.

Acknowledgements

The authors thank Ben Zhu, Ben Dudson, Will Pazner, and Ben Southworth for many thoughtful discussions. 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-840380. This work was supported by LLNL Laboratory Directed Research and Development project PLS-20-ERD-038.
References

[1] J. Wesson, D. J. Campbell, Tokamaks, volume 149, Oxford university press, 2011.

[2] A. H. Boozer, Physics of magnetically confined plasmas, Reviews of modern physics 76 (2005) 1071.

[3] T. D. Rognlien, D. D. Ryutov, N. Mattor, G. D. Porter, Two-dimensional electric fields and drifts near the magnetic separatrix in divertor tokamaks, Physics of Plasmas 6 (1999) 1851–1857.

[4] S. Wiesen, D. Reiter, V. Kotov, M. Baelmans, W. Dekeyser, A. S. Kukushkin, S. W. Lisgo, R. A. Pitts, V. Rozhansky, G. Saibene, et al., The new solps-iter code package, Journal of Nuclear Materials 463 (2015) 480–484.

[5] C. R. Sovinec, T. A. Gianakon, E. D. Held, S. E. Kruger, D. D. Schnack, N. Team, NIMROD: A computational laboratory for studying nonlinear fusion magnetohydrodynamics, Physics of Plasmas 10 (2003) 1727–1732.

[6] C. R. Sovinec, A. H. Glasser, T. A. Gianakon, D. C. Barnes, R. A. Nebel, S. E. Kruger, D. D. Schnack, S. J. Plimpton, A. Tarditi, M. S. Chu, et al., Nonlinear magnetohydrodynamics simulation using high-order finite elements, Journal of Computational Physics 195 (2004) 355–386.

[7] N. M. Ferraro, S. C. Jardin, Calculations of two-fluid magnetohydrodynamic axisymmetric steady-states, Journal of Computational Physics 228 (2009) 7742–7770. URL: https://www.sciencedirect.com/science/article/pii/S0021999109003969. doi:https://doi.org/10.1016/j.jcp.2009.07.015.

[8] M. Hoelzl, G. T. A. Huijsmans, S. J. P. Pamela, M. Bécoulet, E. Nardon, F. J. Artola, B. Nkonga, C. V. Atanasiu, V. Bandaru, A. Bhole, et al., The JOREK non-linear extended MHD code and applications to large-scale instabilities and their control in magnetically confined fusion plasmas, Nuclear Fusion 61 (2021) 065001.

[9] S. I. Braginskii, Transport processes in a plasma, edited by MA Leontovich, Reviews of Plasma Physics 1 (1963) 205.

[10] L. Spitzer Jr., R. Härm, Transport phenomena in a completely ionized gas, Physical Review 89 (1953) 977.

[11] R. D. Hazeltine, J. D. Meiss, Plasma confinement, Courier Corporation, 2003.

[12] W. D. D’haeseleer, W. N. G. Hitchon, J. D. Callen, J. L. Shohet, Flux Coordinates and Magnetic Field Structure, Springer Berlin, Heidelberg, 2012. doi:10.1007/978-3-642-75595-8.

[13] A. S. Richardson, NRL Plasma Formulary, 2019.

[14] R. D. Hazeltine, J. D. Meiss, Plasma confinement, Dover Publications, Inc., Mineola, New York, 2003.
[15] P. Helander, D. J. Sigmar, Collisional transport in magnetized plasmas, Cambridge University Press, New York, 2005.

[16] R. Anderson, J. Andrej, A. Barker, J. Bramwell, J.-S. Camier, J. Cerveny, V. Dobrev, Y. Dudouit, A. Fisher, T. Kolev, W. Pazner, M. Stowell, V. Tomov, I. Akkerman, J. Dahm, D. Medina, S. Zampini, MFEM: A Modular Finite Element Methods library, Computers & Mathematics with Applications 81 (2021) 42–74. doi:10.1016/j.camwa.2020.06.009

[17] R. D. Falgout, U. M. Yang, Hypre: A library of high performance preconditioners, in: P. M. A. Sloot, A. G. Hoekstra, C. J. K. Tan, J. J. Dongarra (Eds.), Computational Science — ICCS 2002, Springer Berlin Heidelberg, 2002, pp. 632–641.

[18] O. C. Zienkiewicz, J. Z. Zhu, The superconvergent patch recovery and a posteriori error estimates. Part 1: The recovery technique, Int. J. Numer. Meth. Engng. 33 (1992) 1331–1364. URL: https://doi.org/10.1002/nme.1620330702 doi:10.1002/nme.1620330702.

[19] O. C. Zienkiewicz, J. Z. Zhu, The superconvergent patch recovery and a posteriori error estimates. Part 2: Error estimates and adaptivity, Int. J. Numer. Meth. Engng. 33 (1992) 1365–1382. URL: https://doi.org/10.1002/nme.1620330703 doi:10.1002/nme.1620330703.

[20] O. C. Zienkiewicz, J. Z. Zhu, A simple error estimator and adaptive procedure for practical engineering analysis, Int. J. Numer. Meth. Engng. 24 (1987) 337–357. URL: https://doi.org/10.1002/nme.1620240206 doi:10.1002/nme.1620240206.

[21] P. G. Ciarlet, P. A. Raviart, General Lagrange and Hermite interpolation in $\mathbb{R}^n$ with applications to finite element methods, Archive for Rational Mechanics and Analysis 46 (1972) 177–199. URL: https://doi.org/10.1007/BF00252458 doi:10.1007/BF00252458.