SMAI-JCM
SMAI Journal of Computational Mathematics

An Accurate SUPG-stabilized Continuous Galerkin Discretization for Anisotropic Heat Flux in Magnetic Confinement Fusion

Golo A. Wimmer, Ben S. Southworth, Koki Sagiyama & Xian-Zhu Tang

Volume 11 (2025), p. 473-496.
https://doi.org/10.5802/smai-jcm.131

The SMAI Journal of Computational Mathematics is a member of the Centre Mersenne for Open Scientific Publishing
http://www.centre-mersenne.org/
Submissions at https://smai-jcm.centre-mersenne.org/ojs/submission
e-ISSN: 2426-8399

SMAI Journal of Computational Mathematics
Vol. 11, 473-496 (2025)

An Accurate SUPG-stabilized Continuous Galerkin Discretization for Anisotropic Heat Flux in Magnetic Confinement Fusion

Golo A. Wimmer [1]
Ben S. Southworth [2]
Koki Sagiyama [3]
Xian-Zhu Tang [4]

[1] Los Alamos National Laboratory
E-mail address: gwimmer@lanl.gov
[2] Los Alamos National Laboratory
[3] Imperial College London
[4] Los Alamos National Laboratory.

Abstract. We present a novel spatial discretization for the anisotropic heat conduction equation, aimed at improved accuracy at the high levels of anisotropy seen in a magnetized plasma, for example, for magnetic confinement fusion. The new discretization is based on a mixed formulation, introducing a form of the directional derivative along the magnetic field as an auxiliary variable and discretizing both the temperature and auxiliary fields in a continuous Galerkin (CG) space. Both the temperature and auxiliary variable equations are stabilized using the streamline upwind Petrov–Galerkin (SUPG) method, ensuring a better representation of the directional derivatives and therefore an overall more accurate solution. This approach can be seen as the CG-based version of our previous work (Wimmer, Southworth, Gregory, Tang, 2024), where we considered a mixed discontinuous Galerkin (DG) spatial discretization including DG-upwind stabilization. We prove consistency of the novel discretization, and demonstrate its improved accuracy over existing CG-based methods in test cases relevant to magnetic confinement fusion. This includes a long-run tokamak equilibrium sustainment scenario, demonstrating a 35% and 32% spurious heat loss for existing primal and mixed CG-based formulations versus 4% for our novel SUPG-stabilized discretization.

2020 Mathematics Subject Classification. 65M60, 80M10.

Keywords. Anisotropic heat conduction, anisotropic diffusion, auxiliary operator, continuous Galerkin, SUPG.

1. Introduction

Plasmas in magnetic confinement fusion exhibit extremely anisotropic heat fluxes [24], conducting heat at a rate up to 10 orders of magnitude greater parallel to magnetic field lines than perpendicular [18]. If the parallel heat flux is not represented accurately in numerical simulations, the resulting increased cross-diffusion perpendicular to magnetic field lines may lead to spuriously short energy confinement times. One way to ensure such an accurate representation is to align the computational mesh with magnetic flux surfaces [13, 23, 24]; however, this may not be possible in more complex magnetohydrodynamic (MHD) simulations including magnetic islands or even stochastic magnetic

This work was supported by the Laboratory Directed Research and Development program of Los Alamos National Laboratory, under project number 20240261ER, as well as the U.S. Department of Energy Office of Fusion Energy Sciences Base Theory Program, at Los Alamos National Laboratory under contract No. 89233218CNA000001. The computations have been performed using resources of the National Energy Research Scientific Computing Center (NERSC), a U.S. Department of Energy Office of Science User Facility operated under Contract No. DE-AC02-05CH11231. Further support was received by the Engineering and Physical Sciences Research Council (grant number EP/W029731/1). Los Alamos National Laboratory report number LA-UR-24-33208.
https://doi.org/10.5802/smai-jcm.131

field lines [14, 36]. In practice, such simulations therefore require numerical methods that accurately represent the parallel heat flux independently of mesh alignment.

Due to the importance of anisotropic heat flux in extended MHD and two-fluids simulations for magnetic fusion plasmas, a range of numerical methods have been developed in recent years in order to tackle this challenge. This includes, but is not limited to, improved accuracy through higher polynomial order discretizations [16, 17], mesh refinement at areas of large expected cross-diffusion errors [38], asymptotic preserving methods (and extensions thereof), aimed at well-posedness in the limit of infinite anisotropy strength [11, 12, 25, 32, 41, 42], and first-order hyperbolic system methods including auxiliary variables for the temperature's gradient components together with pseudo-time advancement terms [10] (and references therein). Another auxiliary variable-based approach is presented in [19], where the auxiliary variable represents the temperature's directional derivative along magnetic field lines. The same idea is expanded on in our previous work [39], where a mixed, discontinuous Galerkin (DG) finite element-based method is derived based on classical DG-upwinding as frequently employed for advection terms in hydrodynamic problems. In particular, we showed the directional gradient's improved representation resulting from transport stabilization leads to greatly improved accuracy when compared to its non-stabilized counterpart from [19] in test cases with extreme anisotropy.

Many of the aforementioned approaches are based on the finite element method, which is often considered in magnetic confinement fusion simulations due to, for instance, their flexibility in meshing complex geometries. Next to the aforementioned DG-based work in [16, 17, 39], this includes additional DG-based approaches such as hybrid DG [15] and local DG [22]. However, DG-based methods may lead to challenges such as an increased number of degrees of freedom (DOFs), as well as an increased complexity in terms of implementation when compared to continuous Galerkin (CG) methods. Further, DG-based implementations often require interior penalty formulations, leading to a potential loss in the order of accuracy. For these reasons, and since magnetic confinement fusion scenarios generally involve a shock-free low Mach number regime, many finite element-based MHD codes used to simulate such regimes, such as [5, 7, 23, 36], deploy CG (or higher order regularity Galerkin) rather than DG methods.

In this paper, we present an extension of our previous DG-based upwind-stabilized work in [39] to CG-based discretizations. As in [39], we consider a mixed formulation of the anisotropic heat flux, using a form of the temperature's directional gradient along magnetic field lines as the auxiliary variable. Analogously to upwinding in the DG case, we apply a form of the Streamline Upwind Petrov–Galerkin (SUPG) method [8] for the two directional derivatives comprising the anisotropic diffusion operator. We motivate our particular choice of SUPG formulation and further demonstrate consistency of the resulting spatial discretization for the anisotropic heat flux equation. Additionally, in a series of numerical tests, we validate its order of accuracy as well as superior performance in realistic test cases when compared to a primal weak formulation and the mixed formulation of [19]. In particular, we study the spread of a temperature perturbation in a 2D magnetic flux surface, an MHD equilibrium in a full-torus tokamak domain including a temperature-dependent parallel conductivity coefficient, as well as the spread of a temperature perturbation through a flux tube in the latter domain. Further, the tests contain parameter scales suitable for MHD models, including the choice of time step and parallel conductivity.

The remainder of this paper is structured as follows: In Section 2, we review the anisotropic heat conduction equation and include a brief description of standard CG-based primal and mixed space discretizations. In Section 3, we introduce our novel SUPG-based scheme, including a discussion on its derivation as well as consistency and the choice of SUPG stabilization parameter. In Section 4, we present and discuss numerical results. Finally, in Section 5, we review our results and discuss possible future work.

2. Background

In this section, we briefly review the anisotropic heat flux equation. Further, we recall existing continuous Galerkin-based methods for the equation's spatial discretization, with which we will compare our novel method in the numerical results section.

Anisotropic heat flux equation. We consider the evolution of temperature T within a plasma domain Ω as governed by a heat flux $\mathbf{q}$ and forcing S according to

$$\frac{\partial T}{\partial t} - \nabla \cdot \mathbf{q} = S, \tag{2.1a}$$

$$\mathbf{q} = \kappa_\parallel \nabla_\parallel T + \kappa_\perp \nabla_\perp T, \tag{2.1b}$$

for parallel and perpendicular heat conductivities $\kappa_\parallel$, $\kappa_\perp$, respectively. The directional gradients are given by

$$\nabla_\parallel T = \mathbf{b}(\mathbf{b} \cdot \nabla T), \quad \nabla_\perp T = \nabla T - \nabla_\parallel T, \tag{2.2}$$

where $\mathbf{b} = \mathbf{B}/|\mathbf{B}|$ denotes the normalized magnetic vector field determining the direction of anisotropy. In the following, we consider a reformulated version of $\mathbf{q}$ that contains an isotropic and a purely anisotropic part

$$\mathbf{q} = \kappa_\Delta \mathbf{b}(\mathbf{b} \cdot \nabla T) + \kappa_\perp \nabla T, \tag{2.3}$$

where $\kappa_\Delta = \kappa_\parallel - \kappa_\perp$. Additionally, the equation is equipped with Dirichlet boundary conditions

$$T(\mathbf{x}, t) = T_{bc}(\mathbf{x}), \quad \mathbf{x} \in \partial\Omega, \tag{2.4}$$

where in magnetic confinement fusion, the domain's boundary is given by the device's plasma-facing wall. The remainder of the discussion in this work proceeds in a similar manner for Neumann boundary conditions, which are outlined briefly in Appendix A. Last, the "forcing" term S contains additional terms arising in the context of magnetohydrodynamic models (see (B.1) in Appendix B). In this paper, we will for simplicity assume S to be a known field, which is a reasonable assumption for instance in view of split-time discretizations in magnetohydrodynamics, where different dynamics — including the heat flux — are computed at separate stages.

Finally, we note that in practice, the heat conductivities κ_Δ, $\kappa_\perp$ may depend on e.g., the temperature, density, and magnetic field, rendering the anisotropic heat conduction problem nonlinear. For instance, under collisional closures of Braginskii [6], the parallel conductivity scales as

$$\kappa_\parallel \propto T^{5/2}; \tag{2.5}$$

for details see Appendix B. In the numerical results section below, we will consider such temperature dependent conductivities for one of the test cases, and use constant in time coefficients otherwise.

Existing spatial discretizations. A straightforward discretization for the anisotropic heat conduction equation (2.1) can be derived by considering T_h in a k^{th} polynomial order continuous Galerkin finite element space $\mathbb{V}_k^{\mathrm{CG}}$ equipped with the Dirichlet boundary conditions (2.4)

$$T_h \in \mathbb{V}_{k,bc}^{\mathrm{CG}} = \{\gamma \in \mathbb{V}_k^{\mathrm{CG}},\ \gamma|_{\partial\Omega} = T_{bc}\}, \tag{2.6}$$

and discretizing the elliptic operators weakly. This leads to

$$\left\langle \gamma, \frac{\partial T_h}{\partial t} \right\rangle + \langle \mathbf{b} \cdot \nabla \gamma, \kappa_\Delta \mathbf{b} \cdot \nabla T_h \rangle + \langle \nabla \gamma, \kappa_\perp \nabla T_h \rangle = \langle \gamma, S \rangle \qquad \forall\, \gamma \in \mathring{\mathbb{V}}_k^{\mathrm{CG}}, \tag{2.7}$$

where $\langle \cdot, \cdot \rangle$ denotes the L^2-inner product, and

$$\mathring{\mathbb{V}}_k^{\mathrm{CG}} = \{\gamma \in \mathbb{V}_k^{\mathrm{CG}},\ \gamma|_{\partial\Omega} = 0\}. \tag{2.8}$$

At higher levels of anisotropy, the accuracy of (2.7) can deteriorate drastically due to an inadequate representation of the directional gradient $\mathbf{b} \cdot \nabla T$. This leads to a spurious leak of parallel heat flux in the perpendicular direction, and in order to avoid this, [19] introduced an auxiliary variable to better represent the directional gradient. Given $T_h \in \mathbb{V}^{\mathrm{CG}}_{k,bc}$, the directional gradient can be argued to be a function in the discontinuous Galerkin space $\mathbb{V}^{\mathrm{DG}}_{k-1}$. In this work, we consider a version of the directional derivative scaled by $\sqrt{\kappa_\Delta}$ (see Remark 3.6 below), and the resulting mixed discretization is given by

$$\left\langle \gamma, \frac{\partial T_h}{\partial t} \right\rangle + \langle \mathbf{b} \cdot \nabla \gamma, \sqrt{\kappa_\Delta} \zeta_h \rangle + \langle \nabla \gamma, \kappa_\perp \nabla T_h \rangle = \langle \gamma, S \rangle \qquad \forall\, \gamma \in \mathring{\mathbb{V}}^{\mathrm{CG}}_k, \tag{2.9a}$$

$$\langle \phi, \zeta_h - \sqrt{\kappa_\Delta} \mathbf{b} \cdot \nabla T_h \rangle = 0 \qquad \forall\, \phi \in \mathbb{V}^{\mathrm{DG}}_{k-1}. \tag{2.9b}$$

Due to the ill conditioning (and corresponding time-step restriction for explicit integration) of discretized second derivatives, the anisotropic heat flux equation is typically discretized implicitly in time. In our numerical results section, we consider for simplicity a midpoint rule in time for (2.7), (2.9) as well as our novel space discretization to be introduced in the next section. Other more general implicit integration schemes could be considered, and we will elaborate on this more in Section 4. For scenarios with temperature dependent conductivity coefficients, this leads to a nonlinear system of equations, which can be solved using a Newton iteration method, or alternatively e.g., a quasi-Newton iteration approach with known estimates for the temperature field appearing in the coefficients.

3. Discretization

Having introduced the anisotropic heat flux equation as well as standard CG-based spatial discretizations thereof, we next present our novel upwind-stabilized method. For this purpose, we start with a brief motivation leading up to our method's definition, followed by a discussion of its properties.

Motivation. In [39], we built on the observation of [19] that a better discrete representation of the directional gradient $\mathbf{b} \cdot \nabla T$ is a key property to avoid spurious cross-diffusion. To improve this representation, we constructed a DG-based version of the mixed scheme (2.9), using classical DG upwinding [26] to discretize the directional gradient. In analogy, for CG-based schemes, one could consider an upwind-type approach such as the Streamline Upwind Petrov Galerkin (SUPG) method [8], where first, an additional term is included to introduce diffusion in the direction of streamlines — thereby targeting numerical noise introduced by advection. Second, additional terms are included to ensure a consistent method. Put together, the two changes can be achieved by considering a modified test function space, leading to a Petrov–Galerkin method where test and trial function spaces are different. For a simple transport equation discretized in a CG space, this modification corresponds to

$$\left\langle \gamma, \frac{\partial \theta_h}{\partial t} + \mathbf{u} \cdot \nabla \theta_h \right\rangle = 0 \quad \xrightarrow{\text{SUPG}} \quad \left\langle \gamma + \tau \mathbf{u} \cdot \nabla \gamma, \frac{\partial \theta_h}{\partial t} + \mathbf{u} \cdot \nabla \theta_h \right\rangle = 0 \qquad \forall\, \gamma \in \mathring{\mathbb{V}}_k, \tag{3.1}$$

for discrete tracer field θ_h, advecting velocity field $\mathbf{u}$, and SUPG stabilization parameter τ. In particular, the SUPG-modified equation is clearly still strongly consistent, in the sense that a solution θ to the strong equation $\frac{\partial \theta}{\partial t} + \mathbf{u} \cdot \nabla \theta = 0$ is also a solution to the SUPG weak form. Further, setting $\gamma = \theta_h$, we find a stabilizing effect

$$\frac{1}{2} \frac{\mathrm{d}}{\mathrm{d}t} \|\theta_h\|_2^2 = \cdots - \|\sqrt{\tau} \mathbf{u} \cdot \nabla \theta_h\|_2^2. \tag{3.2}$$

While upwind stabilization is well-motivated in hyperbolic problems, its use for the anisotropic diffusion equation is less immediately clear, as the latter equation is parabolic. To discuss this further, we

note that the algebraic system of equations for the (implicitly discretized in time) mixed method (2.9) can be written as

$$\begin{pmatrix} \frac{1}{\delta t} M_{CG} + L_\perp & D_\parallel^T \\ -D_\parallel & M_{DG} \end{pmatrix} \begin{pmatrix} \mathbf{T}_c \\ \boldsymbol{\zeta}_c \end{pmatrix} = \begin{pmatrix} R_T \\ R_\zeta \end{pmatrix}, \tag{3.3}$$

for time step δt, CG and DG mass matrices M_{CG}, M_{DG}, respectively, discrete directional gradient $D_\parallel$ corresponding to $\sqrt{\kappa_\Delta}\mathbf{b}\cdot\nabla$ and discrete Laplacian $L_\perp$ weighted by $\kappa_\perp$, where strong Dirichlet boundary conditions for T_h are applied appropriately to the matrix blocks. Further, we have degree of freedom coefficient vectors $\mathbf{T}_c$, $\boldsymbol{\zeta}_c$ to be solved for, and known right-hand side vectors R_T, R_ζ. Additionally, we ignored terms occurring in the case of temperature-dependent conductivity coefficients. Eliminating the auxiliary variable, we obtain

$$\left(\frac{1}{\delta t} M_{CG} + D^T M_{DG}^{-1} D + L_\perp\right) \mathbf{T}_c = \ldots, \tag{3.4}$$

that is we recover a symmetric positive definite matrix $M_{CG} + D_\parallel^T M_{DG}^{-1} D_\parallel + L_\perp$ representing the parabolic problem, where $D_\parallel^T M_{DG}^{-1} D_\parallel$ denotes a discrete form of a strictly anisotropic diffusion operator along $\mathbf{b}$ with diffusion coefficient κ_Δ. For the DG-upwind scheme introduced in [39], the corresponding system of equations can be written as

$$\begin{pmatrix} \frac{1}{\delta t} M_{DG} + L_{DG,\perp} + M_{BC} & G_\parallel^T \\ -G_\parallel & M_{DG} \end{pmatrix} \begin{pmatrix} \mathbf{T}_c \\ \boldsymbol{\zeta}_c \end{pmatrix} = \begin{pmatrix} R_T \\ R_\zeta \end{pmatrix}, \tag{3.5}$$

where $L_{DG,\perp}$ is a suitable (symmetric) DG-based discretization of the Laplacian weighted by $\kappa_\perp$, and $G_\parallel$ is a DG-upwinding-based discretization of the directional gradient $\sqrt{\kappa_\Delta}\mathbf{b}\cdot\nabla$. Additionally, M_{BC} is a boundary penalty term used to enforce the Dirichlet boundary conditions weakly. Similarly to (3.4), upon eliminating the auxiliary variable, we recover an operator for $\mathbf{T}_c$ that was shown in [39] to be symmetric positive definite. In other words, while upwinding is used for a better representation of the directional gradients occurring in the mixed formulation for the anisotropic diffusion equation, the overall method still represents an elliptic operator. With this in mind, we aim to obtain a similar result for CG-based methods including SUPG-stabilization for the directional gradients.

SUPG-based method. We next discuss our novel CG-based discretization, aiming for a stabilized form of the directional gradients while maintaining consistency and keeping in mind the discrete parabolic operator property discussed above. For this purpose, we first define suitable streamline upwind operators.

Definition 3.1. For magnetic field $\mathbf{B}$ and conductivity difference $\kappa_\Delta = \kappa_\parallel - \kappa_\perp$, we set the upwind velocity $\mathbf{s} \coloneqq \sqrt{\kappa_\Delta(T)}\mathbf{B}/|\mathbf{B}|$. For arbitrary functions χ, υ, the advective and flux-based SUPG-modified bilinear forms M_a, M_f are defined by

$$M^a(\chi, \upsilon) \coloneqq \langle \chi + \tau \mathbf{s}\cdot\nabla\chi, \upsilon\rangle, \tag{3.6a}$$

$$M^f(\chi, \upsilon) \coloneqq \langle \chi, \upsilon + \mathbf{s}\cdot\nabla(\tau\upsilon)\rangle. \tag{3.6b}$$

Further, for arbitrary functions ω, υ, the SUPG-modified bilinear form $G_\parallel^f$ is defined by

$$G_\parallel^f(\omega, \upsilon) \coloneqq M^f(\mathbf{s}\cdot\nabla\omega, \upsilon). \tag{3.7}$$

Note that M^a, M^f, and $G_\parallel^f$ are well-defined after discretization if χ, υ, ω are functions of a CG space, and further provided that the stabilization parameter τ is continuous. After discretization, M^a and M^f both are forms corresponding to SUPG-modified mass matrices, and they reduce to a mass matrix as $\tau \to 0$. Further, $G_\parallel^f$ corresponds to an SUPG-modified directional gradient.

Definition 3.2. The SUPG-modified, CG-based mixed discretization of the anisotropic heat flux equation (2.1) for velocity $\mathbf{s} \coloneqq \sqrt{\kappa_\Delta(T)}\mathbf{B}/|\mathbf{B}|$ is given by finding $(T_h, \zeta_h) \in \mathbb{V}_{k,bc}^{\mathrm{CG}} \times \mathbb{V}_k^{\mathrm{CG}}$ such that

$$M^a\left(\gamma, \frac{\partial T_h}{\partial t}\right) + G_\parallel^f(\gamma, \zeta_h) + \langle \nabla\gamma, \kappa_\perp \nabla T_h\rangle - \langle \tau\mathbf{s}\cdot\nabla\gamma, \nabla\cdot(\kappa_\perp\nabla T_h)\rangle = M^a(\gamma, S) \quad \forall\, \gamma \in \mathring{\mathbb{V}}_k^{CG}, \tag{3.8a}$$

$$M^f(\eta, \zeta_h) - G_\parallel^f(T_h, \eta) + \int_{\partial\Omega} \tau\eta(\mathbf{s}\cdot\nabla T_h)(\mathbf{n}\cdot\mathbf{s})\mathrm{d}S = 0 \qquad \forall\, \eta \in \mathbb{V}_k^{CG}, \tag{3.8b}$$

for outward unit normal vector $\mathbf{n}$.

Proposition 3.3. *The mixed discretization* (3.8) *is consistent, in the sense that a strong solution* T *to the non-discretized equation* (2.1) *satisfies* (3.8) *with a sufficiently regular test function space* $\mathbb{V}_T \supset \mathbb{V}_k^{CG}$ *and parameters* $\kappa_\parallel$, $\kappa_\perp$, τ *and* $\mathbf{B}$.

Proof. We want to show that (3.8) holds true with T_h set to a strong solution T in a suitable non-discretized function space $\mathbb{V}_T$ (such as the space of twice continuously differentiable functions $C^2(\Omega)$). For this purpose, we first eliminate ζ_h by using the definition of $G_\parallel^f$ and applying integration by parts in $G_\parallel^f(T_h, \eta)$, which leads to a reformulation of (3.8b) of the form

$$M^f(\eta, \zeta_h) = \langle \mathbf{s}\cdot\nabla T, \eta + \mathbf{s}\cdot\nabla(\tau\eta)\rangle - \int_{\partial\Omega} \tau\eta(\mathbf{s}\cdot\nabla T)(\mathbf{n}\cdot\mathbf{s})\mathrm{d}S = \left\langle \eta, \mathbf{s}\cdot\nabla T - \tau\nabla\cdot\left(\mathbf{s}(\mathbf{s}\cdot\nabla T)\right)\right\rangle, \tag{3.9}$$

where we assumed sufficient regularity for the test functions η, as well as T, $\mathbf{B}$ and τ. Next we set $\eta = \mathbf{s}\cdot\nabla\gamma$ in (3.9) for test function γ appearing in the temperature equation (again assuming sufficient regularity), which yields

$$G_\parallel^f(\gamma, \zeta_h) = M^f(\mathbf{s}\cdot\nabla\gamma, \zeta) = \left\langle \mathbf{s}\cdot\nabla\gamma, \mathbf{s}\cdot\nabla T - \tau\nabla\cdot\left(\mathbf{s}(\mathbf{s}\cdot\nabla T)\right)\right\rangle. \tag{3.10}$$

We can then substitute this in the second term in (3.8a), from which we obtain

$$M^a\left(\gamma, \frac{\partial T}{\partial t}\right) + \left\langle \mathbf{s}\cdot\nabla\gamma, \mathbf{s}\cdot\nabla T - \tau\nabla\cdot\left(\mathbf{s}(\mathbf{s}\cdot\nabla T)\right)\right\rangle + \langle \nabla\gamma, \kappa_\perp\nabla T\rangle - \langle \tau\mathbf{s}\cdot\nabla\gamma, \nabla\cdot(\kappa_\perp\nabla T)\rangle = M^a(\gamma, S) \qquad \forall\, \gamma \in \mathring{\mathbb{V}}_T, \tag{3.11}$$

where analogously to the discrete case, $\mathring{\mathbb{V}}_T$ corresponds to $\mathbb{V}_T$ equipped with homogenous boundary conditions. Finally, collecting terms in (3.11), we obtain

$$\left\langle \gamma, \frac{\partial T}{\partial t}\right\rangle + \langle \mathbf{s}\cdot\nabla\gamma, \mathbf{s}\cdot\nabla T\rangle + \langle \nabla\gamma, \kappa_\perp\nabla T\rangle - \langle \gamma, S\rangle = -\left\langle \tau\mathbf{s}\cdot\nabla\gamma, \frac{\partial T}{\partial t} - \nabla\cdot\left(\mathbf{s}(\mathbf{s}\cdot\nabla T)\right) - \nabla\cdot(\kappa_\perp\nabla T) - S\right\rangle, \tag{3.12}$$

for any test function $\gamma \in \mathring{\mathbb{V}}_T$. This equation is satisfied by the strong solution T as required, since the left-hand side is the standard weak formulation of the anisotropic heat flux equation (2.1), while the right-hand side consists of the inner product of (2.1) in strong form together with $\tau\mathbf{s}\cdot\nabla\gamma$. ∎

We complete this section with a brief discussion and remarks on the spatial discretization (3.8).

Ellipticity. First, given an implicit discretization in time, the SUPG-modified method (3.8) can be written as

$$\begin{pmatrix} \frac{1}{\delta t}M^a + L_\perp^f & {G_\parallel^f}^T \\ -G_\parallel^f - G_{\parallel,b}^f & M^f \end{pmatrix} \begin{pmatrix} \mathbf{T}_c \\ \boldsymbol{\zeta}_c \end{pmatrix} = \begin{pmatrix} R_T \\ R_\zeta \end{pmatrix}, \tag{3.13}$$

where in a slight abuse of notation, here M^a, M^f, and $G_\parallel^f$ denote the matrix version of the corresponding forms described in Definition (3.1), and $L_\perp^f$ corresponds to the last two terms on the left-hand

side of (3.8a). Further, $G^f_{\parallel,b}$ denotes a boundary-type directional gradient matrix corresponding to the (negative version of the) last term on the left-hand side of (3.8b). Eliminating the auxiliary variable as before, we obtain an algebraic form of the heat equation given by

$$\left(\frac{1}{\delta t}M^a + G^{f\,T}_{\parallel}\left(M^f\right)^{-1}\left(G^f_{\parallel} + G^f_{\parallel,b}\right) + L^f_{\perp}\right)\mathbf{T}_c = \dots, \tag{3.14}$$

which retains the same "transpose gradient"-"mass inverse"-"gradient" structure for the anisotropic heat flux term as in (3.14), up to the SUPG-modification in M^f and the additional SUPG-related term $G^f_{\parallel,b}$. This loss in ellipticity of the anisotropic diffusion operator — and also the SUPG-modified isotropic operator $L^f_{\perp}$ — is a result of ensuring the discretization to be consistent by reverting to a Petrov–Galerkin method. Note that as $\tau \to 0$, M^f reduces to the usual mass matrix, $G^f_{\parallel,b}$ vanishes, and $L^f_{\perp}$ reduces to $L_{\perp}$, and we therefore recover the parabolic structure in this limit. As usual for SUPG methods, in practice the Petrov–Galerkin type modification translates to potential instabilities if the SUPG stabilization parameter τ is not chosen carefully. The loss of symmetry and ellipticity seems like a downside from a linear solver perspective, but for very high anisotropy, state-of-the-art elliptic solvers like geometric and algebraic multigrid methods are not robust [39].

Stabilization parameter. In terms of the choice of stabilization parameter, we recall that the SUPG-modification to test functions γ is given by $\tau\mathbf{s}\cdot\nabla\gamma$, for $\mathbf{s} = \sqrt{\kappa_\Delta(T)}\mathbf{B}/|\mathbf{B}|$. The heat conductivity's unit is m^2/s, and therefore for the overall SUPG modification to be non-dimensional, we require the unit of τ to be $\mathrm{s}^{1/2}$. This is different to the standard SUPG method for advection equations such as (3.1), where the streamline velocity $\mathbf{u}$ is in m/s and τ represents a time scale in seconds. The difference follows since here we do not consider a hyperbolic equation. If for instance instead we considered a discretization for a mixed form of an anisotropic wave equation — which could be achieved e.g. by replacing ζ_h by $\frac{\partial\zeta_h}{\partial t}$ in (3.8b) and taking $\sqrt{\kappa_\Delta}$ as a wave speed with unit m/s — τ would scale in seconds again. Guided by common choices of τ in the computational fluid dynamics literature [37], in the numerical results section below we set

$$\tau = \left(\frac{2}{\sqrt{\delta t}} + \frac{k\sqrt{\kappa_\Delta}}{\delta x}\right)^{-1}, \tag{3.15}$$

for local mesh cell length scale δx and polynomial degree k of the CG function space T_h and ζ_h are discretized in. Note that as we refine the mesh and time step, the stabilization contribution gets smaller, reaching $\tau = 0$ in the space-time refined limit. In this case, the mixed SUPG-modified method (3.2) reduces to the standard mixed method (2.9) up to the choice of finite element space for the auxiliary variable ζ_h. In (2.9) the auxiliary variable is placed in a DG space in view of ζ_h as a gradient of the CG field T_h; here we placed ζ_h in a CG space since we apply SUPG — which requires CG spaces — to both the T_h and ζ_h equations. Finally, as mentioned below Definition 3.1, we note that τ occurs within gradient operations in our SUPG-modified mass matrix M^f, and therefore needs to be continuous for M^f to be well-defined. In practice, τ is evaluated as an expression according to (3.15) on quadrature points, including the local cell length scale δx, which we found to work well.

Additional adjustments to τ could for instance include information on the mesh relative to the magnetic field. This is of relevance for tokamak applications with frequently very anisotropic cells (see Figure 4.5 in the numerical results section for a mildly anisotropic case); if a magnetic field line passes through such a cell along the elongated direction, we would expect the parallel diffusion to be significantly better resolved than if the field line passes through the short direction. More generally, an alignment of conductivity and mesh anisotropies may lead to a discrete system of equations that is overall more isotropic in character, and this can be addressed on the level of our stabilization paremeter e.g., by replacing δx in (3.15) by an average distance covered by $\mathbf{B}$ passing through the given cell.

Remark 3.4 (Alternative SUPG-based forms)**.** In the original mixed formulation (2.9), the auxiliary variable ζ directly corresponds to $\mathbf{s} \cdot \nabla T$. For the SUPG-modified method (3.8), this no longer holds true, and more generally, there is some flexibility with the exact setup of our SUPG modification and definition of auxiliary variable ζ. Simpler formulations include e.g. a more straight-forward application of SUPG-based on the original mixed form (2.9), leading to

$$\left\langle \gamma, \frac{\partial T_h}{\partial t} \right\rangle + \langle \mathbf{s} \cdot \nabla \gamma, \zeta_h \rangle + \langle \nabla \gamma, \kappa_\perp, \nabla T_h \rangle = \langle \gamma, S \rangle - \langle \tau \mathbf{s} \cdot \nabla \gamma, T_{rhs} \rangle \qquad \forall\, \gamma \in \mathring{\mathbb{V}}_k^{\mathrm{CG}}, \tag{3.16a}$$

$$\langle \eta + \tau \mathbf{s} \cdot \nabla \eta, \zeta_h - \mathbf{s} \cdot \nabla T_h \rangle = 0 \qquad \forall\, \eta \in \mathbb{V}_k^{\mathrm{CG}}, \tag{3.16b}$$

where T_{rhs} corresponds to the strong form (2.1) of the anisotropic heat flux equation applied to the discrete field T_h. While this form is clearly consistent — T_{rhs} vanishes for strong solutions T and ζ still corresponds to $\mathbf{s} \cdot \nabla T$ in strong form — we found it to lead to worse results than the version defined in Definition 3.2, possibly because it can be shown to not lead to an approximately antisymmetric off-diagonal block structure as seen in (3.3), (3.5) and (3.13). Finally, we also did not find any SUPG-based strategies for the primal form of the anisotropic heat equation — that is without the use of an auxiliary variable — that performed well in practical test cases.

Remark 3.5 (Solver considerations)**.** While we focus on spatial discretizations in this work, we briefly discuss solver considerations here. One possible way to solve for mixed systems such as (3.3) and (3.13) is to eliminate the auxiliary variable in order to obtain a Schur complement in the temperature field T, given by (3.4) and (3.14), respectively. For (3.14), the Schur complement could then be preconditioned using an approximate version in which the inverse of the modified mass matrix M^f is replaced by the inverse of its diagonal. Alternatively, (3.12) in the proof of Proposition 3.3 corresponds to a weak form of the Schur complement before discretization, and replacing the continuous space $\mathring{\mathbb{V}}_T$ by $\mathring{\mathbb{V}}_k^{CG}$ would yield a variational form corresponding to another possible approximate Schur complement. However, standard multigrid methods do not perform well for extremely anisotropic problems due to the stark difference in error frequency along versus across the field lines defining the direction of anisotropy. In tests using the inverse of the diagonal of M^f, we found the Schur complement approach to be particularly challenging for the SUPG-modified formulation leading to the system (3.13), likely due to additional asymmetry introduced by the SUPG-based modifications in the Schur complement (3.14).

Alternatively, one can follow the approach of [39], where the system is considered as two weakly coupled directional derivatives, and multigrid methods specialized on transport such as AIR [29, 30] are applied instead. Such methods exploit the sparsity in stencil and matrix structure that arises due to upwinding, including of SUPG-type [29]. In future work, we aim to combine our SUPG-modified spatial discretization (3.2) with such an approach.

Remark 3.6 (κ_Δ-scaling)**.** Finally, we note that in the original formulation of [19], ζ_h corresponds to $\mathbf{b} \cdot \nabla T_h$ instead of the scaled version $\sqrt{\kappa_\Delta}\mathbf{b} \cdot \nabla T$, and a full factor of κ_Δ is kept in the temperature equation instead. In this work, we use a factor of $\sqrt{\kappa}_\Delta$ in both equations for consistency with our previous work [39], in which this type of multiplicative split was considered in view of the solver strategy based on two weakly coupled directional derivatives as mentioned in the previous remark. In particular, for the SUPG-based discretization (3.8) introduced in this work, the multiplicative split is not strictly required, and it is also possible to derive an analogous discretization based on an auxiliary variable corresponding to the non-scaled term $\mathbf{b} \cdot \nabla T_h$. The required modifications to (3.8) for a non-scaled auxiliary variable then include the need for two separate stabilization parameters τ_ζ and τ_T, which are formulated analogously to (3.15) with $\sqrt{\kappa_\Delta}$ replaced by 1 and κ_Δ, respectively.

4. Numerical results

Having introduced our novel spatial discretization in Definition (3.2), we next consider a series of numerical tests to validate the discretization's order of accuracy as well as to compare it against the existing CG methods (2.7) and (2.9) in scenarios relevant to magnetic confinement fusion simulations. As mentioned in the background section, we couple these spatial discretizations to an implicit midpoint rule in time. Further, for some of the test cases, we will consider a small initial time step that is increased to a final fixed value over a relatively small number of steps, in order to avoid larger time stepping errors related to adjustments of the initial conditions projected or interpolated into the discrete spaces. Small initial time steps related to such initial equilibrations are common in magnetic confinement fusion simulations, and frequently arise automatically in PDE discretizations when using adaptive time stepping schemes (see e.g., Figures 2,3 in [35] and Figure 2 in [40]) The tests are implemented in the finite element library Firedrake [21], which heavily relies on PETSc [2]. The system of equations are solved using the direct solver library MUMPS [1], and in the case of temperature-dependent conductivities, we use the nonlinear Newton solver functionality from PETSc, together with an exact Jacobian computed automatically in Firedrake. For each test case, we use the SUPG parameter τ as specified in (3.15). We will consider 2D quadrilateral and 3D prism meshes, using the standard continuous Galerkin space P_k for our temperature space, and the discontinous Galerkin space dP_{k-1} for ζ_h in the mixed discretization (2.9). Our magnetic field $\mathbf{B}$ is set to be a function of a divergence conforming finite element space $\mathbb{V}_B$ of equal polynomial degree k as our temperature CG space P_k. For the above types of meshes, $\mathbb{V}_B$ corresponds to the quadrilateral Raviart–Thomas space $RT^f_{c_k}$ and a suitable tensor product including the triangular Raviart–Thomas space RT^f_k, respectively [31]. There is no particular reason for this choice of $\mathbb{V}_B$ other than related work on magnetohydrodynamics by the authors, and in general we found similar results for other choices of magnetic field space.

4.1. Convergence

First, we validate the SUPG-based discretization's order of accuracy using a simple 2D test case modeling dissipation of a Gaussian profile along magnetic field lines. The domain Ω is given by a unit square $[0,1]^2$, and we set the magnetic field and initial temperature to

$$\mathbf{B} = (B_x, B_y)^T = [1,0]^T, \quad T_0(x,y) = f(y)\, e^{-(x-0.5)/\sigma^2}, \quad f(y) = \frac{1}{2}(1-\cos(2\pi y)), \tag{4.1}$$

for $\sigma = 0.2$. The conductivities are set to $\kappa_\parallel = 1$ and $\kappa_\perp = 0.01$, and the boundary conditions are set equal to the values of the initial temperature field along the boundary. In the absence of perpendicular diffusion, the problem consists of decoupled 1D diffusion equations applied to Gaussian profiles on an interval $[0,1]$, subject to Dirichlet boundary conditions. The solution to each 1D profile can be computed analytically using separation of variables and a Fourier decomposition given by $T_F = a_0 + \sum_{n=1}^{m} b_n \sin(\pi n x) e^{-\kappa_\parallel \pi^2 t}$. We set $m = 200$ and use $f(y)T_F(\mathbf{x}, 0)$ for the initial projection into the initial discrete field T_{h_0}, and further set our forcing term equal to $S(t) = -\kappa_\perp \Delta(f(y)T_F)$. This way, the effects of perpendicular diffusion are balanced out, and T evolves according to the aforementioned analytic solution. We can then compare our discrete solution T_h against this analytic solution by considering the relative L^2 error

$$e(t) = \frac{\|T_h(t) - T_F(0)\|_2}{\|T_F(0)\|_2}, \tag{4.2}$$

where the finite element field T_h and the Fourier mode expression T_F (with pre-computed coefficients) are evaluated at quadrature points.

The polynomial degree for the temperature CG space is set to $k = 2$, and we consider a regular quadrilateral mesh for Ω, with 10×2^{k_x} cells in each coordinate direction, for $k_x \in \{1, 2, 3\}$. Further, the time step is set to $\delta t = 10^{-5}$ for each spatial resolution, and we run up to $t_{max} = 400\delta t$. The resulting error convergence rates for the two standard primal and mixed as well as SUPG-based discretizations (2.7), (2.9) and (3.8), respectively, are given in Figure 4.1.

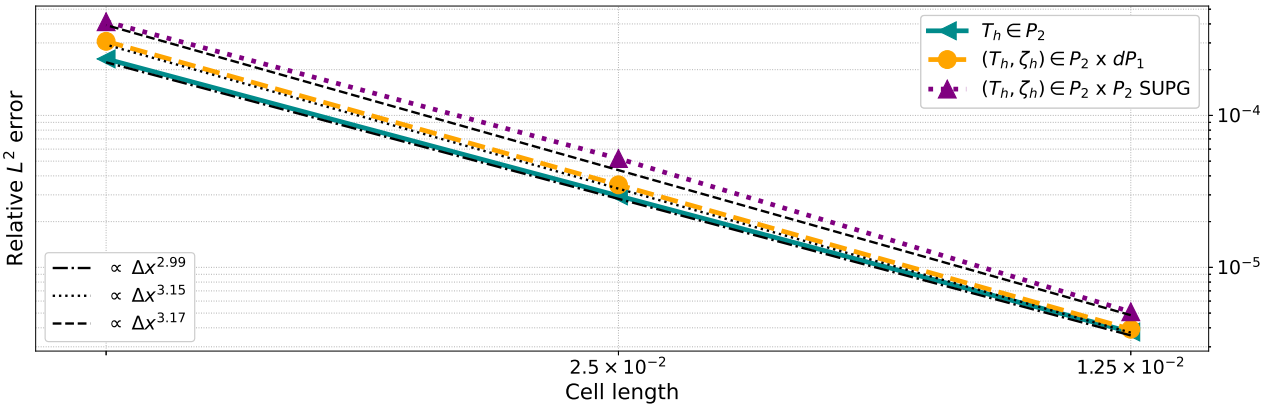


FIGURE 4.1. Convergence plot for relative L^2 errors (4.2) for analytic Gaussian profile diffusion test case (4.1), for primal CG (solid cyan), standard mixed CG (dashed orange), and SUPG-based (dotted purple) formulations (2.7), (2.9), and (3.8), respectively.

For second order CG spaces, the expected convergence rate for the isotropic heat equation is given by 3, and all three CG discretizations approximately lead to such a rate for this test case. Further, as the degree of anisotropy is relatively small and the time step is small with respect to the parallel diffusion, the difference in error between the discretizations is small. When comparing the two mixed discretizations, we observe a slightly elevated error for the SUPG setup; in further tests, we found this to be due to using the CG space P_2 instead of the DG space dP_1 for the auxiliary variable. In other words, the latter DG space may be more suitable for representing the directional derivative $\mathbf{s} \cdot \nabla T$ than P_2; however, as we will see below, using P_2 leads to advantages in more practical test cases with higher anisotropy ratios, as it allows for the SUPG-stabilized approach detailed in the previous section.

4.2. **Magnetic flux surface**

Having discussed our spatial discretization's order of accuracy, we next consider three tokamak simulation related test cases. The first represents a temperature perturbation on an idealized 2D magnetic flux surface (see Figure 4.5). For this purpose, we consider a 2D periodic domain $\Omega = [0, L_x] \times [0, L_y]$ for $(L_x, L_y) = (5, 4)$, with a magnetic field given by

$$\mathbf{B} = (1, m)^T, \tag{4.3}$$

where $m = 20$ denotes the magnetic field's slope with respect to the surface. We then construct the temperature field perturbation in orthogonal coordinates (ξ, ζ), where the ξ coordinate is aligned with the magnetic field lines, and the origin is placed at $(x, y) = (L_y/m, L_y/2)$. Further, the unit lengths for ξ and ζ are defined such that the xy-coordinate origin is at $(-2, -1)$ in $\xi\zeta$-coordinates. We then

set the initial temperature equal to $T_0(\xi,\zeta) = T(\xi,\zeta,t)_{|_{t=0}}$, where

$$T(\xi,\zeta,t) = T_b + \left(1+4\kappa_\parallel \xi_0^2 t\right)^{-1/2} \; e^{-\frac{(\xi_0\xi)^2}{\sqrt{1+4\kappa_\parallel \xi_0^2 t}}} \; (1+\cos(\pi\zeta))/2, \tag{4.4}$$

if $\zeta \in (-1,1)$, and $T = T_b$ otherwise. Further, $T_b = 0.2$, $\xi_0 = 2.5$ and we consider conductivities $\kappa_\parallel = 10$, $\kappa_\perp = 0$. Note that since there is no conductivity in the perpendicular direction, the problem decouples into a series of 1D Gaussians along separate magnetic field lines, and we obtain an analytic solution in time given by (4.4) together with its corresponding periodically shifted solutions with respect to our underlying domain Ω (up to the background temperature value T_b). For the discretization, we consider a pseudo-regular quadrilateral mesh with 120×96 cells, with inner vertices perturbed randomly by a factor of up to $0.1\delta x$ in order to avoid a decreased error due to an equal alignment of the magnetic field with the mesh throughout the domain. Additionally, we start with a small time step of $\delta t = 10^{-4}$, which is increased linearly to $\delta t = 0.1$ over 20 time steps. The simulation is run for the two standard primal and mixed spatial discretizations (2.7) and (2.9), as well as the SUPG-based form (3.8). Results for the qualitative field development and qualitative as well as quantitative error development are shown in Figures 4.2, 4.3 and 4.4, where we consider the relative L^2 error as defined in (4.2) (with T_F replaced by the expression for T from (4.4), together with periodically shifted profiles).

We find that the primal CG formulation exhibits a significant amount of spurious cross diffusion. This can be seen by considering the magnetic field line along $\zeta = 1$ (white straight lines in Figure 4.2), which initially does not contain any of the temperature perturbation. Since there is no perpendicular diffusion, we therefore expect $T = T_b$ along this field line throughout the simulation, which is clearly not the case for the primal CG formulation. Conversely, this is satisfied to a much better degree for both of the mixed formulations. However, the field development of the mixed formulations differs in that the standard mixed CG one leads to spurious grid scale noise (see Figure 4.3). In contrast, the SUPG-based formulation leads to the smallest error, which itself is dominated by noise near the perturbation's initial location. This noise is likely due to an initial condition projection error combined with a time stepping-related error in the perpendicular direction, which cannot be attenuated by the parallel diffusion. Overall, at the end of the simulation's runtime, we find relative L^2 errors of approximately 1.3×10^{-2}, 2.5×10^{-3} and 2.7×10^{-4} for the primal CG, standard mixed CG, and SUPG-based mixed CG formulations, respectively, demonstrating a strong reduction in error through the auxiliary variable as well as the SUPG-stabilization.

4.3. **Tokamak equilibrium**

Next, we consider a 3D tokamak scenario starting from an axisymmetric MHD equilibrium

$$\nabla p = \frac{1}{\mu_0}(\nabla \times \mathbf{B}) \times \mathbf{B}, \tag{4.5}$$

for pressure p and vacuum permeability μ_0. The equilibrium is generated by a Grad–Shafranov solver code [28], which returns p inside the separatrix — that is the boundary confining the closed magnetic field lines; see Figure 4.5 — and the components of $\mathbf{B}$ (in cylindrical coordiantes (R,φ,Z)) on a regular rectangular grid corresponding to a poloidal (i.e., (R,Z)) plane. The magnetic field's safety factor — that is the ratio of toroidal to poloidal rotations of magnetic field lines — varies between approximately $q_0 = 1.5$ at the magnetic axis and $q_{95} = 6$ towards the separatrix. The equilibrium resembles an ITER discharge, with a magnetic field strength of $B_0 = 5.42$ T as well as a pressure of $p_0 = 656$ kPa, both measured at the magnetic axis (see Figure 4.5). We then interpolate the rectangular grid's field data into finite element spaces defined on a 3D mesh. For faster computations — recalling that for simplicity we consider a direct solver rather than a more scalable iterative solver in this work — instead of considering the entire torus, the mesh consists of a toroidally (i.e. φ-directed)

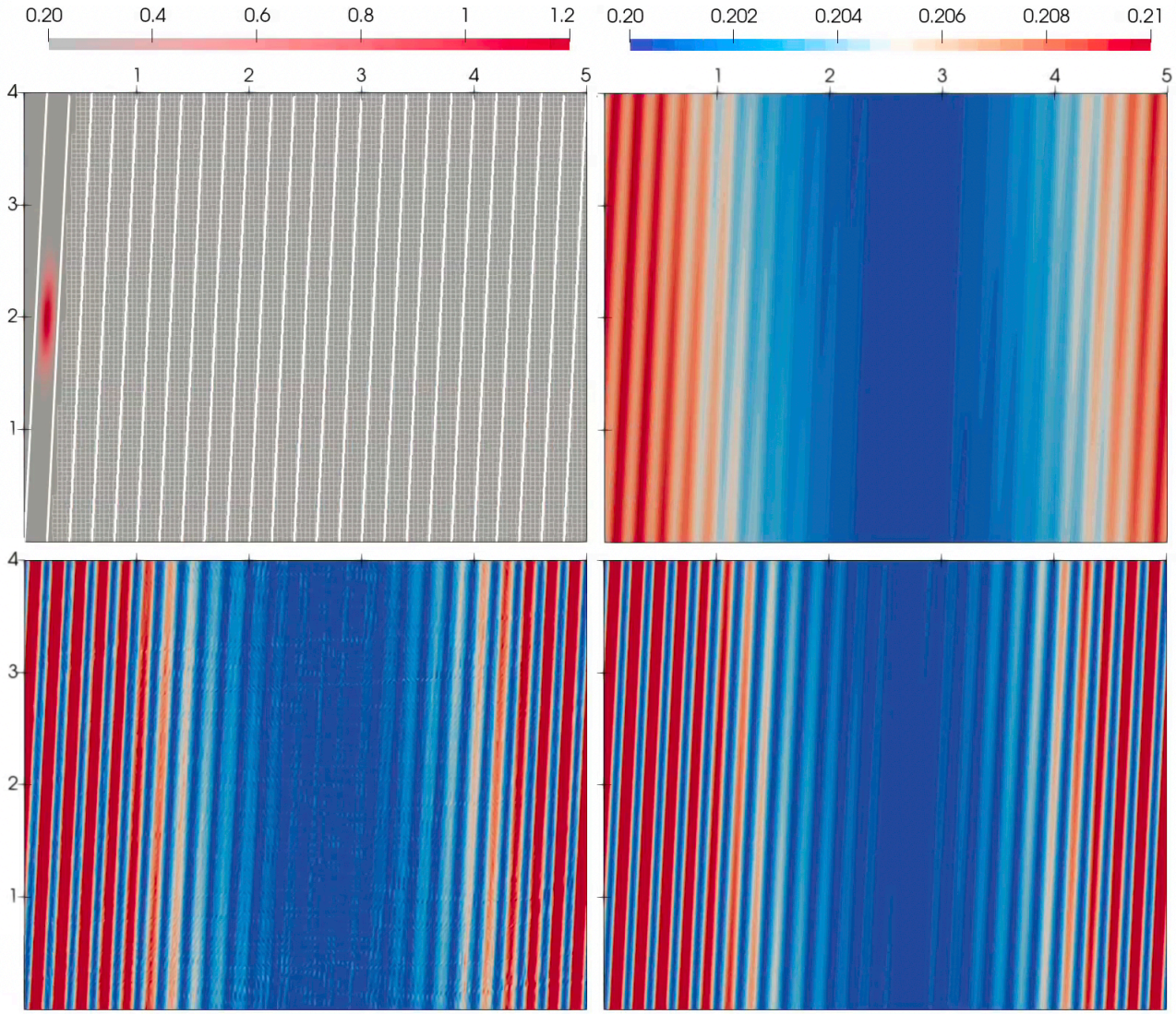


Figure 4.2. Images depicting temperature field T for test case (4.4). Top left: T at initial time, together with straight white lines tracing a single (periodic) magnetic field line and a section of the quadrilateral mesh used for the simulation runs. Top right: primal CG formulation (2.7) at $t \approx 14$. Bottom row, left to right: standard mixed and SUPG-modified mixed formulations (2.9) and (3.8), respectively, at $t \approx 14$.

periodic section that is constructed by periodically extruding an irregular triangular poloidal mesh in the toroidal direction. The extrusion length is $\pi/4$ radians, and the resulting 3D mesh consists of prisms aligned in the toroidal direction. Note that the extrusion direction is curved, and we therefore represent the mesh geometry by polynomials of degree equal to that chosen for our temperature space. Finally, we note that in practice, tokamak meshes are constructed to be aligned with the flux surfaces. However, here we refrain from doing so since we are interested in testing our new spatial discretization for scenarios in which the magnetic flux surfaces are not aligned with the mesh. In particular, while flux surface-aligned meshes can be constructed for initial magnetohydrodynamic equilibria, this need

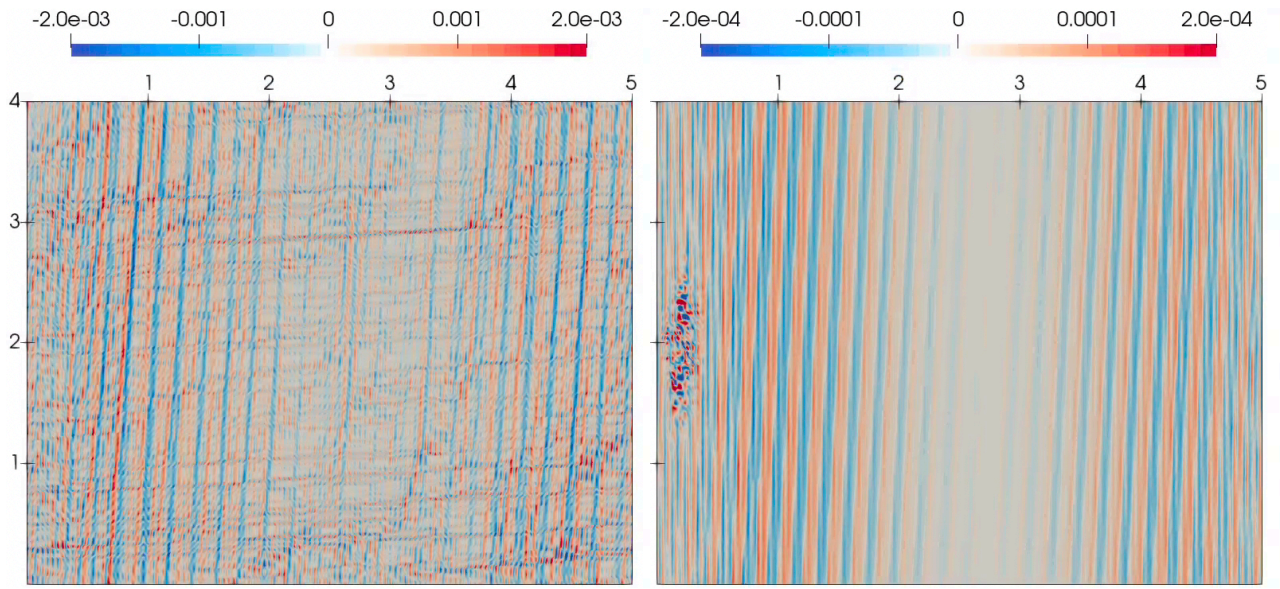


FIGURE 4.3. Images depicting error of temperature field T for test case (4.4). Left to right: standard mixed and SUPG-modified mixed formulations (2.9) and (3.8) with color ranges of $\pm 2 \times 10^{-3}$ and $\pm 2 \times 10^{-4}$, respectively, at $t \approx 14$.

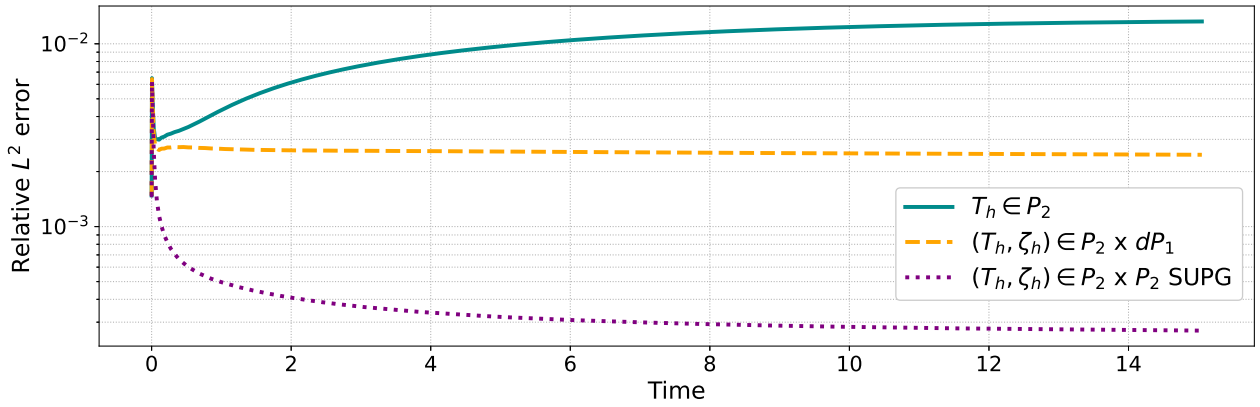


FIGURE 4.4. Relative L^2 error development over time for magnetic flux surface test case, for primal CG (solid cyan), standard mixed CG (dashed orange), and SUPG-based (dotted purple) formulations (2.7), (2.9), and (3.8), respectively.

not be the case at later stages for more complex simulations involving e.g., magnetohydrodynamic instabilities that change the magnetic topology.

The pressure is defined inside the separatrix only, noting that $(\nabla \times \mathbf{B}) \times \mathbf{B} = \mathbf{0}$ outside the separatrix. We add a floor value to p throughout the domain, which corresponds to a value found between the separatrix and the tokamak chamber's wall. This value is typically very small, and for the case of meshes not aligned with the separatrix, it may be smaller than the values resulting from transferring the 2D regular poloidal plane data onto the finite element grid for second or higher order polynomial

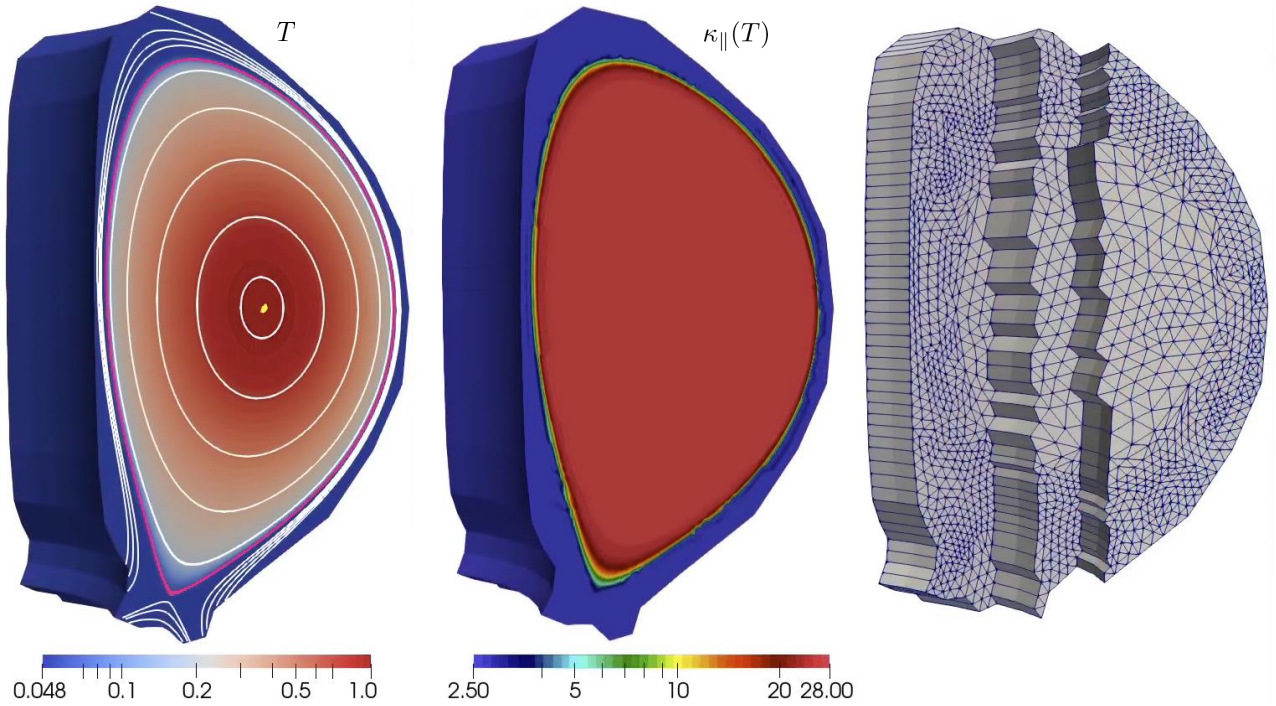


FIGURE 4.5. Images depicting toroidally periodic domain for tokamak test case setup. Left: non-dimensionalized initial temperature field T, and white lines denoting (vertical cross-sections of axisymmetric) magnetic flux surfaces. The non-dimensionalized magnetic field's (R, Z)-component is of magnitude at most approximately 0.3, while its ϕ-component's strength varies between approximately 0.75 and 1.5. The purple curve and yellow dot denote the separatrix and magnetic axis, respectively. Center: initial parallel conductivity $\kappa_\parallel$ defined according to (4.6). Right: cross-section of once (poloidally) refined second order mesh created from toroidally periodic extrusion created from 2D irregular triangular poloidal mesh.

spaces, since the pressure field exhibits a very steep gradient just inside the separatrix. This in turn may lead to negative pressure and therefore temperature values in the simulation, causing errors e.g. when evaluating the temperature-dependent parallel conductivity (2.5). Since we do not use a separatrix-aligned mesh in this test case, we resolve this by adding a larger floor value, given by a factor of $p_b = 0.05 p_0$ of the pressure at the magnetic axis.

Given p and $\mathbf{B}$, we non-dimensionalize the problem and domain with respect to their values $p_b + p_0$ and B_0 at the magnetic axis, as well as the tokamak chamber's length scale. The temperature T and the corresponding boundary condition T_{bc} is computed from p and p_b through a representative density number n_0 and the atomic number Z of the plasma in consideration. Further, we consider a scenario in which the perpendicular diffusion is set to $\kappa_\perp = 0$, while the parallel diffusion is set to

$$\kappa_\parallel = 8.8 \times 10^3 \, f(T)^{5/2}, \quad f(T) = T_l - \sigma_l \ln\left(1 + e^{-\frac{T - T_l}{\sigma_l}}\right), \tag{4.6}$$

for $T_l = 0.1$ and $\sigma_l = 0.04$. This choice follows the Braginskii transport model [6], together with an upper limit enforced through f. The limit reflects physical restrictions to the Braginskii model and is included purely to obtain a more realistic parallel conductivity. In particular, it is not related to any

numerical restrictions; for details on the non-dimensionalization and choice of $\kappa_\parallel$, see Appendix B. For $T < T_l$, we have $f(T) \approx T$, while for $T > T_l$, we have $f(T) \approx T_l$. Since the initial axisymmetric pressure profile's contours are aligned with the magnetic flux surfaces, the absence of perpendicular diffusion implies an analytic solution of the anisotropic heat flux equation given by $T(\mathbf{x}, t) = T(\mathbf{x}, 0)$. For this reason, any change of the discretely computed temperature T_h over time can be attributed to spurious cross-diffusion of temperature in the perpendicular direction.[1]

We consider the spatial discretizations (2.7), (2.9) and (3.8) with CG spaces of degree $k = 2$, as well as $k = 1$ in the case of (3.8). For $k = 2$, we use 3 cells in the toroidal extrusion, and for $k = 1$, we use a once refined version of the underlying 2D triangular poloidal mesh, with 6 cells in the toroidal extrusion, leading to an overall equal number of degrees of freedom as for the $k = 2$ case. These two setups then serve as the base resolution in space, which is refined in the poloidal direction once — noting that spurious cross-diffusion is primarily a problem resulting from an inaccurate resolution of the parallel heat flux' poloidal component — for a total of 2 runs for each choice of spatial discretization (see Figure 4.5 for the once refined mesh for $k = 2$). To avoid oscillatory errors related to initial adjustments of the discrete temperature profile to the large parallel diffusion term, we start with a small time step of $\delta t = 10^{-5}$, which is increased linearly to $\delta t = 5$ over 100 time steps. The simulation is run up to $t_{max} = 10^5$, noting that after non-dimensionalization, the time unit is given by the time taken for an Alfven wave to travel a characteristic distance within the tokamak chamber. In particular, the choice of final time step δt and simulation runtime t_{max} corresponds to values that may be considered in (reduced) magnetohydrodynamic simulations used to study slow, dissipative time scales of interest for resistive magnetohydrodynamic instabilities [4, 9, 33]. On the other hand, the choice of initial time step is somewhat arbitrary, although we note that the initial spin-up phase of 100 time steps is small compared to the $\sim 2 \times 10^4$ time steps for the overall simulation. As mentioned in the beginning of this section, such spin-ups are useful in MHD codes, since the initial error due to interpolating from the GS equilibrium solver grid to the MHD scheme's grid may initially lead to both large parallel diffusion values and significant magnetosonic waves.

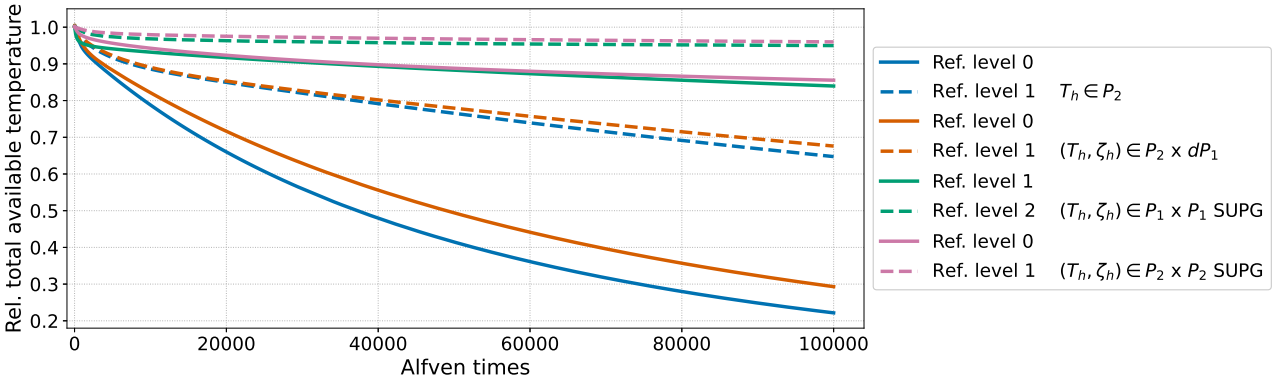
An important quantity with respect to such simulations is the total amount of internal energy stored within the plasma chamber; in our case for the non-dimensionalized anisotropic heat equation, this can be translated to $\|T(t)\|_1$. For the case of $\kappa_\perp = 0$, we expect this quantity to remain constant. In practice after discretization, spurious cross-diffusion will cause temperature to be diffused across the separatrix and then "removed" from the domain via the fixed small Dirichlet boundary conditions, noting that all magnetic field lines outside the separatrix start and end at the domain's boundary. In the worst case, all temperature beyond a constant background profile $T \equiv T_{bc}$ is diffused. Time series of the relative total available temperature

$$\frac{\|T(t)\|_1 - \|T_{bc}\|_1}{\|T(0)\|_1 - \|T_{bc}\|_1}, \tag{4.7}$$

for the different discretizations and spatial resolutions are depicted in Figure 4.6.

We find that for the lower resolutions, the primal CG formulation (2.7) leads to a loss of approximately 78% of the available total temperature, with some improvement when using the mixed formulation (2.9), down to approximately 71%. In contrast, the first and second order runs for the novel SUPG-based discretization (3.8) lead to a significantly reduced loss at approximately 16% and 14%, respectively. In other words, comparing the second order standard mixed and mixed SUPG-modified methods, we obtain a 5-fold decrease in temperature loss. Here, we note that the parallel conductivity's temperature-dependence is an important mechanism that amplifies spurious perpendicular diffusion across the separatrix. This is because in the course of spurious cross-diffusion, the temperature in the region just outside the separatrix is raised, thereby locally increasing the parallel conductivity. This in

[1] An additional error may be attributed to interpolating from the GS equilibrium solver grid onto the finite element grid; this can be avoided e.g., by using the same poloidal mesh in a finite-element based GS solver such as [34].



$t = t_{max}$	$T_h \in P_2$	$(T_h, \zeta_h) \in P_2 \times dP_1$	$(T_h, \zeta_h) \in P_1 \times P_1$ SUPG	$(T_h, \zeta_h) \in P_2 \times P_2$ SUPG
lower ref.	0.223	0.293	0.840	0.856
higher ref.	0.647	0.676	0.950	0.960

Figure 4.6. Evolution of the relative total available temperature (4.7) over time for tokamak equilibrium test case. Blue, brown, green and purple curves correspond to discretizations (2.7) with polynomial degree $k = 2$, (2.9) with $k = 2$, (3.8) with $k = 1$ and (3.8) with $k = 2$, respectively. The lower and higher refinement levels are denoted by solid and dashed lines, respectively. The table depicts the relative total available temperature at the last time step.

turn leads to an increase in spurious cross-diffusion, thereby creating a cycle whose effect is stronger for methods with a less accurate representation of the directional derivative along magnetic field lines. In particular, we expect this mechanism to also hold true in the case of a more realistic, lower initial background temperature outside of the separatrix.

All four runs improve for the higher resolution, with a reduced loss of approximately 35% and 32% for the standard primal and mixed CG discretizations, respectively. However, this is still significantly larger than the losses reported for the SUPG-based discretizations run at the lower refinement level. For the higher refinement level, the SUPG-based runs lead to a loss of only approximately 5% and 4% respectively. This time, comparing the second order standard mixed and mixed SUPG-modified methods, we obtain an 8-fold decrease in temperature loss, from slightly less than a third down to a few percent. Further, in a run not shown here, the second order primal CG discretization run at the next higher (poloidal) refinement leads to a loss of approximately 17%, which is still higher than the loss of the second order SUPG-based discretization run at the lowest refinement level (i.e. a quarter of the resolution).

Overall, the standard primal and mixed CG discretizations lead to losses that may not be deemed acceptable for magnetohydrodynamic simulations run for dissipative time scales, while the SUPG-modified discretizations — especially the second order one — lead to more reasonable losses as low as 4%. Here, we note that the higher refinement level leads to approximately 30.5k degrees of freedom for the CG space P_2. Extrapolating to a full tokamak mesh (with $8 \times 3 = 24$ layers, i.e. 48 degrees of freedom in the toroidal direction), this corresponds to only $244k$ degrees of freedom for the temperature space. In particular, we expect a) an even better performance of our novel SUPG-based discretization in the case of higher resolutions and a mesh aligned with the initial conditions, and b) an improved performance in the case of more complex dynamical scenarios, given these coarse resolution results with a mesh not aligned with the magnetic flux surfaces, as well as the above magnetic flux surface results.

Remark 4.1 (Higher order discretizations)**.** Next to the first and second polynomial order runs shown in Figure 4.6, we also performed a run with the primal weak formulation (2.7) on the lower refinement mesh, with a CG function space that is second and fourth order in the toroidal and poloidal components, respectively. In particular, this leads to the same number of degrees of freedom as the higher refinement runs of Figure 4.6. In this case, we found a relative total available temperature of approximately 0.901 at the last time step. In other words, the fourth order primal CG discretization leads to a loss of approximately 10% in temperature, compared to only approximately 5% for the lowest polynomial order mixed SUPG-based method at the same resolution.

4.4. **Tokamak perturbation**

Next to equilibrium scenarios for tokamak domains, we can also consider temperature perturbations that lead to a non-steady field evolution. For this purpose, we consider the same toroidally periodic extruded domain and magnetic field as for the equilibrium case. This time, however, we set the initial temperature equal to $T \equiv 0.5$ throughout, and apply a perturbation over time through an axisymmetric forcing term, given by

$$S(t) = S_p(R, Z)S_t(t), \tag{4.8}$$

$$S_t(t) = \left(\tanh\left((t - t_s)/t_\sigma\right) + 1\right)/2 \qquad S_p = \begin{cases} s_m(1 + \cos(f_\rho \pi \rho))/2 & \rho < 1/f_\rho, \\ 0 & \rho > 1/f_\rho, \end{cases} \tag{4.9}$$

where $t_s = 40$, $t_\sigma = 15$ specify the forcing's growth over time, and $s_m = 1$, $f_\rho = 7$, $\rho = \sqrt{(R - R_0)^2 + (Z - Z_0)^2}$ specify its location in space, for RZ-coordinate (R_0, Z_0). In our case, we set (R_0, Z_0) to a particular mesh vertex location at approximately $(2.4, -0.22)$; see Figure 4.7 (noting that we consider a tokamak domain that is non-dimensionalized with length scale $L_0 = 2$). The parallel conductivity is set to a fixed value of $\kappa_\parallel = 10$, and further $\kappa_\perp = 0$. As before, we start with a time step of $\delta t = 10^{-5}$, which this time is increased linearly to $\delta t = 0.02$ over 100 time steps. The simulation is run up to $t_{max} = 120$ Alfven times, using the mesh as depicted in Figure 4.5. Since there is no perpendicular diffusion, the temperature perturbation should be diffused strictly along magnetic field lines which form a magnetic flux tube, and in Figure 4.7, we marked two approximate bounding surfaces, which together form a volume that contains 95% of the perturbation. In the course of the simulation, we expect the perturbation to steadily be diffused throughout the tube it lies in. Results for the standard mixed and SUPG discretizations (2.9), (3.8), respectively, are depicted in Figure 4.7, noting that the primal discretization (2.7) leads to a qualitatively similar result as (2.9).

We find that as expected, the temperature perturbation generated by the forcing term $S(t)$ is spread across the region encompassed by the two aforementioned bounding flux surfaces. For the standard mixed discretization (2.9), we find a significant amount of cross-diffusion, leaking temperature across the two bounding flux surfaces. For the SUPG-based discretization, this is far less so the case, and the temperature is well contained between the bounding surfaces. In particular, both discretizations conserve the total temperature up to boundary conditions[2], and we find a maximum temperature of approximately 0.65 for the SUPG setup, compared to only 0.6 for the standard mixed one. Finally, both methods lead to areas with a temperature less than the background value 0.5; this is likely due to small negative values in the discrete forcing term, and especially the choice of relatively coarse time step and midpoint-rule for the time discretization.

[2]In the absence of boundary conditions, we can set $\gamma \equiv 1$ in (2.9a) and (3.8a) and obtain $\frac{d}{dt}\int_\Omega T_h \, dx = \int_\Omega S \, dx$.

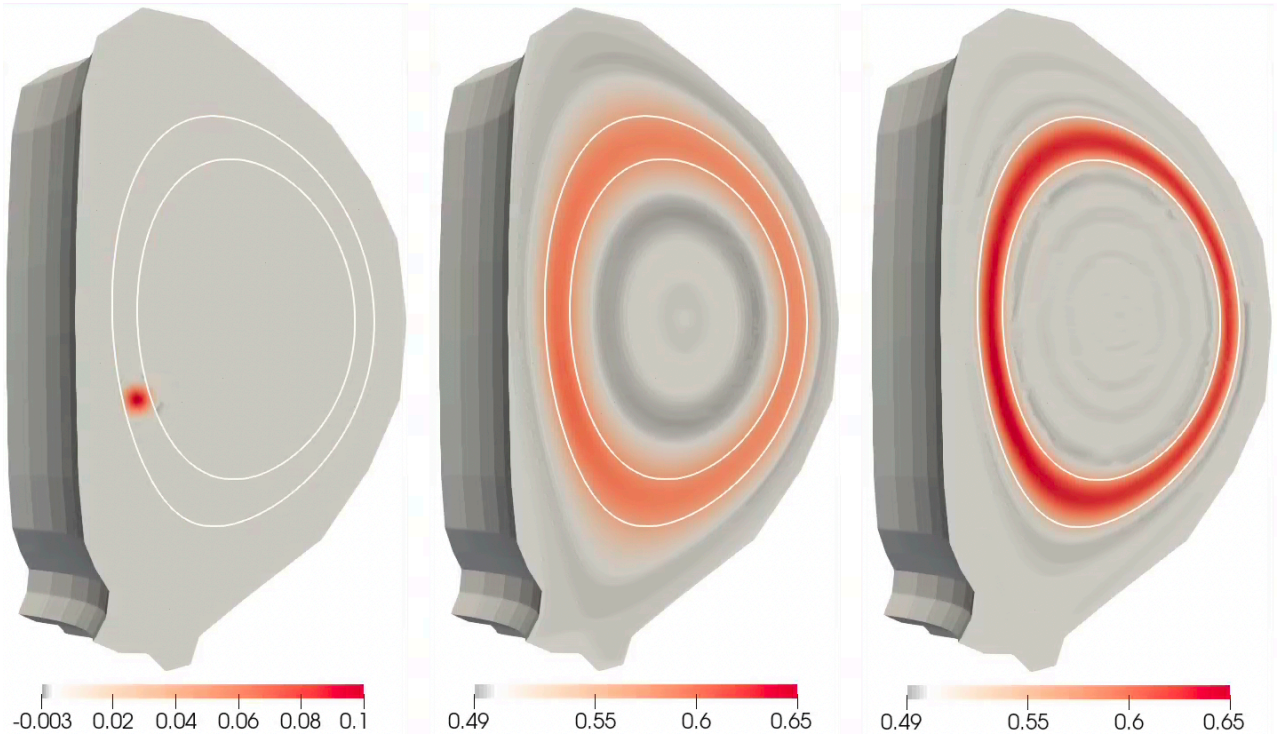


Figure 4.7. Images for perturbation test (4.9), at $t \approx 120$ Alfven times. Left: forcing $S(t)$. Center and right: temperature field for standard mixed and SUPG discretizations (2.9), (3.8), respectively. White lines denote two magnetic flux surfaces forming a magnetic flux tube that contains 95% of the temperature generated by $S(t)$.

5. Conclusion

In this work, we presented a novel CG-based spatial discretization for the anisotropic heat conduction equation in the context of magnetic confinement fusion. It contains an auxiliary variable that serves as a form of the temperature's directional derivative, and further includes an SUPG-stabilization method to reduce spurious heat flux across magnetic field lines. We proved the method's consistency with respect to strong solutions of the anisotropic heat flux equation, and further motivated and discussed our choice of SUPG formulation, leading to a parabolic character of the discretization up to consistency-related modifications.

In numerical tests, we verified the spatial discretization's order of accuracy. Further, we considered a test case representing the spread of a temperature perturbation along a 2D magnetic flux surface with rational winding number, demonstrating a superior performance of our novel method when compared to existing CG-based discretizations. Finally, we considered two realistic tokamak scenarios representing a magnetohydrodynamic equilibrium and a perturbation spreading inside a magnetic flux tube, demonstrating a comparably far smaller spurious cooling rate and a more confined temperature spread, respectively, at time scales of interest for resistive magnetohydrodynamic simulations.

Possible future extensions to this work include studies in the context of anisotropic heat flux as part of a magnetohydrodynamic model, examining the discretization's accuracy when the temperature and magnetic fields are subject to fluid flow. Further, we are working on efficient solvers for our novel discretization in conductivity and time step regimes typically considered for tokamak simulations.

Appendix A. Neumann boundary conditions

Magnetohydrodynamic simulations for magnetic confinement fusion typically consider Dirichlet boundary conditions for the temperature along the reactor wall, which are set to a fixed low value corresponding to the temperature in the near vacuum region outside the separatrix (see Figure 4.5). On the other hand, when Neumann boundary conditions are considered, this is usually done relative to the magnetic field $\mathbf{B}$, while assuming a negligible perpendicular heat flux $\kappa_\perp|_{\partial\Omega} \approx 0$ at the boundary. This fits naturally with our scheme since it is based on a separation between isotropic and purely anisotropic heat flux components. Assuming a Neumann boundary condition of the form $\mathbf{b}\cdot\nabla T = g$, and further assuming a negligible perpendicular heat flux at the boundary, the SUPG-modified, CG-based mixed discretization is given by finding $(T_h, \zeta_h) \in \mathbb{V}_k^{\mathrm{CG}} \times \mathbb{V}_k^{\mathrm{CG}}$ such that

$$M^a\left(\gamma, \frac{\partial T_h}{\partial t}\right) + G_\parallel^f(\gamma, \zeta_h) - \int_{\partial\Omega} \gamma(\mathbf{s}\cdot\mathbf{n})(\zeta_h + \mathbf{s}\cdot\nabla(\tau\zeta_h))\mathrm{d}S + \langle\nabla\gamma, \kappa_\perp\nabla T_h\rangle - \langle\tau\mathbf{s}\cdot\nabla\gamma, \nabla\cdot(\kappa_\perp\nabla T_h)\rangle = M^a(\gamma, S) \qquad \forall\, \gamma \in \mathbb{V}_k^{CG}, \tag{A.1a}$$

$$M^f(\eta, \zeta_h) - G_\parallel^f(T_h, \eta) + \int_{\partial\Omega} \tau\eta\kappa_\Delta g(\mathbf{n}\cdot\mathbf{b})\mathrm{d}S = 0 \qquad \forall\, \eta \in \mathbb{V}_k^{CG}. \tag{A.1b}$$

In practice, in the context of magnetic confinement fusion, Neumann boundary conditions may be considered e.g., when only parts of the plasma are simulated, with the boundary given partially by a fixed magnetic flux surface $\partial\Omega_p$. This then leads to homogeneous Neumann Boundary conditions on $\partial\Omega_p$ with $\kappa_\perp = 0$ and $\mathbf{B}\cdot\mathbf{n}|_{\partial\Omega_p} = 0$. Finally, this may be paired with the aforementioned Dirichlet boundary conditions at the domain's outer boundary $\partial\Omega_w$ corresponding to the reactor wall. In our case, this setting can be achieved by considering test and trial function spaces $\mathbb{V}_{k,bc}^{\mathrm{CG}}$, $\mathring{\mathbb{V}}_k^{CG}$, respectively, that are adjusted for Dirichlet boundary conditions along $\partial\Omega_w$ (but not $\partial\Omega_p$). The resulting scheme reads as (3.8) with $\partial\Omega$ replaced by $\partial\Omega_w$, noting that any boundary integrals at the inner, Neumann boundary condition-type boundary $\partial\Omega_p$ drop out since $\mathbf{s}\cdot\mathbf{n}|_{\partial\Omega_p} = (\sqrt{\kappa}_\Delta/\mathbf{B})(\mathbf{B}\cdot\mathbf{n})|_{\partial\Omega_p} = 0$.

Appendix B. Non-dimensionalization

To derive a non-dimensionalized form of the anisotropic heat flux equation that is appropriate for the tokamak test cases considered in the numerical results section 4, we consider parameter choices common for magnetohydrodynamic models. In such models the ion temperature equation may be given by [24]

$$\frac{n_i}{\gamma - 1}\left(\frac{\partial T_i}{\partial t} + \mathbf{V}_i\cdot\nabla T_i\right) + n_i T_i \nabla\cdot\mathbf{V}_i - \nabla\cdot(\kappa_{i_\parallel}\nabla_\parallel T_i + \kappa_{i_\perp}\nabla_\perp T_i) = S, \tag{B.1}$$

for ion temperature T_i, ion density n_i, adiabatic index $\gamma = \frac{5}{3}$, ion velocity $\mathbf{V}_i$, ion conductivities $\kappa_{i_\parallel}$, $\kappa_{i_\perp}$, and forcing term S containing additional contributions such as Ohmic heating, heat exchange, radiative cooling and external sources. The conductivites can be described by the Branginskii plasma model [6], which leads to

$$\kappa_{i_\parallel} = \frac{3.9 n_i T_i \tau_i}{m_i}, \quad \kappa_{i_\perp} = \frac{2 n_i T_i}{m_i \omega_i^2 \tau_i}. \tag{B.2}$$

The additional quantities in (B.2) denote the ion mass m_i, ion collisional time τ_i and ion gyrofrequency ω_i, respectively. The latter two in turn are defined by

$$\tau_i = \frac{12\pi^{3/2}\epsilon_0^2\sqrt{m_i}T^{3/2}}{n_i Z^4 e^4 \ln\Lambda}\,\mathrm{s}, \quad \omega_i = \frac{Ze|\mathbf{B}|}{m_i}\,\mathrm{s}^{-1}, \tag{B.3}$$

for vacuum permittivity ϵ_0, elementary charge e, Coulomb logarithm $\ln\Lambda$, and atomic number Z. For our numerical results in Section 4, we consider the following parameters and physical constants:

$$\begin{aligned} &Z = 1, && \ln\Lambda \approx 15, && m_i = 1.673 \times 10^{-27}\,\mathrm{kg}, \\ &\epsilon_0 = 8.854 \times 10^{-12}\,\mathrm{Fm}^{-1}, && e = 1.6 \times 10^{-19}\,\mathrm{C}, && \mu_0 = 4\pi \times 10^{-7}\,\mathrm{NA}^{-2}, \end{aligned} \tag{B.4}$$

where μ_0 denotes the vacuum permeability and will be used further below.

We non-dimensionalize (B.1) with respect to typical MHD equilibrium quantities in tokamaks, using a plasma chamber with a prescribed minor radius L_0 and a typical ion density number n_0. Further, we consider a magnetic field strength and plasma pressure which are given by B_0 and p_0, respectively, at the magnetic axis, as computed by a Grad–Shafranov equilibrium code. For the equilibrium used in the tokamak test case in Section 4, the quantities are given by

$$p_0 = 656\,\mathrm{kPa}, \quad B_0 = 5.42\,\mathrm{T}, \quad n_0 = 10^{20}\,\mathrm{m}^{-3}, \quad L_0 = 2\,\mathrm{m}, \quad t_A = 1.83 \times 10^{-7}\,\mathrm{s}, \tag{B.5}$$

where $t_A = L_0/v_A$ denotes the characteristic Alfven wave time scale, for Alfven wave speed $v_A = B_0/\sqrt{\mu_0 n_0 m_i}$. These are representative values for an ITER discharge. In the following, we will assume for simplicity charge neutrality and further the ion and electron temperature to be approximately equal, so that the total plasma pressure at the magnetic axis can be written in terms of the ion temperature T_0 according to

$$p_0 \approx (1+Z) n_0 T_0 = 2 n_0 T_0 \quad \Longrightarrow \quad T_0 = 3.28 \times 10^{-15}\,\mathrm{J}, \tag{B.6}$$

which corresponds to approximately 20.5 keV at magnetic axis. Note that as in the formulas above, T_0 is given in Joule (with $T_0 = k_B T_K$, for Boltzmann constant k_B and temperature in Kelvin T_K). Given these choices, the above quantities (B.3), (B.2) evaluate to

$$\tau_i = \frac{12\pi^{3/2}\epsilon_0^2\sqrt{m_i}T_0^{3/2}}{n_0 Z^4 e^4 \ln\Lambda}\widetilde{n}^{-1}\widetilde{T}^{3/2} = \tau_{i_0}\widetilde{n}^{-1}\widetilde{T}^{3/2}, \qquad \omega_i = \frac{Ze|\mathbf{B}_0|}{m_i}|\widetilde{\mathbf{B}}| = \omega_{i_0}|\widetilde{\mathbf{B}}|, \tag{B.7a}$$

$$\kappa_{i_\parallel} = \frac{3.9 n_0 T_0 \tau_{i_0}}{m_i}\widetilde{T}^{3/2}\kappa_{i_{\parallel 0}}\widetilde{T}^{3/2}, \qquad \kappa_{i_\perp} = \frac{2 n_0 T_0}{m_i \omega_{i_0}^2 \tau_{i_0}}\widetilde{n}^2\widetilde{T}^{-1/2}|\widetilde{\mathbf{B}}|^{-2} = \kappa_{i_{\perp 0}}, \tag{B.7b}$$

where tildes denote non-dimensionalized quantities. Further, the system's characteristic collisional time, gyrofrequency and conductivities, which depend on physical constants as well as our choice of Z, $\ln\Lambda$, B_0, T_0 and n_0, evaluate to

$$\tau_{i_0} \approx 4.1 \times 10^{-2}\,\mathrm{s}, \qquad \omega_{i_0} \approx 5.2 \times 10^{8}\,\mathrm{s}^{-1}, \tag{B.8a}$$

$$n_0^{-1}\kappa_{i_{\parallel 0}} \approx 3.13 \times 10^{11}\,\mathrm{m^2 s^{-1}}, \qquad n_0^{-1}\kappa_{i_{\perp 0}} \approx 3.56 \times 10^{-4}\,\mathrm{m^2 s^{-1}}. \tag{B.8b}$$

Altogether, the non-dimensionalized form of the ion temperature equation (B.1) reads

$$\frac{\widetilde{n} n_0}{\gamma - 1}\frac{T_0}{t_A}\frac{\partial \widetilde{T}}{\partial \widetilde{t}} - \frac{1}{L_0}\widetilde{\nabla}\cdot\left(\kappa_{i_{\parallel 0}}\,\widetilde{T}^{5/2}\frac{1}{L_0}\widetilde{\nabla}_\parallel(T_0\widetilde{T}) + \kappa_{i_{\perp 0}}\,\widetilde{n}^2|\widetilde{\mathbf{B}}|^{-2}\widetilde{T}^{-1/2}\frac{1}{L_0}\widetilde{\nabla}_\perp(T_0\widetilde{T})\right) = \widetilde{S}, \tag{B.9}$$

noting that $\mathbf{b}$ in the directional gradients is transformed to $\mathbf{b} = \widetilde{\mathbf{B}}/|\widetilde{\mathbf{B}}|$. Multiplying (B.9) by $(\gamma-1)t_A/(T_0 n_0)$, we then obtain

$$\widetilde{n}\frac{\partial \widetilde{T}}{\partial \widetilde{t}} - \widetilde{\nabla}\cdot\left(K_\parallel\widetilde{\kappa}_\parallel\,\widetilde{\nabla}_\parallel\widetilde{T} + K_\perp\widetilde{\kappa}_\perp\,\widetilde{\nabla}_\perp\widetilde{T}\right) = \frac{(\gamma-1)t_A}{T_0 n_0}\widetilde{S}, \tag{B.10}$$

for conductivity parameters

$$K_\parallel = (\gamma-1)\frac{t_A}{L_0^2 n_0/\kappa_{i_{\parallel 0}}} = (\gamma-1)\frac{t_A}{t_\parallel}, \qquad \widetilde{\kappa}_\parallel = \widetilde{T}^{5/2}, \tag{B.11}$$

$$K_\perp = (\gamma-1)\frac{t_A}{L_0^2 n_0/\kappa_{i_{\perp 0}}} = (\gamma-1)\frac{t_A}{t_\perp}, \qquad \widetilde{\kappa}_\perp = \widetilde{n}^2|\widetilde{\mathbf{B}}|^{-2}\widetilde{T}^{-1/2}, \tag{B.12}$$

where $t_\parallel$ and $t_\perp$ denote diffusive time scales. In this work, we assume for simplicity a constant density field $\widetilde{n} = 1$, leading to the non-dimensionalized anisotropic heat flux equation (2.1), with $\kappa_\parallel = K_\parallel \widetilde{\kappa}_\parallel$ and $\kappa_\perp = K_\perp \widetilde{\kappa}_\perp$.

For our choice of spacial and temporal scales L_0, t_A, respectively, from (B.5), we obtain

$$K_\parallel \approx 8.8 \times 10^3, \quad K_\perp \approx 1 \times 10^{-11}, \tag{B.13}$$

and these will be our conductivity magnitudes at the magnetic axis, since our non-dimensionalization implies $\widetilde{T} = |\widetilde{\mathbf{B}}| = 1$ here. Conversely, we note that the smallest temperature values occur in the region between the separatrix and the plasma facing wall. A typical value in this region is given by $T_b = 10$ eV, corresponding to a factor of $T_b = 4.9 \times 10^{-4}$ of T_0. Additionally, towards the outer wall region (with respect to the radial coordinate r), we have $|\widetilde{\mathbf{B}}| \approx 0.75$. Correspondingly, in this region, the conductivity magnitudes are given by

$$K_\parallel \widetilde{\kappa}_\parallel = K_\parallel T_b^{5/2} \approx 4.7 \times 10^{-5}, \quad K_\perp \widetilde{\kappa}_\perp = K_\perp |\widetilde{\mathbf{B}}|^{-2} T_b^{-1/2} \approx 8 \times 10^{-10}, \tag{B.14}$$

which together with (B.13) implies a parallel to perpendicular anisotropy ratio between approximately 5 and 15 orders of magnitude.

The underlying assumptions for the Braginskii closure [6] of parallel thermal conduction do not hold in the high plasma temperature limit where the magnetic field line length L_c is shorter than the mean free path λ_{mfp}. This is when the formal perturbative expansion of the distribution function in Braginskii's analysis no longer applies because the Knudsen number $Kn \equiv \lambda_{\text{mfp}}/L_c$ is large. A straightforward scaling analysis of Braginskii's parallel thermal conduction flux for electrons has $q_{e\parallel} = \kappa_\parallel \nabla_\parallel T_e \sim n_e v_{th,e} T_e Kn$, with n_e the electron density, $v_{th,e}$ the electron thermal speed, and T_e the electron temperature. In the collisional regime $Kn \ll 1$. When Kn reaches order unity or greater, the normal expectation is that parallel heat flux would be bounded by electron thermal speed, yielding the so-called flux limiting form of $q_{e\parallel} \sim n_e v_{th,e} T_e$. [3] In a range of tokamak plasma problems, more exotic parallel electron conduction physics can appear [20, 27, 43]. They all have the effect of suppressing the parallel heat flux far below the Braginskii values at high plasma temperature. For our purposes, this limiting behavior is mimicked by deploying the limiting function f in (4.6) in the numerical results section. With the limiter in place, the maximum anisotropy ratio reduces to approximately 12 orders of magnitude, which in practice may be further reduced since with the aforementioned limitations, the perpendicular conductivity may also be larger than predicted by the Braginskii model because of plasma micro-turbulence. In general, the limiting case for particle transport is complex, and our choice of limit T_l in (4.6) is somewhat arbitrary and set roughly to the temperature near the separatrix, noting that the separatrix marks a transition zone from long mean free paths and high temperatures to shorter mean free paths and low temperatures. In results not shown here, we also performed a comparison without the use of a limiter (i.e., such that $f(T) = T$ in (4.6)), and found that the standard primal and mixed CG discretizations' accuracy greatly deteriorates. This is much less so the case for the SUPG-based one, leading to a more favorable comparison than the one presented in the numerical results section.

Finally, we note that the time scales of interest in MHD simulations for tokamaks are typically inferred from dissipative dynamics, including $t_{\kappa_\perp}$ and the plasma current decay time scale t_R due to resistivity. After non-dimensionalization, the latter quantity is formulated through the so-called Lundquist number S, which denotes the ratio between t_A and t_R. Depending on the underlying physics assumptions for resistivity, S may lie in the range 10^6 – 10^{12} [24, 33]. Considering the lower end of this range, we therefore have that given a non-dimensionalized unit time scale of 1 Alfven time, simulations of interest may be run up to $t_{max} = 10^6$ Alfven times. In this case, the tokamak test case in Section 4 with its runtime of $t_{max} = 10^5$ Alfven times therefore corresponds to a simulation up to many growth times of the resistive modes [33].

References

[1] P. R. Amestoy, I. S. Duff, and J.-Y. L'Excellent. Multifrontal parallel distributed symmetric and unsymmetric solvers. *Comput. Methods Appl. Mech. Eng.*, 184(2-4):501–520, 2000.

[2] S. Balay, S. Abhyankar, M. Adams, J. Brown, P. Brune, K. Buschelman, L. Dalcin, A. Dener, V. Eijkhout, W. Gropp, et al. PETSc users manual, 2019.

[3] A. R. Bell. Non-Spitzer heat flow in a steadily ablating laser-produced plasma. *Phys. Fluids*, 28(6):2007–2014, 1985.

[4] D. Biskamp. *Nonlinear Magnetohydrodynamics*. Cambridge Monographs on Plasma Physics. Cambridge University Press, 1993.

[5] J. Bonilla, J. N. Shadid, X.-Z. Tang, M. M. Crockatt, P. Ohm, E. G. Phillips, R. P. Pawlowski, S. Conde, and O. Beznosov. On a fully-implicit VMS-stabilized FE formulation for low Mach number compressible resistive MHD with application to MCF. *Comput. Methods Appl. Mech. Eng.*, 417: article no. 116359, 2023.

[6] S. I. Braginskii. Transport processes in a plasma. *Rev. Plasma Phys.*, 1: article no. 205, 1965.

[7] J Breslau, N Ferraro, and S Jardin. Some properties of the M3D-C1 form of the three-dimensional magnetohydrodynamics equations. *Phys. Plasmas*, 16(9): article no. 092503, 2009.

[8] A. N. Brooks and T. J. R. Hughes. Streamline upwind/Petrov-Galerkin formulations for convection dominated flows with particular emphasis on the incompressible Navier-Stokes equations. *Comput. Methods Appl. Mech. Eng.*, 32(1-3):199–259, 1982.

[9] L. Chacón, D. A. Knoll, and J. M. Finn. An implicit, nonlinear reduced resistive MHD solver. *J. Comput. Phys.*, 178(1):15–36, 2002.

[10] A. S. Chamarthi, H. Nishikawa, and K. Komurasaki. First order hyperbolic approach for Anisotropic Diffusion equation. *J. Comput. Phys.*, 396:243–263, 2019.

[11] P. Degond, A. Lozinski, J. Narski, and C. Negulescu. An asymptotic-preserving method for highly anisotropic elliptic equations based on a micro–macro decomposition. *J. Comput. Phys.*, 231(7):2724–2740, 2012.

[12] F. Deluzet and J. Narski. A two field iterated asymptotic-preserving method for highly anisotropic elliptic equations. *Multiscale Model. Simul.*, 17(1):434–459, 2019.

[13] B. D. Dudson et al. BOUT++: Recent and current developments. *J. Plasma Phys.*, 81(1): article no. 365810104, 2015.

[14] J. P. Freidberg. *Ideal Magnetohydrodynamics*. Plenum Press, 1987.

[15] G. Giorgiani, H. Bufferand, F. Schwander, E. Serre, and P. Tamain. A high-order non field-aligned approach for the discretization of strongly anisotropic diffusion operators in magnetic fusion. *Comput. Phys. Commun.*, 254: article no. 107375, 2020.

[16] D. Green, X. Hu, J. Lore, L. Mu, and M. L. Stowell. An efficient high-order numerical solver for diffusion equations with strong anisotropy. *Comput. Phys. Commun.*, 276: article no. 108333, 2022.

[17] D. Green, X. Hu, J. Lore, L. Mu, and M. L. Stowell. An Efficient High-Order Solver for Diffusion Equations with Strong Anisotropy on Non-Anisotropy-Aligned Meshes. *SIAM J. Sci. Comput.*, 46(2):S199–S222, 2024.

[18] S. Günter, K. Lackner, and C. Tichmann. Finite element and higher order difference formulations for modelling heat transport in magnetised plasmas. *J. Comput. Phys.*, 226(2):2306–2316, 2007.

[19] S. Günter, Q. Yu, J. Krüger, and K. Lackner. Modelling of Heat Transport in Magnetised Plasmas Using Non-Aligned Coordinates. *J. Comput. Phys.*, 209(1):354–370, 2005.

[20] Z. Guo and X.-Z. Tang. Parallel Heat Flux from Low to High Parallel Temperature along a Magnetic Field Line. *Phys. Rev. Lett.*, 108: article no. 165005, 2012.

[21] D. A. Ham, P. H. J. Kelly, L. Mitchell, C. J. Cotter, R. C. Kirby, K. Sagiyama, N. Bouziani, S. Vorderwuelbecke, T. J. Gregory, J. Betteridge, D. R. Shapero, R. W. Nixon-Hill, C. J. Ward, P. E. Farrell, P. D. Brubeck, I. Marsden, T. H. Gibson, M. Homolya, T. Sun, A. T. T. McRae, F. Luporini, A. Gregory, M. Lange, S. W. Funke, F. Rathgeber, G.-T. Bercea, and G. R. Markall. *Firedrake User Manual*. Imperial College London and University of Oxford and Baylor University and University of Washington, first edition edition, 2023.

[22] M. Held, M. Wiesenberger, and A. Stegmeir. Three discontinuous Galerkin schemes for the anisotropic heat conduction equation on non-aligned grids. *Comput. Phys. Commun.*, 199:29–39, 2016.

[23] M. Hoelzl et al. The JOREK non-linear extended MHD code and applications to large-scale instabilities and their control in magnetically confined fusion plasmas. *Nucl. Fusion*, 61(6): article no. 065001, 2021.

[24] S. Jardin. *Computational methods in plasma physics*. CRC Press, 2010.

[25] S. Jin. Efficient asymptotic-preserving (AP) schemes for some multiscale kinetic equations. *SIAM J. Sci. Comput.*, 21(2):441–454, 1999.

[26] D. Kuzmin. *A guide to numerical methods for transport equations*. 2010.

[27] J. Li, Y. Zhang, and X.-Z. Tang. Staged cooling of a fusion-grade plasma in a tokamak thermal quench. *Nucl. Fusion*, 63(6): article no. 066030, 2023.

[28] S. Liu, Q. Tang, and X.-Z. Tang. A Parallel Cut-Cell Algorithm for the Free-Boundary Grad–Shafranov Problem. *SIAM J. Sci. Comput.*, 43(6):B1198–B1225, 2021.

[29] T. A. Manteuffel, S. Münzenmaier, J. Ruge, and B. S. Southworth. Nonsymmetric reduction-based algebraic multigrid. *SIAM J. Sci. Comput.*, 41(5):S242–S268, 2019.

[30] T. A. Manteuffel, J. Ruge, and B. S. Southworth. Nonsymmetric algebraic multigrid based on local approximate ideal restriction (lAIR). *SIAM J. Sci. Comput.*, 40(6):A4105–A4130, 2018.

[31] A. T. T. McRae, G.-T. Bercea, L. Mitchell, D. A. Ham, and C. J. Cotter. Automated generation and symbolic manipulation of tensor product finite elements. *SIAM J. Sci. Comput.*, 38(5):S25–S47, 2016.

[32] J. Narski and M. Ottaviani. Asymptotic Preserving scheme for strongly anisotropic parabolic equations for arbitrary anisotropy direction. *Comput. Phys. Commun.*, 185(12):3189–3203, 2014.

[33] W. Park, E. V. Belova, G. Y. Fu, X.-Z. Tang, H. R. Strauss, and L. E. Sugiyama. Plasma simulation studies using multilevel physics models. *Phys. Plasmas*, 6(5):1796–1803, 1999.

[34] D. A. Serino, Q. Tang, X.-Z. Tang, T. V. Kolev, and K. Lipnikov. An adaptive Newton-based free-boundary Grad-Shafranov solver. `https://arxiv.org/abs/2407.03499`, 2024.

[35] B. S. Southworth, S. Walton, S. B. Roberts, and H. Park. Moment-based adaptive time integration for thermal radiation transport. `https://arxiv.org/abs/2501.17301`, 2025.

[36] C. R. Sovinec et al. Nonlinear magnetohydrodynamics simulation using high-order finite elements. *J. Comput. Phys.*, 195(1):355–386, 2004.

[37] T. Tezduyar. Stabilization parameters and local length scales in SUPG and PSPG formulations. In *Proceedings of the Fifth World Congress on Computational Mechanics*, 2002. paper no. 81508.

[38] C. J. Vogl, I. Joseph, and M. Holec. Mesh Refinement for Anisotropic Diffusion in Magnetized Plasmas. `https://arxiv.org/abs/2210.16442`, 2022.

[39] G. A. Wimmer, B. S. Southworth, T. J. Gregory, and X.-Z. Tang. A fast algebraic multigrid solver and accurate discretization for highly anisotropic heat flux I: open field lines. *SIAM J. Sci. Comput.*, 46(3):A1821–A1849, 2024.

[40] G. A. Wimmer, B. S. Southworth, and Q. Tang. A structure-preserving discontinuous Galerkin scheme for the Cahn-Hilliard equation including time adaptivity. *J. Comput. Phys.*, 537: article no. 114097, 2025.

[41] C. Yang, F. Deluzet, and J. Narski. Preserving the accuracy of numerical methods discretizing anisotropic elliptic problems. `https://arxiv.org/abs/1911.11482`, 2019.

[42] C. Yang, F. Deluzet, and Jacek. Narski. On the accuracy of numerical methods for the discretization of anisotropic elliptic problems. *J. Comput. Phys.*, 521: article no. 113568 (22 pages), 2025.

[43] Y. Zhang, J. Li, and X.-Z. Tang. Cooling flow regime of a plasma thermal quench. *Europhys. Lett.*, 141(5): article no. 54002, 2023.