arXiv:2105.08765v4 [math.NA] 14 Apr 2026

Moving Mesh with Streamline Upwind Petrov-Galerkin (MM-SUPG) Method for Time-dependent Convection-Dominated Convection-Diffusion Problems

Xianping Li* Matthew McCoy†

Time-dependent convection-dominated convection-diffusion problems are considered. We develop a moving mesh streamline upwind Petrov-Galerkin (MM-SUPG) method by combining residual-based SUPG stabilization with a metric-based moving mesh PDE (MM-PDE) approach. The key feature of the method is the interaction between mesh adaptation and stabilization: the evolving mesh modifies both the direction and magnitude of the SUPG stabilization through the local element geometry, while stabilization improves robustness in convection-dominated regimes. For isotropic diffusion, numerical results show that the proposed method reduces spurious oscillations and provides improved resolution of sharp layers compared with fixed mesh methods, while yielding accuracy comparable to moving mesh finite element methods without SUPG in some cases. For anisotropic diffusion, we introduce a weighted tensor that incorporates both the diffusion tensor and the convection field, and construct a metric tensor via intersection to guide mesh adaptation. Under suitable assumptions on the elementwise tensors and mesh geometry, we establish sufficient conditions for the discrete maximum principle (DMP) of the fully discrete scheme. The analysis is based on quantitative bounds for the convection terms, and requires structural conditions on the alignment between the velocity field and the diffusion tensor. Numerical experiments demonstrate that the proposed metric improves monotonicity properties and reduces undershoots, while maintaining overall accuracy.

1 Introduction

We consider the time-dependent convection-diffusion equation defined over the time-independent domain $\Omega \subset \mathbb{R}^2$ with a Lipschitz boundary and denote $\Omega_T = \Omega \times [0, T]$. The equation is given by:

*College of Integrative Sciences and Arts, Arizona State University, AZ, 85212, USA (*xianping@asu.edu*)
†Valorem Reply, Kansas City, MO, 64157, USA (*matthewmccoy6@gmail.com*)

$$u_t - \varepsilon \nabla \cdot (\mathbb{D} \nabla u) + \boldsymbol{b}(\boldsymbol{x}, t) \cdot \nabla u = f(\boldsymbol{x}, t) \qquad \text{in} \qquad \Omega \times (0, T], \tag{1.1}$$

$$u(\boldsymbol{x}, t) = g(\boldsymbol{x}, t) \qquad \text{on} \qquad \partial\Omega \times [0, T], \tag{1.2}$$

$$u(\boldsymbol{x}, 0) = u_0(\boldsymbol{x}) \qquad \text{in} \qquad \Omega \times \{t = 0\}. \tag{1.3}$$

Here, $\mathbb{D} = \mathbb{D}(\boldsymbol{x})$ is the diffusivity tensor that is symmetric and positive-definite (SPD), $\boldsymbol{b}$ represents an incompressible velocity field, while g, u_0, and f are sufficiently smooth functions. The parameter $\varepsilon \in (0, 1]$ is the scaling factor for the diffusivity coefficient. When $\mathbb{D}$ has the same eigenvalues, the problem is isotropic. In the anisotropic case, the eigenvalues of $\mathbb{D}$ are different, which indicates that diffusion in one direction is faster than in the other.

Our focus lies in the convection-dominated scenario where ε is much smaller than $\|\boldsymbol{b}\|_{L^\infty}$. Such problems arise in various fields, including pollutant dispersal in rivers, atmospheric pollution, the Stefan problem, turbulent transport, and are particularly intriguing in the investigation of the Navier–Stokes equations with large Reynolds numbers, see [1] for relevant applications. When ε is small, non-physical oscillations tend to arise in the solution unless an extremely fine spatial-temporal discretization is employed, incurring substantial computational costs. To enhance stability, artificial diffusion in the direction of $\boldsymbol{b}$ is proposed in [2] to be added to the standard Galerkin finite element method.

The streamline upwind Petrov-Galerkin (SUPG) method has proven to be particular effective in numerically addressing boundary and interior layer problems in both time-dependent and time-independent scenarios [3, 4, 5, 6]. The motivation behind SUPG is to employ upwinding of the test functions along the streamlines, aiming to enhance the diffusion of solution. This concept is most clearly illustrated in the 1D time-independent case using finite difference stencils:

$$\frac{\partial u}{\partial x} = -\varepsilon \frac{\partial^2 u}{\partial x^2}. \tag{2}$$

When applying the finite difference method with only the center stencil, one obtains a numerical approximation that, depending on ε, tends to be under-diffusive. Conversely, utilizing only an upwind stencil leads to excessive diffusion, known as over-diffusion. As the exact solution typically lies somewhere in between these extremes, it prompts the adoption of upwinding the center scheme [7]. This method was initially proposed by Brooks and Hughes for incompressible Navier-Stokes problems [2], and has since proven to be fruitful in studying those equations [8, 9]. Comparison among different methods for the steady-state form of equation (1) have been performed in [10].

Efforts have been made to numerically solve equation (1.1) for time-independent solutions, particularly focusing on steady flows [3, 11], by employing adaptive mesh methods in conjunction with SUPG, especially when $\varepsilon \ll 1$. Notably, mesh adaptation has emerged as a potent tool in addressing convection-dominated problems. Our investigation targets a specific class of problems characterized by two key properties: artificial (non-physical) oscillations and sharp layers, encompassing both interior and boundary layers. Consequently, our interest lies in devising a mesh strategy grounded in anisotropic mesh adaptation techniques.

Several notable methodologies have been explored for solving equation (1). For instance, in [12], Lagrange-Galerkin error analysis demonstrated first-order uniform convergence in ε. Additionally, ε-uniform convergence theory for finite element method with homogeneous boundary conditions was investigated in [13]. Carey et al. introduced adaptive mesh refinement (AMR) in [14], while Chinchapatnam et al. explored unsymmetric and symmetric meshless collocation techniques with radial basis functions in [15]. Of particular interest are the so-called SOLD (Spurious Oscillations at Layers Diminishing) schemes, which have been extensively studied in [16, 17, 18]. These schemes involve the

addition of either isotropic or anisotropic artificial diffusion to the Galerkin formulation. Furthermore, in [19], local discontinuous Galerkin (LDG) methods were employed on a class of layer-adapted Shishkin and Bakhvalov-type meshes, accompanied by error analysis, to approximate equation (1).

For time-dependent problems, Wang et al. utilized the singular boundary method with $f = 0$ in [20], employing an exponential variable transform along with the Fourier transform. Additionally, unfitted finite element methods have gained traction for tackling equation (1) in time-dependent domains [21]. These types of unfitted FEMs are known by various names such as XFEM, CutFEM, Finite Cell Method, and TraceFEM (refer to the sources in [21] along with [22, 23]). It is worth noting, as mentioned in [24], that unfitted techniques may encounter conditioning difficulties, which are addressed using specialized finite element spaces and aggregation technique.

Recent developments include variational multiscale stabilized formulations combined with adaptive mesh strategies (e.g., [25]), as well as residual-based stabilization on anisotropic meshes. Compared with these approaches, the present work focuses on integrating SUPG stabilization with a metric-based moving mesh PDE framework, with particular emphasis on discrete maximum principle (DMP) preservation for anisotropic diffusion.

In all the above mentioned methods, only isotropic diffusivity tensors were considered. However, in many practical applications, the diffusivity tensor is anisotropic. When the mesh is not aligned in the principal diffusion direction, non-physical oscillations may occur when using standard computational methods [26]. Special care must be taken to ensure that the numerical solution satisfies the discrete maximum principle (DMP). For problem with only diffusion, metric tensors has been developed in [26] for time-independent problems to adapt the mesh such as $\mathbb{M}_{DMP} = \theta_K \mathbb{D}_K^{-1}$ based on DMP satisfaction and $\mathbb{M}_{DMP+adap} = C_K(det(\mathbb{D}))^{\frac{1}{d}}\mathbb{D}_K^{-1}$ based on DMP satisfaction as well as minimization of interpolation errors. For time-dependent problems, condition for time step size is also developed in [27] to ensure that the numerical solutions satisfy DMP. Lu et al. investigated the satisfaction of DMP for anisotropic diffusion-convection-advection problems in [28]. However, it is even more challenge for numerical solutions of convection-dominated convection-diffusion problems to satisfy the DMP, which will be addressed in this paper.

In this paper, we propose to integrate the moving mesh partial differential equation (MMPDE) method with streamline upwind Petro-Galerkin (SUPG) method and call it as moving mesh SUPG (MM-SUPG) method. It works for both time-independent and time-dependent flows characterized by initial sharp layers, and can deal with both isotropic and anisotropic diffusivity tensors. It has been observed that, even for isotropic diffusivity, oscillations can persist with adapted meshes [29]. The challenge in numerically solving equation (1) becomes evident particularly when $\varepsilon \ll 1$ and the initial profile experiences convection over time or when the profile features sharp regular, parabolic, corner, or interior layers that develop or exist initially. Meanwhile, when $\mathbb{D}$ is anisotropic, non-physical solutions may appear in the numerical solution, which is a violation of DMP [26, 27]. To address this issue, we introduce a metric tensor to control the mesh movement and develop conditions for mesh and time step size so that the numerical solution satisfies DMP.

The novelty of the proposed MM-SUPG method lies not merely in combining moving mesh adaptation with SUPG stabilization, but in analyzing and exploiting their interaction. The effectiveness of SUPG stabilization depends strongly on mesh geometry, as both the streamline derivative $\boldsymbol{b}\cdot\nabla u_h$ and the stabilization parameter τ_K are mesh-dependent. Meanwhile, the moving mesh method dynamically aligns elements according to solution features and anisotropic diffusion. This coupling leads to a nontrivial interplay: mesh alignment influences the direction and strength of SUPG stabilization, which in turn affects oscillation control and monotonicity. In particular, for anisotropic diffusion, the proposed metric tensor construction balances alignment with both the convection field and diffu-

sion tensor, enabling improved discrete maximum principle behavior. This interaction between mesh adaptation and stabilization is a central contribution of the present work.

The remainder of this paper is organized as follows. In Section 2, we give a brief introduction for SUPG and MMPDE, and propose our MM-SUPG method. Section 3 presents numerical examples for isotropic diffusivity case to demonstrate the effectiveness of our approach, while also delineating the constraints of the moving mesh method. In Section 4, we investigate the satisfaction of discrete maximum principle for anisotropic diffusion. Finally, a summary and some conclusions are presented in Section 5.

2 Moving Mesh with Streamline Upwind Petrov-Galerkin (MM-SUPG) Method

Firstly, we give a brief introduction to SUPG method and MMPDE method. Then we explain in details our MM-SUPG method, the integration of MMPDE with SUPG.

2.1 SUPG method

Consider the convection-diffusion equation (1.1) with homogeneous Dirichlet boundary conditions (1.2) and initial condition (1.3). The SUPG method is a stabilized finite element method that introduces artificial diffusion to the Galerkin formulation. Let $L^2(\Omega)$ and $H^k(\Omega)$, $k \geq 1$ denote the standard Lebesgue and Sobolev spaces, respectively. Let $H_g^1(\Omega) = \{v \in H^1(\Omega),\, v|_{\partial\Omega} = g\}$ and $H_0^1(\Omega) = \{v \in H^1(\Omega),\, v|_{\partial\Omega} = 0\}$. The weak formulation of (1.1) is to find $u \in H_g^1(\Omega)$ such that

$$\int_\Omega u_t v - \varepsilon \int_\Omega \nabla \cdot (\mathbb{D}\nabla u)v + \int_\Omega (\boldsymbol{b} \cdot \nabla u)v = \int_\Omega fv, \quad \forall v \in H_0^1(\Omega). \tag{3}$$

We consider an affine family of simplicial triangulations $\{\mathcal{T}_h\}$ for Ω and denote the piecewise linear finite element space as $V^h = (v \in H^1(\Omega),\, v|_K \in \mathcal{P}_1^K, \forall K \in \mathcal{T}_h)$, where $\mathcal{P}_1^K$ is the linear polynomial space in element K. Denote by $U^h = V^h \cap H_g^1(\Omega)$ and $U_0^h = V^h \cap H_0^1(\Omega)$. The standard Galerkin method is to find $u_h \in U^h$ such that

$$\int_\Omega (u_h)_t v_h - \varepsilon \sum_{K\in\mathcal{T}_h} \nabla \cdot (\mathbb{D}_K \nabla u_h)v_h + \int_\Omega (\boldsymbol{b} \cdot \nabla u_h)v_h = \int_\Omega fv_h, \quad \forall v_h \in U_0^h, \tag{4}$$

where $\mathbb{D}_K$ is the restriction of $\mathbb{D}$ to element K.

The essential idea of SUPG involves altering the test function space to operate within the realm of functions that upwind the trial functions along the streamlines. This can be represented as the space:

$$U_h^{\mathrm{SUPG}} = \left\{ v_h + \tau_K \boldsymbol{b} \cdot \nabla v_h \mid v_h \in U_0^h \right\}, \tag{5}$$

where τ_K is the stabilization parameter. Therefore, one may perceive the space U_h^{SUPG} as a suitable stabilized test space for singularly perturbed problems. The weak formulation for SUPG method is given as follows: find $u_h \in U^h(\Omega)$ such that

$$\begin{aligned}\int_\Omega (u_h)_t v_h - \varepsilon \sum_{K\in\mathcal{T}_h} \nabla \cdot (\mathbb{D}_K \nabla u_h)(v_h + \tau_K \boldsymbol{b} \cdot \nabla v_h) \\ + \sum_{K\in\mathcal{T}_h} (\boldsymbol{b} \cdot \nabla u_h)(v_h + \tau_K \boldsymbol{b} \cdot \nabla v_h) = \int_\Omega f(v_h + \tau_K \boldsymbol{b} \cdot \nabla v_h), \quad \forall v_h \in U_0^h.\end{aligned} \tag{6}$$

The stabilization parameter τ_K is chosen following classical SUPG formulations, where it scales with the local element size and inverse flow magnitude to balance numerical diffusion and consistency; see [2, 8, 30], and related developments in moving mesh contexts [3]. A conventional choice of τ_K on mesh $\mathcal{T}_h$ is

$$\tau_{\mathcal{T}_h}^K = \begin{cases} c_0 \text{diam}(K)_{\mathcal{T}_h} & \text{if } \text{Pe}_K > 1 \text{ convection-dominated} \\ c_1 \text{diam}(K)_{\mathcal{T}_h}^2 & \text{if } \text{Pe}_K \leq 1 \text{ diffusion-dominated} \end{cases}, \tag{7}$$

where Pe_K denotes the element-wise Péclet number given by

$$\text{Pe}_K = \frac{\|\boldsymbol{b}\|_\infty \text{diam}(K)_{\mathcal{T}_h}}{2\varepsilon}. \tag{8}$$

τ_K determines local convection-dominated or diffusion-dominated regimes on the current mesh $\mathcal{T}_K$. There is flexibility in defining $\text{diam}(K)_{\mathcal{T}_h}$ and in selecting c_0 and c_1. This paper focuses on convection-dominated regimes characterized by large element Péclet numbers $P_{e_K} \gg 1$, which may arise due to small ε and/or insufficient mesh resolution relative to the flow magnitude.. Typically, individual bilinear forms in the SUPG formulation are estimated, and $\tau_{\mathcal{T}_h}^K$ is derived based on those errors. In [3], such a method was employed using error estimates in the discrete norm for time independent $\boldsymbol{b}$

$$\|w\|_{\varepsilon,k}^2 = \varepsilon \|\nabla w\|^2 + \sum_{K \in \mathcal{T}_K} \tau_{\mathcal{T}_h}^K \|\boldsymbol{b} \cdot \nabla w\|_{L^2(K)}^2, \, \forall w \in H_0^1(\Omega). \tag{9}$$

However, our method relies on different error estimates for defining the metric tensor along with a mesh functional approach. When a reaction/source term is present, stabilization parameters must be modified to account for its contribution to the residual, as discussed in [31, 32, 33, 34].

In the case of anisotropic mesh adaptation, $\text{diam}(K)_{\mathcal{T}_h}$ will be updated with each new mesh as the elements stretch and shrink. In this paper, we define $\tau_{\mathcal{T}_h}^K$ as

$$\tau_{\mathcal{T}_h}^K = \frac{\text{diam}(K)_{\mathcal{T}_h}}{2\|\boldsymbol{b}\|_\infty} \xi(\text{Pe}_K), \tag{10}$$

which is based on the stabilization parameter selected in [11]. Common choices of $\text{diam}(K)_{\mathcal{T}_h}$ include the length of the longest edge of the element K projected onto $\boldsymbol{b}$ and the diameter of K in the direction of $\boldsymbol{b}$. In this paper, for each $K \in \mathcal{T}_h$, we define

$$\text{diam}(K) = \sup_{\boldsymbol{x}_1, \boldsymbol{x}_2 \in K} \|\boldsymbol{x}_1 - \boldsymbol{x}_2\|, \tag{11}$$

and

$$\xi(\text{Pe}_K) = \min\left\{1, \frac{\text{Pe}_K}{3}\right\}. \tag{12}$$

2.2 Moving mesh partial differential equations (MMPDE)

Moving mesh methods are instrumental in addressing domains where the solution exhibits excessively large gradient, such as sharp layers. Traditional methods like fitted operator methods, finite difference methods, finite element methods with uniform or Shishkin-type meshes may not be able to catch the sharp gradient without a hyper-refined mesh or a priori knowledge about the solution. The MMPDE approach automatically concentrates mesh elements near the region where the solution exhibits large

gradient, hence improves the computational efficiency. In this subsection, we give a brief overview of MMPDE. More details can be found in [35].

We consider an affine family of simplicial triangulations $\{\mathcal{T}_h\}$ for Ω. Denote by $(K, \mathcal{P}_1^K, \Sigma_K)$ the finite element space. Let N, N_v, and N_t denote the number of mesh elements, the number of vertices, and the number of time steps, respectively. Let $\mathbb{M} = \mathbb{M}(\boldsymbol{x})$ be a symmetric and uniformly positive definite metric tensor defined on Ω with $c_1\mathbb{I} \leq \mathbb{M}(\boldsymbol{x}) \leq c_2\mathbb{I}$ for all $\boldsymbol{x} \in \Omega$. We take the $\mathbb{M}$-uniform mesh approach that takes a nonuniform mesh and views it as a uniform one in the metric specified by $\mathbb{M}$. Essentially, $\mathbb{M}$ controls the concentration and alignment of the mesh elements.

We consider two affine-equivalent finite elements, one for the reference elements and another for the physical mesh element. There exists an invertible affine mapping between the finite elements, denoted by F_K:

$$(\hat{K}, \mathcal{P}_1^{\hat{K}}, \Sigma_{\hat{K}}) \xrightarrow{F_K} (K, \mathcal{P}_1^K, \Sigma_K). \tag{13}$$

For each element $K \in \mathcal{T}_h$, denote by $F_K : \hat{K} \to K$ the affine mapping between K and the reference element $\hat{K}$ so that $|\hat{K}| = 1$. Denote by $\boldsymbol{x}_j^K$ for $j = 0$, 1, and 2 the vertices of K and $\boldsymbol{\xi}_j^K$ the corresponding vertices of $\hat{K}$. Then the affine map F_K maps $\boldsymbol{\xi}_j^K$ to $\boldsymbol{x}_j^K$, i.e., $F_K(\boldsymbol{\xi}_j^K) = \boldsymbol{x}_j^K$ for $j = 0$, 1, and 2. $\hat{K}$ is called the reference element and belongs to $\hat{\mathcal{T}}_h$, the reference mesh.

Denote by $\mathbb{M}_K$ the volume-average of $\mathbb{M}$ over element K, i.e.,

$$\mathbb{M}_K = \frac{1}{|K|}\int_K \mathbb{M}(\boldsymbol{x})\mathrm{d}\boldsymbol{x},$$

and

$$\sigma_h = \sum_{K\in\mathcal{T}_h}\int_K \sqrt{\det(\mathbb{M}_K)}\mathrm{d}\boldsymbol{x} = \sum_{K\in\mathcal{T}_h}|K|\sqrt{\det(\mathbb{M}_K)}.$$

F_K and $\mathbb{M}$ give rise to the so-called equidistribution and alignment conditions in $\mathbb{R}^2$:

1. equidistribution: $|K|\sqrt{\det(\mathbb{M}_K)} = \dfrac{\sigma_h}{N}$ for all $K \in \mathcal{T}_h$, and
2. alignment: $\dfrac{1}{2}\mathrm{tr}\left(F_K'\mathbb{M}_K(F_K')^T\right) = \sqrt{\det\left(F_K'\mathbb{M}_K(F_K')^T\right)}$ for all $K \in \mathcal{T}_h$,

where F_K' is the Jacobian matrix of F_K. These two completely characterize a non-uniform mesh: the equidistribution condition controls the size and shape of mesh elements, while the alignment condition controls the orientation of mesh elements. The new physical mesh $\mathcal{T}_h$ is found by minimizing the energy functional (also called the mesh functional) I given in discrete form by

$$I(\mathcal{T}_h) = \sum_{K\in\mathcal{T}_h}|K|G(\mathbb{J}_K, \det(\mathbb{J}_K), \mathbb{M}_K), \tag{14}$$

where $\mathbb{J}_K = (F_K')^{-1}$ and

$$\begin{aligned} G(\mathbb{J}_K, \det(\mathbb{J}_K), \mathbb{M}_K) =& \alpha\sqrt{\det(\mathbb{M}_K)}\left(\mathrm{tr}\left(\mathbb{J}_K\mathbb{M}_K\mathbb{J}_K^T\right)\right)^p \\ &+ (1-2\alpha)2^p\sqrt{\det(\mathbb{M}_K)}\left(\frac{\det(\mathbb{J}_K)}{\sqrt{\det(\mathbb{M}_K)}}\right)^p, \end{aligned} \tag{15}$$

with $\alpha \in (0, \frac{1}{2}]$ and $p > 1$ as dimensionaless parameters. From numerical experiments, $\alpha = \frac{1}{3}$ and $p = \frac{3}{2}$ work well for most problems [36].

The moving mesh PDE (MMPDE) method is developed by Huang et al. [35] to find a new physical mesh by minimizing $I(\mathcal{T}_h)$, which is equivalent to the steepest descent method [37]. The MMPDE is a partial differential equation (also called the mesh PDE) that is given by

$$\frac{\mathrm{d}\boldsymbol{x}_i}{\mathrm{d}t} = -\frac{P_i}{\gamma}\left(\frac{\partial I}{\partial \boldsymbol{x}_i}\right)^T,\ i = 1, \ldots, N_v, \tag{16}$$

where γ is a user parameter meant to adjust the time scale of the mesh movement and $P_j = \det\left(\mathbb{M}(x_j)\right)^{\frac{p-1}{2}}$ is chosen so that the MMPDE is invariant under the scaling transformation of $\mathbb{M}$.

The physical PDE and mesh PDE can be solved either simultaneously or alternatively. When the PDEs are solved simultaneously, the mesh can respond quickly to any change in the physical solution. However, this procedure is highly nonlinear. In our work, we employ the alternative approach. The mesh PDE is firstly solved based on the physical solution and the mesh at the current time step. Then mesh nodes are relocated based on the solutions from mesh PDE. Lastly, the physical PDE is solved on the new mesh for the next time step. The alternate procedure allows for flexibility and potential efficiency at each time level. One disadvantage of this approach is a lag in time, which can be addressed by employing multiple iterations on the same time level before advancing to the next. It is always beneficial to start with a initially adapted mesh, especially when sharp layers exist from the beginning.

In this paper, for moving mesh for isotropic diffusion case, we use the metric tensor defined by

$$\mathbb{M} = \det\left(\mathbb{I} + |H_K(u)|\right)^{-\frac{1}{6}}\left(\mathbb{I} + |H_K(u)|\right), \tag{17}$$

where $\mathbb{I}$ is the identity matrix, $H_K(u)$ is the recovered Hessian using least squares fitting to the values of the solution function at the mesh vertices, and $|H_K(u)| = \Lambda \mathrm{diag}(|\lambda_1|, |\lambda_2|)\Lambda^T$ with $\Lambda \mathrm{diag}(\lambda_1, \lambda_2)\Lambda^T$ being the eigen-decomposition of $H_K(u)$. The Hessian is computed $H(u)(\boldsymbol{x}_j) \approx H(q)(\boldsymbol{x}_j)$, where q is the algebraic polynomial determined by least squares fitting. For completeness, some details for the least squares fitting method (see Section 5.3 in [35]) are provided here.

Specifically, if u_j is the approximation of u at node $\boldsymbol{x}_j = (x_j, y_j)$ on mesh $\mathcal{T}_h$. Let $\boldsymbol{x}_{j_i}$ for $i = 1, \ldots, N_j$ be the N_j neighboring vertices near $\boldsymbol{x}_j$ (including $\boldsymbol{x}_j$). Denote by $\boldsymbol{x}_j^c$ the center of these points and define

$$H_j^x = \max_{i=1,\ldots,N_j} |x_{j_i} - x_j^c|, \quad H_j^y = \max_{i=1,\ldots,N_j} |y_{j_i} - y_j^c|. \tag{18}$$

Then the polynomial q is in (Legendre) polynomial form

$$q(x, y) = \sum_{k=0}^{2}\sum_{m=0}^{2-k} a_{k,m} P_k\left(\frac{x - x_j^c}{H_j^x}\right) P_m\left(\frac{y - y_j^c}{H_j^y}\right), \tag{19}$$

and is determined by

$$\min_{a_{0,0},\ldots,a_{2,0}} \sum_{i=1}^{N_j} (q(x_{j_i}, y_{j_i}) - u_{j_i})^2. \tag{20}$$

Then it is clear that $\nabla u(\boldsymbol{x}_j) \approx \nabla q(\boldsymbol{x}_j)$. This metric tensor is optimal in the L^2 norm of linear interpolation error on triangular meshes [37]. Moreover, it is clearly symmetric and one may replace $|H_K(u)|$ with $\alpha^{-1}|H_K(u)|$, where α is called the intensity parameter and helps ensure that $\mathbb{M}$ is positive definite. This monitor function is derived on anisotropic error estimates (see [35], chapter 5, for a more detailed discussion and the derivation of (17) and choice of α).

2.3 Moving mesh with Streamline Upwind Petrov-Galerkin (MM-SUPG) method

Now we apply the MMPDE method to solve the SUPG formulation (6). Write $\boldsymbol{u}_h(\boldsymbol{x},t) = \sum_{j=1}^{N_v} u_j(t)\varphi_j(\boldsymbol{x})$, where $\{\varphi_j\}_{j=1}^{N_v}$ is the set of linear basis functions in U_0^h. Also, since φ_j is a linear function, the element-wise integral average can be defined explicitly

$$\frac{1}{|K|}\int_K \varphi_j\varphi_i \mathrm{d}\boldsymbol{x} = \begin{cases} \dfrac{1}{(d+1)(d+2)} & \text{for} \quad i \neq j \\ \dfrac{2}{(d+1)(d+2)} & \text{for} \quad i = j \end{cases} . \tag{21}$$

Approximate the global integral as

$$\int_\Omega g(\boldsymbol{x},t)\mathrm{d}\boldsymbol{x} = \sum_{K\in\mathcal{T}_h} |K| \int_K g(\boldsymbol{x},t)\mathrm{d}\boldsymbol{x} \tag{22}$$

$$\approx \sum_{K\in\mathcal{T}_h} |K| \sum_{j=1}^{m} \lambda_j g(\boldsymbol{\omega}_j), \tag{23}$$

where $\boldsymbol{\omega}_j$ are the quadrature nodes for element K and λ_j are weights such that $\sum_{j=1}^m \lambda_j = 1$. For any tensor $\mathbb{F}$, we then define the element-wise average by

$$\mathbb{F}_K = \frac{1}{|K|}\int_K \mathbb{F}(\boldsymbol{x})\mathrm{d}\boldsymbol{x} \approx \sum_{j=1}^{N} \lambda_j \mathbb{F}(\boldsymbol{\omega}_j). \tag{24}$$

The semi-discrete matrix formulation for (6) is given by

$$\boldsymbol{M}\boldsymbol{u}_h'(t) + \boldsymbol{A}(t)\boldsymbol{u}_h(t) = \boldsymbol{f}(t), \tag{25}$$

where $\boldsymbol{u} = [u_1, \ldots, u_{N_v}]^T$, $\boldsymbol{M}$, $\boldsymbol{A}$, and $\boldsymbol{f}$ are given by

$$M_{ij} = \sum_{K\in\mathcal{T}_h} (\varphi_j, \varphi_i)_{0,K} + \tau_K(\varphi_j, \boldsymbol{b}\cdot\nabla\varphi_i)_{0,K}, \tag{26}$$

$$\begin{aligned} A_{ij}(t) = \sum_{K\in\mathcal{T}_h} & (\boldsymbol{b}\cdot\nabla\varphi_j, \varphi_i)_{0,K} + |K|(\nabla\varphi_i)^T\varepsilon\mathbb{D}_K\nabla\varphi_j \\ & + \tau_K\left((\boldsymbol{b}\cdot\nabla\varphi_j, \boldsymbol{b}\cdot\nabla\varphi_i)_{0,K} + (\varepsilon\mathbb{D}_K\nabla\varphi_j, \nabla(\boldsymbol{b}\cdot\nabla\varphi_i))_{0,K}\right), \end{aligned} \tag{27}$$

$$f_i(t) = \sum_{K\in\mathcal{T}_h} (f, \varphi_i)_{0,K} + \tau_K(f, \boldsymbol{b}\cdot\nabla\varphi_i)_{0,K}. \tag{28}$$

We use a θ-scheme to discretize in time. For (25), we approximate

$$\boldsymbol{u}_h'(t_m) \approx \frac{\boldsymbol{u}_h(t_{m+1}) - \boldsymbol{u}_h(t_m)}{\Delta t}, \quad m = 1, \ldots, N_t. \tag{29}$$

Then the corresponding θ-scheme is given by

$$\begin{aligned} M\frac{\boldsymbol{u}_h(t_{m+1}) - \boldsymbol{u}_h(t_m)}{\Delta t} + \theta A\boldsymbol{u}_h(t_{m+1}) + (1-\theta)A\boldsymbol{u}_h(t_m) \\ = \theta\boldsymbol{f}(t_{m+1}) + (1-\theta)\boldsymbol{f}(t_m), \quad m = 1, \ldots, N_t. \end{aligned} \tag{30}$$

Grouping the $\boldsymbol{u}_h(t_{m+1})$ terms, we may rewrite (30) as

$$\left(\frac{M}{\Delta t}+\theta A\right)\boldsymbol{u}_h(t_{m+1})=\theta\boldsymbol{f}(t_{m+1})+(1-\theta)\boldsymbol{f}(t_m)+\left(\frac{M}{\Delta t}-(1-\theta)A\right)\boldsymbol{u}_h(t_m), \tag{31}$$

for $m=1,\ldots,N_t$. Hence, we may write a system of ODEs

$$\tilde{\boldsymbol{A}}\boldsymbol{u}_h(t_{m+1})=\tilde{\boldsymbol{f}}^{m+1},\quad m=1,\ldots,N_t, \tag{32}$$

where

$$\tilde{\boldsymbol{A}}=\boldsymbol{M}+\theta\Delta t\boldsymbol{A}, \tag{33}$$

$$\tilde{\boldsymbol{f}}^{m+1}=\theta\Delta t\boldsymbol{f}(t_{m+1})+\Delta t(1-\theta)\boldsymbol{f}(t_m)+(\boldsymbol{M}-\Delta t(1-\theta)\boldsymbol{A})\,\boldsymbol{u}_h(t_m). \tag{34}$$

In this work, we use $\theta=0.5$, which corresponds to the Crank-Nicolson method. Fixed time step is used in the computations. In Section 4 we derive conditions on Δt so that DMP is satisfied.

To generate the adapted mesh at each physical time level, we minimize the discrete meshing functional $I(\mathcal{T}_h)$ defined in (14) by solving the mesh PDE (16). The derivatives of the meshing functional are assembled element-by-element following the MMPDE framework in [35, 37] and the mesh PDE can be written in the form

$$\frac{\mathrm{d}\boldsymbol{x}_i}{\mathrm{d}t}=\frac{P_i}{\gamma}\sum_{K\in\omega_i}|K|\boldsymbol{v}_{i_K}^K,\ i=1,\ldots,N_v, \tag{35}$$

where ω_i is the element patch associated with the i-th vertex, i_K is the local index of vertex i in element K, and $\boldsymbol{v}_{i_K}^K$ denotes the local mesh velocity contributed by element K. Here the local velocities are obtained from the derivatives of $G(\mathbb{J}_K,\det(\mathbb{J}_K),\mathbb{M}_K)$ defined in (15) with respect to $\mathbb{J}_K$ and $\det(\mathbb{J}_K)$, where $\mathbb{J}_K=(F_K')^{-1}$. Boundary vertices are constrained to remain on $\partial\Omega$, thus boundary nodes are allowed to slide along the boundary with corner nodes being fixed.

To obtain the new mesh, we integrate the discretized mesh PDE (35) from the current mesh $\mathcal{T}_h^m$ to the updated mesh $\mathcal{T}_h^{m+1}$. In our implementation, this system of ODEs is solved in pseudo-time using MATLAB *ode15s* solver. After the integration is completed, the new mesh $\mathcal{T}_h^{m+1}$ is obtained directly from the updated vertex coordinates.

A key point is that the matrices $\boldsymbol{M}$ and $\boldsymbol{A}$, and the stabilization parameter τ_K are all assembled on the current physical mesh. Since the mesh changes in time, these quantities are recomputed after each mesh update. In particular, τ_K is evaluated on the updated mesh, so the amount of streamline stabilization adapts dynamically to the local mesh geometry. We emphasize that we do not solve the MMPDE and the physical PDE fully simultaneously. Instead, we use a staggered procedure:

(1) fix the physical solution and update the mesh;

(2) freeze the new mesh and solve the SUPG system for the physical solution.

This decouples the nonlinear mesh movement from the stabilized finite element solution and is considerably simpler and more robust in practice. The algorithm is given as follows.

Algorithm 1: MM-SUPG Method

for $m=1,...,N_t$, do:

$t_{m+1}=t_m+\Delta t$,

(1) Metric construction.
Using $\boldsymbol{u}_h^m$ on $\mathcal{T}_h^m$, construct the metric tensor $\mathbb{M}^m = \mathbb{M}(\boldsymbol{u}_h^n)$.

(2) Mesh movement.
Fixing $\mathbb{M}^m$, solve the MMPDE for the mesh coordinates from $\boldsymbol{x}^m$ to an updated vertex set $\boldsymbol{x}^{m+1}$, yielding the new mesh $\mathcal{T}_h^{m+1}$.

(3) Solution interpolation.
Interpolate $\boldsymbol{u}_h^m$ from $\mathcal{T}_h^m$ onto $\mathcal{T}_h^{m+1}$, denoted by $\tilde{\boldsymbol{u}}_h^m$.

(4) Recompute geometric quantities.
Using $\tilde{\boldsymbol{u}}_h^m$ on $\mathcal{T}_h^{m+1}$, compute $\mathrm{diam}(K)_{\mathcal{T}_h}$, $\tau_{\mathcal{T}_h}^K$ and basis-function gradients.

(5) Assemble SUPG system.
Assemble the mesh-dependent matrices $\boldsymbol{M}^{M+1}$, $\boldsymbol{A}^{m+1}$, and $\boldsymbol{f}^{m+1}$.

(6) Solve for physical solution.
Solve the following linear system for the solution at t_{m+1}.

$$\boldsymbol{M}^{m+1} + \theta\Delta t\boldsymbol{A}^{m+1}\boldsymbol{u}_h^{m+1} = \theta\Delta t\boldsymbol{f}^{m+1} + \Delta t(1-\theta)\boldsymbol{f}^m + \left(\boldsymbol{M}^{m+1} - \Delta t(1-\theta)\boldsymbol{A}^{m+1}\right)\boldsymbol{u}_h^m.$$

The effectiveness of MM-SUPG stems from the complementary roles of its components: moving mesh adaptation reduces interpolation error by aligning elements with solution features, while SUPG stabilization suppresses spurious oscillations along streamlines. Together, they yield improved stability without excessive numerical diffusion.

3 Results for Isotropic Diffusion

Now, we present some numerical examples for isotropic diffusion case where we take the diffusivity tensor $\mathbb{D}$ as an identity matrix. The value of diffusivity is given by ε. We consider two different examples and compare results from four different methods: fixed mesh finite element method (FM-FEM) without SUPG, fixed mesh finite element method with SUPG (FM-SUPG), moving mesh finite element method (MM-FEM) without SUPG (which is the MMPDE method), and our moving mesh with SUPG (MM-SUPG) method. We denote the MMPDE method as MM-FEM for consistency in what follows.

Note that for $\varepsilon = O(1)$, all methods produce similar results, confirming that stabilization and mesh adaptation are primarily beneficial in convection-dominated regimes.

Example 3.1. This example is similar to the simulated example presented in [38]. The domain is $\Omega = (0, 10) \times (-1.25, 1.25)$ and an initial smooth solution is given by

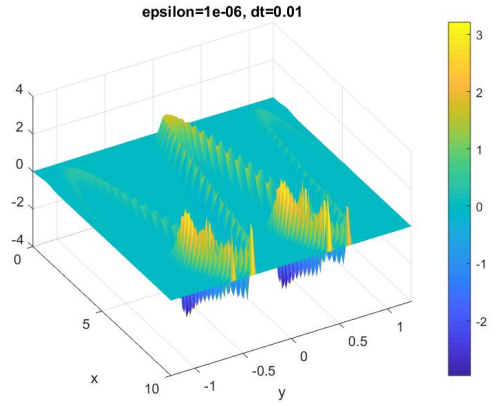
$$u_0(x, y) = \mathrm{e}^{-\frac{(x-1)^2+y^2}{0.25}} + \mathrm{e}^{-\frac{(x-1)^2+(y-0.5)^2}{0.25}} + \mathrm{e}^{-\frac{(x-1)^2+(y+0.5)^2}{0.25}}. \tag{36}$$

The velocity field is given by

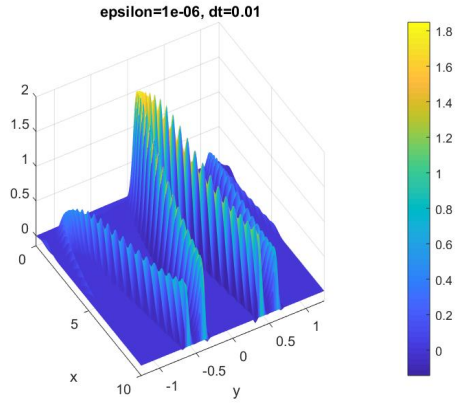
$$\boldsymbol{b} = \left(1 - \exp(\theta x)\cos(2\pi y), \frac{\theta}{2\pi}\exp(\theta x)\sin(2\pi y)\right), \tag{37}$$

with

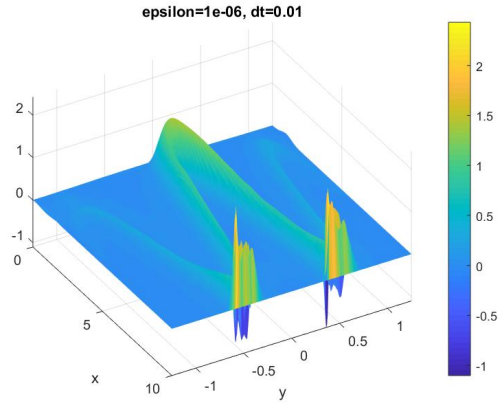
$$\theta = \frac{1}{2\varepsilon} - \sqrt{\left(\frac{1}{2\epsilon}\right)^2 + 4\pi^2}. \tag{38}$$



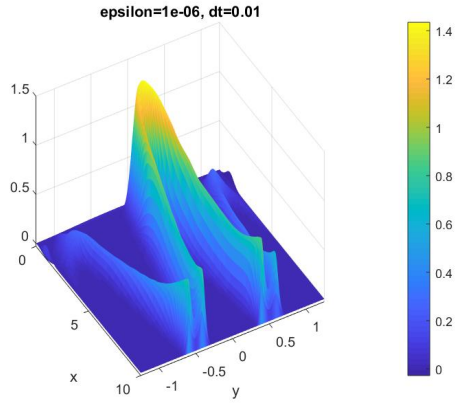
(a): FM-FEM



(b): FM-SUPG



(c): MM-FEM



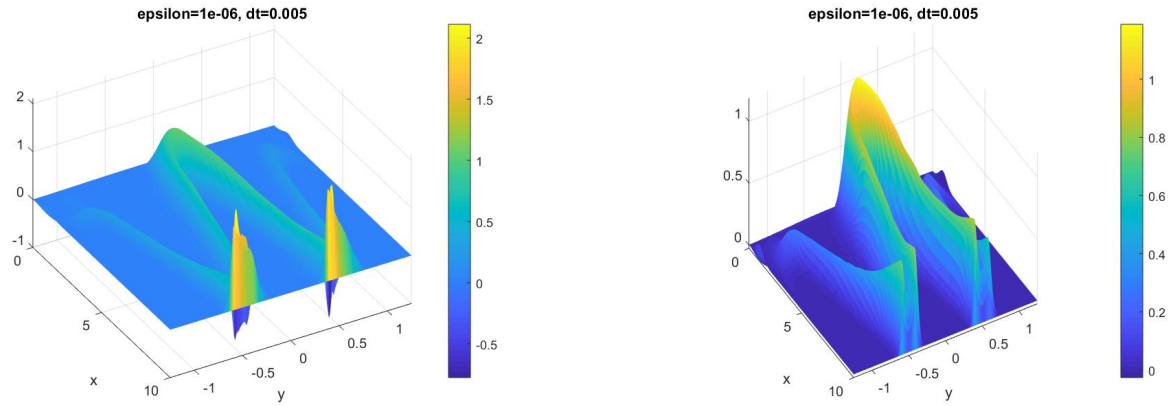
(d): MM-SUPG

Figure 1: Example 3.1 – side view of four different methods with $\Delta t = 0.01$, $\boldsymbol{b}$ is given by (37), $N = 32,768$ at $T = 5.0$.

For this example, $\varepsilon = 10^{-6}$ is used in the computation, which gives $\theta \approx -3.948 \times 10^{-5}$. We compute the diffusion of the initial solution from 0 to 5 seconds with Dirichlet boundary conditions. Time step $\Delta t = 10^{-2}$ and $N = 32,768$ elements are used in the computations. Allowing the convection process to be carried out, we see fine layers form on the interior and boundary of Ω.

The results are shown in Figure 1 with side view profiles. From Figures 1(a) and (c), we see that for the classic Galerkin method, both FM-FEM and MM-FEM, the numerical solutions have non-physical oscillations form near the outflow boundary at $x = 10$. Figure 1(b) shows that FM-SUPG resolves the boundary layers present in FM-FEM and MM-FEM but has artificial oscillations on the interior of Ω. On the other hand, Figure 1(d) shows that MM-SUPG results in the smoothest surface among all four methods, and both interior oscillations and oscillations near the outflow have been resolved. The computations are also performed with smaller Δt as 5×10^{-3}, and the results are shown in Figure 2.

As can be seen, even with smaller time step, MM-FEM still has layers at the outflow while MM-SUPG performs much better.



(a): MM-FEM (b): MM-SUPG

Figure 2: Example 3.1 – side view of MM-FEM and MM-SUPG methods with $\Delta t = 0.005$, $\boldsymbol{b}$ is given by (37), $N = 32,768$ at $T = 5.0$.

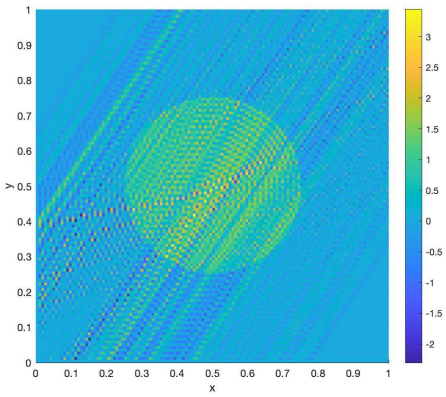
Example 3.2. Here we consider another example (taken from [16]) with exact solution

$$u(x,y,t) = 16\sin(\pi t)x(1-x)y(1-y) \times \left(\frac{1}{2} + \frac{1}{\pi}\arctan\left(2\varepsilon^{-1/2}(0.25^2 - (x-0.5)^2 - (y-0.5)^2)\right)\right), \tag{39}$$

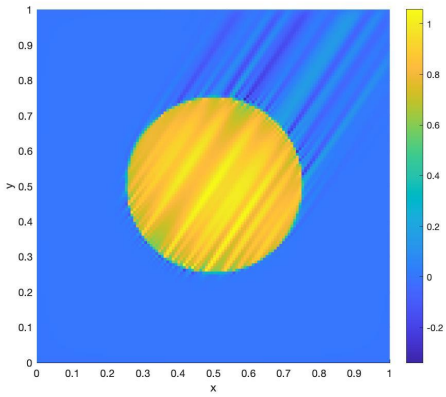
with Dirichlet boundary conditions. We compute to time $T = 0.5$ with time step $\Delta t = 10^{-3}$. The severity of the interior layer is proportional to ε and its thickness is $\sqrt{\varepsilon}$. We also use $\boldsymbol{b} = (2,3)$ and $\varepsilon = 10^{-6}$.

For coarse mesh with fewer elements (that is, smaller N), we observe artificial oscillations in the direction of $\boldsymbol{b}$, including MM-FEM, for example, with $N = 8192$. Increasing the number of mesh elements helps to improve the results for MM-FEM. Figure 3 shows the top view of the solutions at $T = 0.5$ for the four different methods with $N = 32768$, while Figure 4 displays the side view of the solution profiles. As can be seen, significant smearing occurs in the solution obtained using the FM-FEM and FM-SUPG, though the latter shows an improvement. On the other hand, both the MM-FEM and MM-SUPG perform well with $N = 32768$ elements in the mesh. This is supported by the H^1-seminorm errors in Table 1 and numerical solutions shown in Figure 4. The L^2-norm errors in Table 2 shows that MM-SUPG performs comparable even better than MM-FEM for this example. We observe pronounced mesh movement in the direction of $\boldsymbol{b}$ for MM-FEM and MM-SUPG, as shown in Figure 5. The rate of convergence in the H^1-seminorm and L^2 norm, based on errors from Tables 1 and 2 are shown in Figure 6. The reference slopes correspond to expected convergence rates: $O(h)$ in H^1 and $O(h^{3/2}$ in L^2 using MMPDE framework for convection-dominated problems.

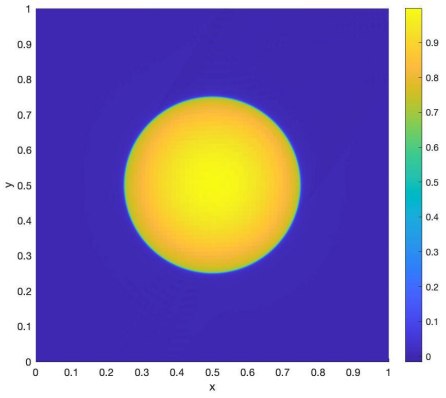
Within this example, we also consider a time-dependent flow given by $\boldsymbol{b} = (y - t, x - t)$. This time-dependent velocity field introduces rotational and time-varying transport, testing robustness of the method under non-uniform flow. The convergence rates in the H^1-seminorm and L^2 norm can be found in Figure 7, where numerical values are recorded in Tables 3 and 4. Again, both MM-FEM and MM-SUPG perform much better than FM-FEM and FM-SUPG with larger N.



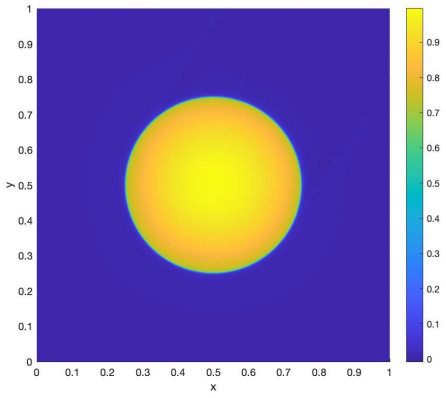
(a): FM-FEM



(b): FM-SUPG



(c): MM-FEM



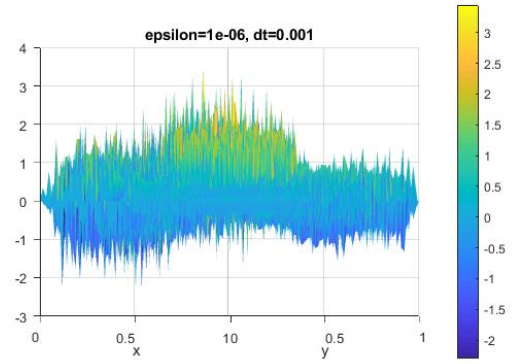
(d): MM-SUPG

Figure 3: Example 3.2 – top view with four different methods with $\Delta t = 0.001$ for $\boldsymbol{b} = (2, 3)$, $N = 32768$ at $T = 0.5$. The observed discreteness in (a) is due to the use of piecewise linear finite elements on a fixed mesh where numerical oscillation is significant, which results into obvious color changes in the top view of the solution.

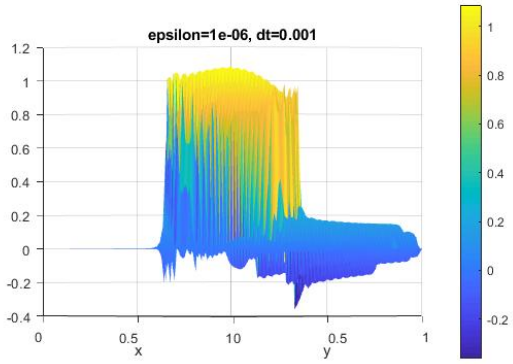
Table 1: Example 3.2 – H^1-seminorm errors for $\varepsilon = 10^{-6}$ at $\Delta t = 0.001$ for $\boldsymbol{b} = (2, 3)$.

N	MM-FEM	MM-SUPG	FM-FEM	FM-SUPG
512	12.1038	10.9715	27.6628	18.3147
2048	10.5624	11.0912	26.7879	15.3485
8192	4.4580	4.3027	34.9201	14.7499
32768	0.7387	0.9278	39.3491	10.7836
131072	0.2657	0.2101	39.2190	7.7920

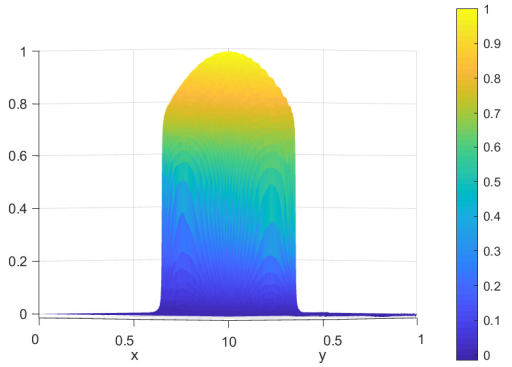
We see small oscillations along the top of the surface in the direction of the outflow in Figures 8, even for large N. However, for larger N, we see that DMP is not violated. Both methods converge to the same approximate solution as N increases. In Figure 4 we see that the MM-SUPG method is a noticeable improvement compared to all methods. In Figure 8, only slight oscillations are observable on the backside of the surface (in the direction of the streamlines) for MM-FEM. This is more prominent



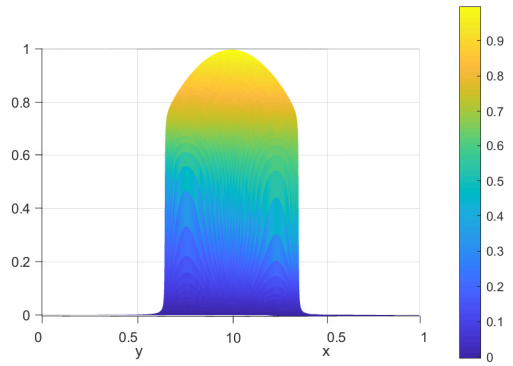
(a): FM-FEM



(b): FM-SUPG



(c): MM-FEM



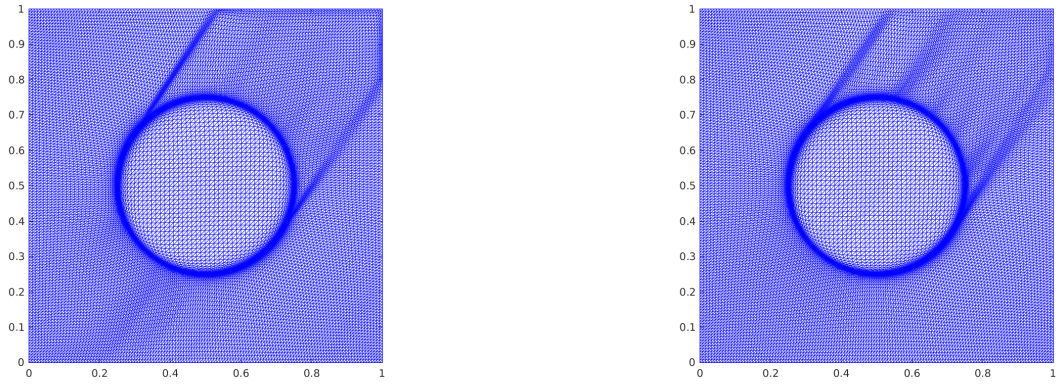
(d): MM-SUPG

Figure 4: Example 3.2 – side view with the four different methods with $\Delta t = 0.001$, $\boldsymbol{b} = (2, 3)$, $N = 32768$ at $T = 0.5$.

Table 2: Example 3.2 – L^2 norm errors for $\varepsilon = 10^{-6}$ at $\Delta t = 0.001$ for $\boldsymbol{b} = (2, 3)$.

N	MM-FEM	MM-SUPG	FM-FEM	FM-SUPG
512	0.2633	0.1289	1.5475	0.4138
2048	0.1127	0.0823	0.8708	0.2871
8192	0.0475	0.0096	0.5818	0.1573
32768	0.0019	0.0010	0.3720	0.0510
131072	0.0004	0.0005	0.1995	0.0180

for smaller N in the MM-FEM case. The different results are due to the fact that Example 3.1 involves strong advective transport of sharp features, where SUPG stabilization plays a dominant role. In contrast, Example 3.2 has a smoother structure where mesh adaptation alone already captures the solution well, reducing the relative impact of SUPG. Overall, MM-SUPG provides improved stability and reduced oscillations compared to fixed mesh methods, while offering comparable accuracy to MM-FEM.



(a): MM-FEM (b): MM-SUPG

Figure 5: Example 3.2 – Mesh movement for $T = 0.5$ and $\boldsymbol{b} = (2, 3)$, $N = 32768$ at $T = 0.5$.

Table 3: Example 3.2 – H^1-seminorm errors for $\varepsilon = 10^{-6}$ at $\Delta t = 0.001$ for $\boldsymbol{b} = (y - t, x - t)$.

N	MM-FEM	MM-SUPG	FM-FEM	FM-SUPG
512	11.1473	11.7017	16.7382	16.6045
2048	10.6529	9.7031	13.0265	12.8651
8192	5.8337	4.7036	12.9658	12.3947
32768	0.7614	0.9654	11.0075	9.8446
131072	0.1614	0.2054	8.7996	7.1405

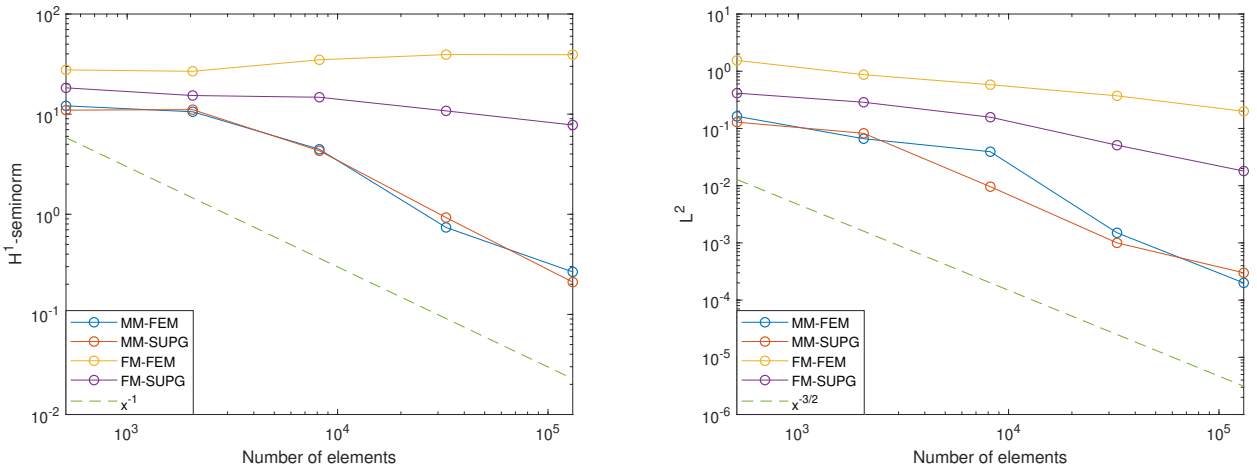
4 Discrete Maximum Principle for Anisotropic Diffusion

In this section, we consider the anisotropic diffusion case where the diffusivity tensor $\mathbb{D}$ has different eigenvalues, which indicates diffusion is faster in one direction (the principal diffusion direction) than in the other direction.

In the isotropic case, mesh adaptation primarily improves resolution of steep gradients. In contrast, for anisotropic diffusion, the orientation of the mesh relative to the diffusion tensor becomes critical. The discrete maximum principle is highly sensitive to mesh geometry and coefficient alignment. In the MM-SUPG framework, both the stiffness matrix and stabilization terms depend on the evolving mesh. Misalignment may lead to violation of the discrete maximum principle (DMP), even in stabilized formulations. Therefore, the design of metric tensors must account for both diffusion anisotropy and convection direction. The proposed tensor and metric intersection strategy aim to balance the

Table 4: Example 3.2 – L^2 errors for $\varepsilon = 10^{-6}$ at $\Delta t = 0.001$ for $\boldsymbol{b} = (y - t, x - t)$.

N	MM-FEM	MM-SUPG	FM-FEM	FM-SUPG
512	0.1629	0.1181	0.3340	0.1843
2048	0.0662	0.0530	0.2326	0.1193
8192	0.0392	0.0130	0.1724	0.0903
32768	0.0015	0.0011	0.1175	0.0411
131072	0.0002	0.0002	0.0602	0.0161



(a): H^1-seminorm error. (b): L^2 norm error.

Figure 6: Example 3.2 – Errors corresponds to Tables 1 and 2. H^1-seminorm and L^2 norm errors with $\varepsilon = 10^{-6}$ and $\Delta t = 0.001$ with flow field $\boldsymbol{b} = (2, 3)$ at $T = 0.5$.

competing effects of convection and anisotropic diffusion.

We say the solution u to (1) satisfies the weak maximum principle if $u \in C^{2,1}(\Omega_T) \cap C(\overline{\Omega}_T)$ and $f \leq 0$ in Ω_T [39]:

$$\max_{(\boldsymbol{x},t)\in\overline{\Omega}_T} u \leq \max\left\{0, \max_{(\boldsymbol{x},t)\in\Gamma_T} u(\boldsymbol{x},t)\right\}, \tag{40}$$

where $\Gamma_T = \overline{\Omega}_T - \Omega_T$ is the parabolic boundary.

Let S_i^K be the face opposite to the vertex $\boldsymbol{a}_i^K$, the ith vertex of element K. Introduce $\boldsymbol{q}$ vectors as [27]:

$$\boldsymbol{q}_i^K = \frac{\boldsymbol{n}_i^K}{h_i^K}, \tag{41}$$

where $\boldsymbol{n}_i^K$ is the unit inward normal pointing to $\boldsymbol{a}_i^K$, and h_i^K is the distance from vertex $\boldsymbol{a}_i^K$ to face S_i^K. Figure 9 shows an illustration of $\boldsymbol{q}$ vectors.

In [40] it was shown that

$$\nabla\varphi_i|_K = \boldsymbol{q}_i^K. \tag{42}$$

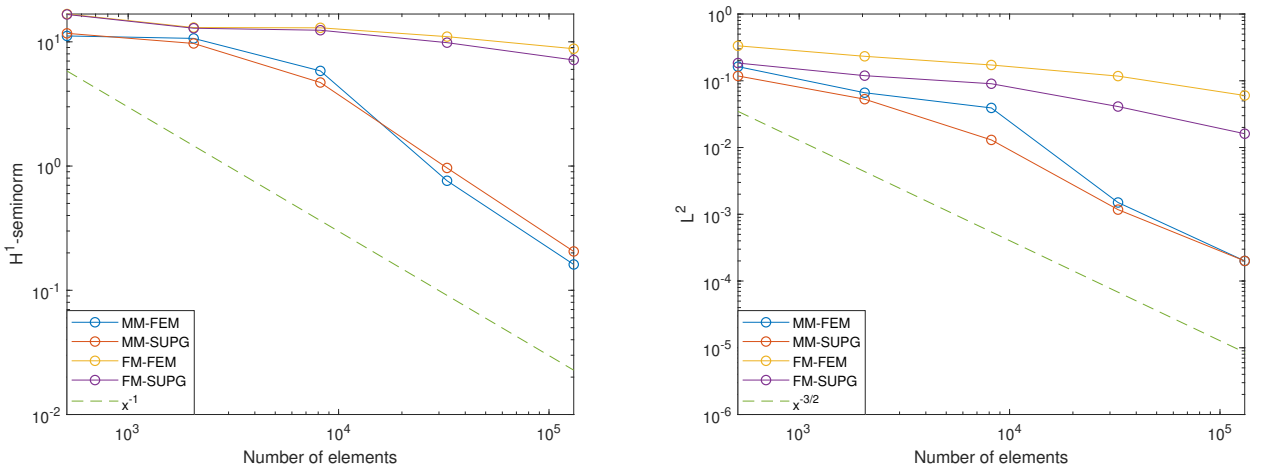
Therefore,

$$\int_K (\nabla\varphi_i)^2 \mathrm{d}\boldsymbol{x} = \|\boldsymbol{q}\|^2|K| = \frac{|K|}{(h_i^K)^2}. \tag{43}$$

Using the Divergence Theorem on an element K, we have

$$\int_K \varphi_i(\boldsymbol{b}\cdot\nabla\varphi_i)\mathrm{d}\boldsymbol{x} = \frac{1}{2}\int_{\partial K}\varphi_i^2(\boldsymbol{b}\cdot\boldsymbol{n})\mathrm{d}\boldsymbol{s} - \frac{1}{2}\int_K(\nabla\cdot\boldsymbol{b})\varphi_i^2\mathrm{d}\boldsymbol{x}. \tag{44}$$

Equation (44) does not determine the sign of the local convection term in general, since the boundary flux contribution on ∂K need not vanish elementwise. Therefore, in the following analysis, we do not



(a): H^1-seminorm error. (b): L^2 norm error.

Figure 7: Example 3.2 – Errors corresponds to Tables 3 and 4. H^1-seminorm and L^2 norm errors with $\varepsilon = 10^{-6}$ and $\Delta t = 0.001$ with flow field $\boldsymbol{b} = (y - t, x - t)$ at $T = 0.5$.

use (44) to infer positivity or negativity of $\int_K \varphi_i(\boldsymbol{b} \cdot \nabla\varphi_i)\mathrm{d}\boldsymbol{x}$. Instead, we use quantitative bounds as follows. Using Cauchy–Schwarz inequality, we obtain

$$\left| \int_K \phi_j (b \cdot \nabla \phi_i)\, dx \right| \le \|\phi_j\|_{L^2(K)} \|b \cdot \nabla \phi_i\|_{L^2(K)} \tag{45}$$

$$\le \|\phi_j\|_{L^2(K)} \|b\|_{\infty,K} \|\nabla \phi_i\|_{L^2(K)}. \tag{46}$$

Since

$$\int_K \phi_j^2\, dx = \frac{2|K|}{(d+1)(d+2)} \quad \text{and} \quad \|\nabla\phi_i\|^2_{L^2(K)} = \frac{|K|}{(h_i^K)^2},$$

it follows that

$$\left| \int_K \phi_j (b \cdot \nabla \phi_i)\, dx \right| \le \frac{\sqrt{2}\,|K|}{\sqrt{(d+1)(d+2)}\, h_i^K} \|b\|_{\infty,K}. \tag{47}$$

A simpler estimate is obtained by noting that $\nabla\phi_i$ is constant on K, $\int_K \varphi_i \mathrm{d}\boldsymbol{x} = \frac{K}{d+1}$, and $|\nabla\varphi_i| = \frac{1}{h_i^K}$:

$$\left| \int_K \phi_j (b \cdot \nabla \phi_i)\, dx \right| \le \int_K |\phi_j|\, |b|\, |\nabla \phi_i|\, dx \tag{48}$$

$$\le \|b\|_{\infty,K} |\nabla\phi_i| \int_K \phi_j\, dx = \frac{|K|}{(d+1)h_i^K} \|b\|_{\infty,K}.$$

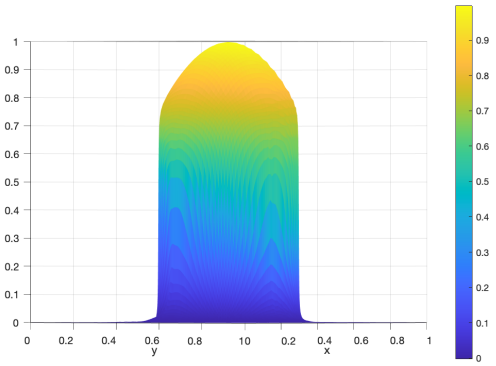
Thus,

$$\left| \int_K \phi_j (b \cdot \nabla \phi_i)\, dx \right| \le \frac{|K|}{(d+1)h_i^K} \|b\|_{\infty,K}. \tag{49}$$

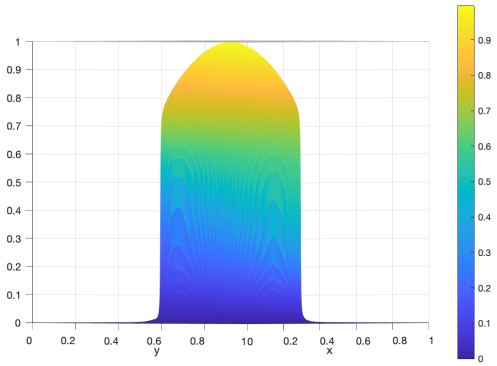
Similarly,

$$\left| \int_K (b \cdot \nabla \phi_i)(b \cdot \nabla \phi_j)\, dx \right| \le \int_K |b|^2 |\nabla\phi_i||\nabla\phi_j|\, dx \tag{50}$$

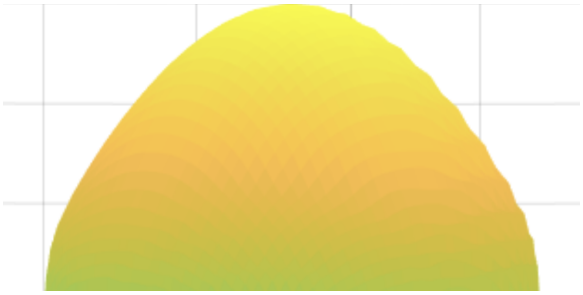
$$\le \|b\|^2_{\infty,K} |\nabla\phi_i||\nabla\phi_j||K| = \frac{|K|}{h_i^K h_j^K} \|b\|^2_{\infty,K},$$



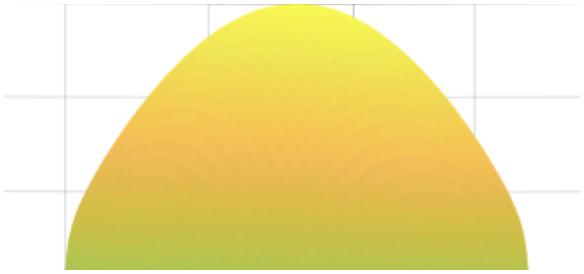
(a): MM-FEM



(b): MM-SUPG



(c): Enlarged view of (a)



(d): Enlarged view of (b)

Figure 8: Example 3.2 – side view with $\Delta t = 0.001$ and $\boldsymbol{b} = (y - t, x - t)$, $\varepsilon = 10^{-6}$, and $N = 32768$ for MM-FEM and MM-SUPG at $T = 0.5$.

which gives

$$\left| \int_K (b \cdot \nabla \phi_i)(b \cdot \nabla \phi_j)\, dx \right| \leq \frac{|K|}{h_i^K h_j^K} \|b\|_{\infty,K}^2. \tag{51}$$

For $\mathbb{D}_K$ in (4), following [27], we define the diffusion-dependent map $\mathbb{D}_K^{-1/2}\boldsymbol{x} : K \to \tilde{K}$. Define the $\mathbb{D}_K$-norm by $\|\boldsymbol{x}\|_{\mathbb{D}_K} = \sqrt{\boldsymbol{x}^T \mathbb{D}_K \boldsymbol{x}}$. Then we have the following relations

$$\tilde{\boldsymbol{a}}_i^K = \mathbb{D}_K^{-1/2}\boldsymbol{a}_i^K, \quad \tilde{S}_i^K = \mathbb{D}_K^{-1/2} S_i^K, \quad |\tilde{K}| = \det(\mathbb{D}_K)^{1/2}|K|, \quad \tilde{\boldsymbol{q}}_i^K = \mathbb{D}_K^{1/2}\boldsymbol{q}_i^K, \quad \tilde{h}_i^K = \|\tilde{\boldsymbol{q}}_i^K\|^{-1}. \tag{52}$$

Denote by $\tilde{\alpha}_{ij}^K$ the dihedral angle between faces $\tilde{S}_i^K$ and $\tilde{S}_j^K$. Then

$$\cos(\tilde{\alpha}_{ij}^K) = -\frac{(\tilde{\boldsymbol{q}}_i^K)^T \tilde{\boldsymbol{q}}_j^K}{\|\tilde{\boldsymbol{q}}_i^K\| \|\tilde{\boldsymbol{q}}_j^K\|} = -\frac{(\boldsymbol{q}_i^K)^T \mathbb{D}_K \boldsymbol{q}_j^K}{\|\boldsymbol{q}_i^K\|_{\mathbb{D}_K} \|\boldsymbol{q}_j^K\|_{\mathbb{D}_K}}, \quad i \neq j. \tag{53}$$

Similarly, on the element K, we have the relation

$$\cos(\alpha_{ij}) = -\boldsymbol{n}_i \cdot \boldsymbol{n}_j = -\frac{\boldsymbol{q}_i \cdot \boldsymbol{q}_j}{\|\boldsymbol{q}_i\| \|\boldsymbol{q}_j\|}. \tag{54}$$

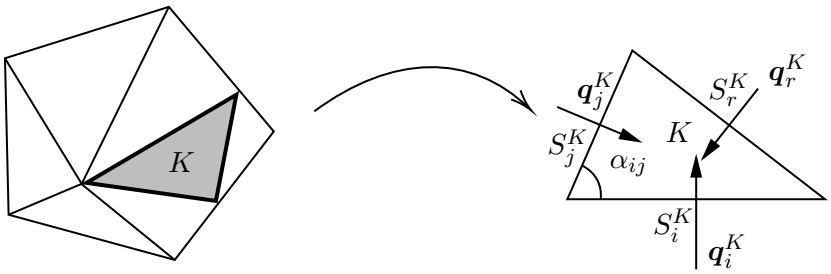


Figure 9: Illustration of faces and $\boldsymbol{q}$ vectors.

It follows from (41) that

$$\frac{h_i^K}{\sqrt{\rho_{max}(\mathbb{D}_K)}} \leq \tilde{h}_i^K \leq \frac{h_i^K}{\sqrt{\rho_{min}(\mathbb{D}_K)}}, \tag{55}$$

where $\rho_{min}(\mathbb{D}_K)$ and $\rho_{max}(\mathbb{D}_K)$ are the min and max eigenvalues of $\mathbb{D}_K$, respectively.

The difficulty with the SUPG formulation (6) is the coupling of the diffusion tensor with the flow field $\boldsymbol{b}$ in (27) as follows:

$$(\mathbb{D}\nabla u_h, \nabla(\boldsymbol{b}\cdot\nabla\varphi^h))_{0,K}. \tag{56}$$

For general $\boldsymbol{b} = (b_1(x_1,x_2), b_2(x_1,x_2))$, consider

$$\frac{\partial \boldsymbol{b}}{\partial x_i} \otimes \mathbb{D}_{(i)}, \quad i = 1,\ldots,d, \tag{57}$$

where $\mathbb{D}_{(i)}$ is the vector ith row of $\mathbb{D}$. An essential assumption to the following analysis is that the tensor defined by

$$\mathbb{K} = \sum_{i=1}^{d} \frac{\partial \boldsymbol{b}}{\partial x_i} \otimes \mathbb{D}_{(i)}, \tag{58}$$

is symmetric and positive definite. This highlights a significant relationship between $\dfrac{\partial \boldsymbol{b}}{\partial x_i}$ and the ith row of $\mathbb{D}$ when using SUPG. For each element $K \in \mathcal{T}_h$, and denote the elementwise average of $\mathbb{K}$ as

$$\mathbb{K}_K = \frac{1}{|K|}\int_K \mathbb{K}(\boldsymbol{x})d\boldsymbol{x}, \tag{59}$$

according to (24).

In the analysis below, we assume that $\mathbb{D}_K$ and $\mathbb{K}_K$ are symmetric and positive definite. This assumption imposes a structural alignment between the velocity field and the diffusion tensor. While restrictive, it enables tractable analysis of DMP preservation. This assumption is not expected to hold in general flows but serves as a sufficient condition.

The added diffusion integral term in (6) can be written in general form as

$$\tau_K \int_K (\mathbb{D}\nabla u)\cdot\nabla(\boldsymbol{b}\cdot\nabla v)\mathrm{d}\boldsymbol{x} = \tau_K \sum_{i=1}^{d}\int_K (\nabla v)^T \left(\frac{\partial \boldsymbol{b}}{\partial x_i}\otimes\mathbb{D}_{(i)}\right)\nabla u \mathrm{d}\boldsymbol{x} = \tau_K(\nabla v)^T\mathbb{K}_K\nabla u. \tag{60}$$

The analysis of the numerical scheme relies on the following fundamental lemma [41].

Lemma 4.1. *If $\tilde{\boldsymbol{A}}$ in (32) is an M-matrix and has nonnegative row sums, then the numerical scheme satisfies DMP.*

The following lemmas were used in [26] to analyze DMP with $\boldsymbol{b} = \boldsymbol{0}$.

Lemma 4.2. *If the mesh satisfies the anisotropic nonobtuse angle condition (ANAC)*

$$0 < \tilde{\alpha}_{ij}^K \leq \frac{\pi}{2}, \quad \forall i, j = 1, \ldots, d+1, i \neq j, \forall K \in \mathcal{T}_h, \tag{61}$$

then for any $\mathbb{F}_K$ of SPD-type, the matrix defined by

$$\sum_{K \in \mathcal{T}_h} |K| (\nabla \varphi_i)^T \mathbb{F}_K \nabla \varphi_j, \quad i =, 1 \ldots N_{vi}, j = 1, \ldots, N_v \tag{62}$$

is an M-matrix and has nonnegative row sums. Moreover, for any element $K \in \mathcal{T}_h$, and $i, j = 1, \ldots, d+1$,

$$(\nabla \varphi_i)^T \mathbb{F}_K \nabla \varphi_j = \begin{cases} -\dfrac{\cos(\tilde{\alpha}_{ij}^K)}{\tilde{h}_i^K \tilde{h}_j^K} & \text{for} \quad i \neq j \\ \dfrac{1}{(\tilde{h}_i^K)^2} & \text{for} \quad i = j \end{cases}. \tag{63}$$

When $d = 2$, the case $i \neq j$ simplifies to

$$|K| (\nabla \varphi_i)^T \mathbb{F}_K \nabla \varphi_j = -\frac{\sqrt{\det(\mathbb{F}_K)}}{2} \cot(\tilde{\alpha}_{ij}^K). \tag{64}$$

We establish DMP in a series of steps. We first investigate spatial estimates, which are critical to understanding the required structure of the matrices $\boldsymbol{A}$ and $\boldsymbol{M}$. Temporal estimates are then derived in order to apply Lemma 4.1.

4.1 General flow forms

In this section we investigate the structure of $\mathbb{K}$ specified by $\boldsymbol{b}$. For $d = 2$, we have

$$\frac{\partial \boldsymbol{b}}{\partial x_1} \otimes \mathbb{D}_{(1)} = \begin{pmatrix} \dfrac{\partial b_1}{\partial x_1} D_{11} & \dfrac{\partial b_1}{\partial x_1} D_{12} \\ \dfrac{\partial b_2}{\partial x_1} D_{11} & \dfrac{\partial b_2}{\partial x_1} D_{12} \end{pmatrix}, \tag{65}$$

and

$$\frac{\partial \boldsymbol{b}}{\partial x_2} \otimes \mathbb{D}_{(2)} = \begin{pmatrix} \dfrac{\partial b_1}{\partial x_2} D_{21} & \dfrac{\partial b_1}{\partial x_2} D_{22} \\ \dfrac{\partial b_2}{\partial x_2} D_{21} & \dfrac{\partial b_2}{\partial x_2} D_{22} \end{pmatrix}. \tag{66}$$

There are two ways in which we consider $\mathbb{K}$. If (65) and (66) are of SPD-type, then $\mathbb{K}$ will be as well. If not, then we need to analyze $\mathbb{K}$ in its general form given by

$$\mathbb{K} = \begin{pmatrix} \frac{\partial b_1}{\partial x_1} D_{11} + \frac{\partial b_1}{\partial x_2} D_{21} & \frac{\partial b_1}{\partial x_1} D_{12} + \frac{\partial b_1}{\partial x_2} D_{22} \\ \frac{\partial b_2}{\partial x_1} D_{11} + \frac{\partial b_2}{\partial x_2} D_{21} & \frac{\partial b_2}{\partial x_1} D_{12} + \frac{\partial b_2}{\partial x_2} D_{22} \end{pmatrix}. \tag{67}$$

Hence, using the symmetric condition, we would require

$$\frac{\partial}{\partial x_1}(b_2 D_{11} - b_1 D_{12}) = \frac{\partial}{\partial x_2}(b_1 D_{22} - b_2 D_{21}). \tag{68}$$

Assume first that these matrices are of SPD-type so that general form of $\mathbb{K}$ is of SPD-type.

Theorem 4.3. *Suppose $\mathbb{D}$, equations (65), and (66) are of SPD-type with constant real-valued entries and $D_{12} = D_{21} \neq 0$. Then there are functions $f_1(x_2, t)$, $g_1(x_2, t)$, $f_2(x_1, t)$, and $g_2(x_1, t)$ such that*

$$b_1(x_1, x_2, t) = \frac{D_{11}D_{22}f_1(x_1, t) - D_{12}D_{21}f_2(x_2, t)}{\det(\mathbb{D})}, \tag{69}$$

and

$$b_2(x_1, x_2, t) = \frac{D_{12}D_{22}(f_1(x_1, t) - f_2(x_2, t))}{\det(\mathbb{D})}, \tag{70}$$

for $D_{22} \neq 0$. Likewise,

$$b_1(x_1, x_2, t) = \frac{D_{11}D_{21}(g_2(x_2, t) - g_1(x_1, t))}{\det(\mathbb{D})}, \tag{71}$$

and

$$b_2(x_1, x_2, t) = \frac{D_{11}D_{22}g_2(x_2, t) - D_{12}D_{21}g_1(x_1, t)}{\det(\mathbb{D})}, \tag{72}$$

whenever $D_{11} \neq 0$.

Proof. Assume $D_{12} = D_{21} \neq 0$ and $D_{22} \neq 0$. The other case is very similar. Assuming symmetry as mentioned, we have

$$\frac{\partial b_1}{\partial x_1} = \left(\frac{D_{11}}{D_{12}}\right)\frac{\partial b_2}{\partial x_1}, \tag{73}$$

$$b_1(x_1, x_2, t) = \left(\frac{D_{11}}{D_{12}}\right) b_2(x_1, x_2, t) + f_2(x_2, t), \tag{74}$$

for some $f_2(x_2, t)$. Additionally,

$$\frac{\partial b_1}{\partial x_2} = \left(\frac{D_{21}}{D_{22}}\right)\frac{\partial b_2}{\partial x_2}, \tag{75}$$

$$b_1(x_1, x_2, t) = \left(\frac{D_{21}}{D_{22}}\right) b_2(x_1, x_2, t) + f_1(x_1, t), \tag{76}$$

for some function $f_1(x_1, t)$. Equating the two equations for b_1 gives

$$\left(\frac{D_{11}}{D_{12}}\right) b_2(x_1, x_2, t) + f_2(x_2, t) = \left(\frac{D_{21}}{D_{22}}\right) b_2(x_1, x_2, t) + f_1(x_1, t), \tag{77}$$

$$\left(\frac{D_{11}}{D_{12}} - \frac{D_{21}}{D_{22}}\right) b_2(x_1, x_2, t) = f_1(x_2, t) - f_2(x_1, t), \tag{78}$$

$$b_2(x_1, x_2, t) = \frac{D_{12}D_{22}(f_1(x_1, t) - f_2(x_2, t))}{\det(\mathbb{D})}. \tag{79}$$

Plugging this into the equation for b_1, we have

$$\begin{aligned} b_1(x_1, x_2, t) &= \left(\frac{D_{11}}{D_{12}}\right) \frac{D_{12}D_{22}(f_1(x_1,t) - f_2(x_2,t))}{\det(\mathbb{D})} + f_2(x_2,t) \\ &= \frac{D_{11}D_{22}(f_1(x_1,t) - f_2(x_2,t)) + f_2(x_2,t)(D_{11}D_{22} - D_{12}D_{21})}{\det(\mathbb{D})} \\ &= \frac{D_{11}D_{22}f_1(x_1,t) - D_{12}D_{21}f_2(x_2,t)}{\det(\mathbb{D})}. \end{aligned} \tag{80}$$

Note that since $\mathbb{D}$ is of SPD-type, $\det(\mathbb{D}) \neq 0$. □

Remark 4.1. If $D_{12} = D_{21}$ is not zero, the general form of $\boldsymbol{b}$ is given by

$$\frac{1}{\det(\mathbb{D})} \left(D_{22}f_1(x_1,t)\mathbb{D}_{(1)} - D_{12}f_2(x_2,t)\mathbb{D}_{(2)}\right), \tag{81}$$

for some functions f_1 and f_2 whenever $D_{22} \neq 0$. Similarly, if $D_{11} \neq 0$, then the general form of $\boldsymbol{b}$ is given by

$$\frac{1}{\det(\mathbb{D})} \left(D_{11}g_2(x_2,t)\mathbb{D}_{(2)} - D_{21}g_1(x_1,t)\mathbb{D}_{(1)}\right). \tag{82}$$

If D_{11} and D_{22} are both not zero, then $g_1 = -\dfrac{D_{22}}{D_{21}}f_1$ and $g_2 = -\dfrac{D_{12}}{D_{11}}f_2$.

Corollary 4.3.1. *If $\mathbb{D} = diag(D_{11}, D_{22})$, with D_{11} and $D_{22} \neq 0$, then there are functions f and g such that*

$$\boldsymbol{b}(x_1, x_2, t) = \begin{pmatrix} \dfrac{f(x_1,t)}{D_{11}} \\ \dfrac{g(x_2,t)}{D_{22}} \end{pmatrix}, \tag{83}$$

and

$$\mathbb{K} = \begin{pmatrix} \dfrac{\partial f}{\partial x_1} & 0 \\ 0 & \dfrac{\partial g}{\partial x_2} \end{pmatrix}. \tag{84}$$

Proof. The form of $\boldsymbol{b}$ as shown in (83) follows from a similar argument as shown in the proof of the previous theorem. □

Based on the Remark 4.1, we require

$$\frac{D_{22}}{\det(\mathbb{D})} \left(D_{11}\frac{\partial f_1}{\partial x_1} - D_{12}\frac{\partial f_2}{\partial x_2}\right) \tag{85}$$

to share the same sign as the divergence of $\boldsymbol{b}$.

4.2 Mass and stiffness matrix structure: estimates on τ_K

We first establish non-negative diagonal entries for the mass and stiffness matrices.

Lemma 4.4. *Suppose $\mathbb{K}_K$ and $\mathbb{D}_K$ are symmetric and positive definite and the meshes associated with $\mathbb{D}_K$ and $\mathbb{K}_K$ satisfy ANAC. Then the diagonal entries of the mass and stiffness matrices satisfy $M_{ii} > 0$ and $A_{ii} > 0$ provided*

$$\tau_K \|b\|_{\infty,K} \leq \frac{h_i^K}{d+2}, \tag{86}$$

and

$$\frac{\rho_{\min}(\mathbb{D}_K) + \tau_K \rho_{\min}(\mathbb{K}_K)}{(h_i^K)^2} > \frac{1}{2\varepsilon\tau_K(d+1)(d+2)}. \tag{87}$$

Proof. We consider a patch of elements ω_i that share a vertex $\boldsymbol{a}_i$. Using (49) and (21), from (26), we obtain

$$M_{ii} \geq \sum_{K\in\omega_i} \left(\frac{2|K|}{(d+1)(d+2)} - \tau_K \frac{|K|}{(d+1)h_i^K} \|b\|_{\infty,K} \right),$$

which is positive under (86).

Applying Young's inequality

$$\phi_i(b\cdot\nabla\phi_i) + \tau_K(b\cdot\nabla\phi_i)^2 \geq -\frac{1}{4\tau_K}\phi_i^2,$$

for A_{ii} in (27), we have,

$$\begin{aligned}
A_{ii} &= \sum_{K\in\omega_i} \int_K \left(\varphi_i \boldsymbol{b}\cdot\nabla\varphi_i + \tau_K(\boldsymbol{b}\cdot\nabla\varphi_i)^2 + \tau_K\left(\varepsilon\mathbb{D}_K\nabla\varphi_i\right)\cdot\nabla(\boldsymbol{b}\cdot\nabla\varphi_i)\right) \mathrm{d}\boldsymbol{x} + |K|(\nabla\varphi_i)^T\varepsilon\mathbb{D}_K\nabla\varphi_i \\
&= \sum_{K\in\omega_i} \int_K \left(\varphi_i \boldsymbol{b}\cdot\nabla\varphi_i + \tau_K(\boldsymbol{b}\cdot\nabla\varphi_i)^2\right) \mathrm{d}\boldsymbol{x} + \tau_K|K|(\nabla\varphi_i)^T\varepsilon\mathbb{K}_K\nabla\varphi_i + |K|(\nabla\varphi_i)^T\varepsilon\mathbb{D}_K\nabla\varphi_i \\
&\geq \sum_{K\in\omega_i} \int_K -\frac{1}{4\tau_K}\phi_i^2 \mathrm{d}\boldsymbol{x} + |K|(\nabla\varphi_i)^T\varepsilon(\tau_K\mathbb{K}_K + \mathbb{D}_K)\nabla\varphi_i \\
&= \sum_{K\in\omega_i} -\frac{|K|}{2\tau_K(d+1)(d+2)} + |K|(\nabla\varphi_i)^T\varepsilon(\tau_K\mathbb{K}_K + \mathbb{D}_K)\nabla\varphi_i \\
&\geq \sum_{K\in\omega_i} |K| \left(\varepsilon\frac{\rho_{min}(\mathbb{D}_K) + \tau_K\rho_{min}(\mathbb{K}_K)}{(h_i^K)^2} - \frac{1}{2\tau_K(d+1)(d+2)}\right),
\end{aligned} \tag{88}$$

which yields positivity under (87). □

It is important to keep in mind that this estimate is only for those elements associated with A_{ii} on the interior of Ω. This estimate highlights the significant relationship between τ_K and the convection term.

Lemma 4.5. *Suppose $\mathbb{K}_K$ and $\mathbb{D}_K$ are symmetric and positive definite and the meshes associated with $\mathbb{D}_K$ and $\mathbb{K}_K$ satisfy ANAC. For $i \neq j$, define*

$$\delta_{ij}^K = \rho_{\min}(\varepsilon\mathbb{D}_K)\cos(\widetilde{\alpha}_{ij}^{\mathbb{D}}) + \tau_K\rho_{\min}(\mathbb{K}_K)\cos(\check{\alpha}_{ij}^{\mathbb{K}}) - \frac{h_i^K}{d+1}\|b\|_{\infty,K} - \tau_K\|b\|_{\infty,K}^2, \tag{89}$$

where $\widetilde{\alpha}_{ij}^{\mathbb{D}}$ *and* $\widetilde{\alpha}_{ij}^{\mathbb{K}}$ *are the angles in the meshes associated with* $\mathbb{D}_K$ *and* $\mathbb{K}_K$*, respectively. If*

$$\delta_{ij}^K \geq 0 \quad \text{for all } K \in \mathcal{T}_h,\ i \neq j, \tag{90}$$

then $A_{ij} \leq 0$.

Proof. For $i \neq j$, from (27) we have

$$A_{ij} = \sum_{K \in \omega_i \cap \omega_j} \left(\int_K \varphi_i (\boldsymbol{b} \cdot \nabla \varphi_j) \mathrm{d}\boldsymbol{x} + \tau_K \int_K (\boldsymbol{b} \cdot \nabla \varphi_j)(\boldsymbol{b} \cdot \nabla \varphi_i) \mathrm{d}\boldsymbol{x} + |K| (\nabla \varphi_i)^T (\varepsilon \mathbb{D}_K + \varepsilon \tau_K \mathbb{K}_K) \nabla \varphi_j \right).$$

Applying Lemma 4.2 to $\varepsilon \mathbb{D}_K$ and $\varepsilon \mathbb{K}_K$, we have

$$|K| \nabla \varphi_i)^T (\varepsilon \mathbb{D}_K) \nabla \varphi_j \leq -\frac{|K|}{h_i^K h_j^K} \rho_{min}(\varepsilon \mathbb{D}_K) \cos(\widetilde{\alpha}_{ij}^{\mathbb{D}}), \tag{91}$$

and

$$|K| \nabla \varphi_i)^T (\varepsilon \mathbb{K}_K) \nabla \varphi_j \leq -\frac{|K|}{h_i^K h_j^K} \rho_{min}(\varepsilon \mathbb{K}_K) \cos(\widetilde{\alpha}_{ij}^{\mathbb{K}}). \tag{92}$$

Using inequalities (49) and (51), we have

$$\begin{aligned} A_{ij} &\leq \sum_{K \in \omega_i \cap \omega_j} \frac{|K|}{h_i^K h_j^K} \left(\frac{h_i^K}{d+1} \|b\|_{\infty,K} + \tau_K \|b\|_{\infty,K}^2 - \rho_{min}(\varepsilon \mathbb{D}_K) \cos(\widetilde{\alpha}_{ij}^{\mathbb{D}}) - \tau_K \rho_{min}(\varepsilon \mathbb{K}_K) \cos(\widetilde{\alpha}_{ij}^{\mathbb{K}}) \right) \\ &= -\sum_{K \in \omega_i \cap \omega_j} \frac{|K|}{h_i^K h_j^K} \delta_{ij}^K. \end{aligned} \tag{93}$$

Therefore, $A_{ij} \leq 0$ if (90) holds. □

In real computations, there is only one mesh at each physical time, thus $\widetilde{\alpha}_{ij}^{\mathbb{D}}$ and $\widetilde{\alpha}_{ij}^{\mathbb{K}}$ will be referred to the same angle, which will be discussed in later subsection 4.4.

Lemma 4.6. *For* $i \neq j$*, if*

$$\tau_K \|b\|_{\infty,K} \leq \frac{h_i^K}{d+2} \quad \text{for all } K \in \mathcal{T}_h, \tag{94}$$

then $M_{ij} \geq 0$.

Proof. For $i \neq j$, from (26) and (21)

$$M_{ij} = \sum_{K \in \omega_i \cap \omega_j} \left(\frac{|K|}{(d+1)(d+2)} + \tau_K \int_K \varphi_j (\boldsymbol{b} \cdot \nabla \varphi_i) \mathrm{d}\boldsymbol{x} \right). \tag{95}$$

Applying (49), we have

$$M_{ij} \geq \sum_{K \in \omega_i \cap \omega_j} \left(\frac{|K|}{(d+1)(d+2)} - \tau_K \frac{|K|}{(d+1) h_i^K} \|b\|_{\infty,K} \right), \tag{96}$$

which is nonnegative under condition (94). □

Combining the previous two lemmas, we have the following result immediately.

Lemma 4.7. *Assume the hypotheses of Lemmas 4.5 and 4.6. Then*

$$M_{ij} - \Delta t (1-\theta) A_{ij} \geq 0, \quad i \neq j. \tag{97}$$

4.3 Temporal estimates

In this section we derive upper and lower bounds on the fixed time step Δt so that DMP is satisfied.

Define

$$\eta_{ij}^K := \frac{|K|}{h_i^k h_j^K}\delta_{ij}^K, \quad \mu_{ij}^K := \frac{|K|}{(d+1)(d+2)} + \tau_K \frac{|K|}{(d+1)h_i^K}\|b\|_{\infty,K}. \tag{98}$$

Lemma 4.8. *Assume that for each element K in $\mathcal{T}_h$:*

(i) The elementwise tensors $\mathbb{D}_K$ and $\mathbb{K}_K$ are symmetric and positive definite.

(ii) The meshes associated with $\mathbb{D}_K$ and $\mathbb{K}_K$ satisfy the anisotropic nonobtuse angle condition (ANAC).

(iii) The stabilization parameter τ_K satisfies

$$\tau_K\|b\|_{\infty,K} \le C_1 h_K, \tag{99}$$

and

$$\varepsilon(\rho_{min}(\mathbb{D}_K) + \tau_K\rho_{min}(\mathbb{K}_K)) > C_2\frac{h_K^2}{\tau_K}, \tag{100}$$

for positive constants C_1, and C_2 independent of the mesh.

(iv) For all $i \neq j$, δ_{ij}^K defined in (89) is positive.

Then, if

$$\Delta t \ge \frac{1}{\theta}\max_{i\neq j}\frac{\sum_{K\in\omega_i\cap\omega_j}\mu_{ij}^K}{\sum_{K\in\omega_i\cap\omega_j}\eta_{ij}^K}, \tag{101}$$

the matrix $\tilde{\boldsymbol{A}} = \boldsymbol{M} + \theta\Delta t\boldsymbol{A}$ is an M-matrix.

Proof. We first show that $\tilde{\boldsymbol{A}}$ is a Z-matrix. For $i \neq j$, from (26) and (21)

$$M_{ij} = \sum_{K\in\omega_i\cap\omega_j}\left(\frac{|K|}{(d+1)(d+2)} + \tau_K\int_K \varphi_j(\boldsymbol{b}\cdot\nabla\varphi_i)\mathrm{d}\boldsymbol{x}\right). \tag{102}$$

Applying (49), we have

$$M_{ij} \le \sum_{K\in\omega_i\cap\omega_j}\left(\frac{|K|}{(d+1)(d+2)} + \tau_K\frac{|K|}{(d+1)h_i^K}\|b\|_{\infty,K}\right) = \sum_{K\in\omega_i\cap\omega_j}\mu_{ij}^K, \tag{103}$$

and from (93), we obtain

$$A_{ij} \le -\sum_{K\in\omega_i\cap\omega_j}\eta_{ij}^K. \tag{104}$$

Thus,

$$\tilde{\boldsymbol{A}}_{ij} = M_{ij} + \theta\Delta t A_{ij} \le \sum_{K\in\omega_i\cap\omega_j}\mu_{ij}^K - \theta\Delta t\sum_{K\in\omega_i\cap\omega_j}\eta_{ij}^K \le 0 \tag{105}$$

provided (101) holds. Hence all off-diagonal entries of $\tilde{\boldsymbol{A}}$ are nonpositive. Combined with the positivity of the diagonal entries of $\boldsymbol{M}$ and $\boldsymbol{A}$, $\tilde{\boldsymbol{A}}$ is a Z-matrix.

Next, we show that $\tilde{\boldsymbol{A}}$ is an M-matrix. Reorder the system into interior and boundary nodes, by (33), we can write

$$\tilde{\boldsymbol{A}} = \begin{bmatrix} \boldsymbol{B} & \boldsymbol{C} \\ 0 & \boldsymbol{I} \end{bmatrix}, \quad \boldsymbol{B} := \boldsymbol{M}_{11} + \theta \Delta t \boldsymbol{A}_{11}.$$

Since $\boldsymbol{C} \leq 0$, it suffices to show that $\boldsymbol{B}$ is an M-matrix.

Firstly, the block $\boldsymbol{B}$ is also a Z-matrix because $\tilde{\boldsymbol{A}}$ is a Z-matrix. Next, we show $\boldsymbol{B}$ is positive-definite.

Let $\boldsymbol{v} = (v_1, v_2, \ldots, v_{N_{vi}})$ be a nonzero vector in $\mathbb{R}^{N_{vi}}$ and write $v_h = \sum_{j=1}^{N_{vi}} v_j \varphi_j \in U_0^h$. Then,

$$\boldsymbol{v}^T \boldsymbol{B} \boldsymbol{v} = \boldsymbol{v}^T \boldsymbol{M}_{11} \boldsymbol{v} + \theta \Delta t \boldsymbol{v}^T \boldsymbol{A}_{11} \boldsymbol{v}, \tag{106}$$

and the interior mass bloack satisfies

$$\boldsymbol{v}^T \boldsymbol{M}_{11} \boldsymbol{v} = \sum_{i,j=1}^{N_{vi}} v_i M_{ij} v_j. \tag{107}$$

Using the definition of M_{ij} in (26), this becomes

$$\boldsymbol{v}^T \boldsymbol{M}_{11} \boldsymbol{v} = \sum_{K \in \mathcal{T}_h} \left(\int_K v_h^2 \mathrm{d}\boldsymbol{x} + \tau_K \int_K v_h (\boldsymbol{b} \cdot \nabla v_h) \mathrm{d}\boldsymbol{x} \right). \tag{108}$$

Applying Cauchy-Schwarz and inverse inequality, we have

$$\left| \int_K v_h (\boldsymbol{b} \cdot \nabla v_h) \mathrm{d}\boldsymbol{x} \right| \leq C_{inv} h_K^{-1} \|b\|_{\infty,K} \|v_h\|_{L^2(K)}^2. \tag{109}$$

Therefore,

$$\begin{aligned} \boldsymbol{v}^T \boldsymbol{M}_{11} \boldsymbol{v} &\geq \sum_{K \in \mathcal{T}_h} \left(\|v_h\|_{L^2(K)}^2 - \tau_K C_{inv} h_K^{-1} \|b\|_{\infty,K} \|v_h\|_{L^2(K)}^2 \right) \\ &= \sum_K \left(1 - C \tau_K h_K^{-1} \|b\|_{\infty,K} \right) \|v_h\|_{L^2(K)}^2 \\ &> 0, \end{aligned}$$

provided τ_K satisfies the condition (99).

Similarly, from (27) and (60), we get,

$$\begin{aligned} \boldsymbol{v}^T \boldsymbol{A}_{11} \boldsymbol{v} = &\sum_{K \in \mathcal{T}_h} \left(\int_K v_h (\boldsymbol{b} \cdot \nabla v_h) \mathrm{d}\boldsymbol{x} + \tau_K \int_K (\boldsymbol{b} \cdot \nabla v_h)^2 \mathrm{d}\boldsymbol{x} \right) \\ &+ \sum_{K \in \mathcal{T}_h} \varepsilon |K| (\nabla v_h)^T (\mathbb{D}_K + \tau_K \mathbb{K}_K) \nabla v_h. \end{aligned} \tag{110}$$

Applying Young's inequality to the scaler quantities v_h and $\boldsymbol{b} \cdot \nabla v_h$, we have

$$v_h (b \cdot \nabla \phi_i) + \tau_K (b \cdot \nabla \phi_i)^2 \geq -\frac{1}{4\tau_K} v_h^2. \tag{111}$$

Integrate over K and substitue into (110), we have

$$\boldsymbol{v}^T \boldsymbol{A}_{11} \boldsymbol{v} \geq \sum_{K \in \mathcal{T}_h} \left(-\frac{1}{4\tau_K} \|v_h\|_{L^2(K)}^2 + \varepsilon |K| (\nabla v_h)^T (\mathbb{D}_K + \tau_K \mathbb{K}_K) \nabla v_h \right).$$

Using SPD properties of $\mathbb{D}_K$ and $\mathbb{K}_K$, we obtain

$$(\nabla v_h)^T(\mathbb{D}_K + \tau_K\mathbb{K}_K)\nabla v_h \geq (\rho_{min}(\mathbb{D}_K) + \tau_K\rho_{min}(\mathbb{K}_K))|\nabla v_h|^2. \tag{112}$$

Since v_h is linear on each element K, ∇v_h is constant on K, so

$$|K||\nabla v_h|^2 = \|\nabla v_h\|^2_{L^2(K)}. \tag{113}$$

Therefore,

$$\boldsymbol{v}^T\boldsymbol{A}_{11}\boldsymbol{v} \geq \sum_{K\in\mathcal{T}_h}\left(-\frac{1}{4\tau_K}\|v_h\|^2_{L^2(K)} + \varepsilon(\rho_{min}(\mathbb{D}_K) + \tau_K\rho_{min}(\mathbb{K}_K))\|\nabla v_h\|^2_{L^2(K)}\right).$$

Applying inverse inequality

$$\|v_h\|_{L^2(K)} \leq C_{inv}h_K\|\nabla v_h\|^2_{L^2(K)}, \tag{114}$$

we have

$$\boldsymbol{v}^T\boldsymbol{A}_{11}\boldsymbol{v} \geq \sum_K\left(\varepsilon(\rho_{min}(\mathbb{D}_K) + \tau_K\rho_{min}(\mathbb{K}_K)) - C^2_{inv}\frac{h_K^2}{4\tau_K}\right)\|\nabla v_h\|^2_{L^2(K)} > 0, \tag{115}$$

provided the condition (100) holds. Thus, $\boldsymbol{v}^T\boldsymbol{B}\boldsymbol{v} > 0$ for all $\boldsymbol{v} \neq \boldsymbol{0}$. So $\boldsymbol{B}$ is positive definite.

Since $\boldsymbol{B}$ is a Z-matrix that is positive definite, it follows that $\boldsymbol{B}$ is a nonsingular M-matrix. Hence $\boldsymbol{B}^{-1} \geq 0$. Then,

$$\tilde{\boldsymbol{A}}^{-1} = \begin{bmatrix}\boldsymbol{B}^{-1} & -\boldsymbol{B}^{-1}\boldsymbol{C}\\ 0 & \boldsymbol{I}\end{bmatrix} \geq 0,$$

since $\boldsymbol{C} \leq 0$. Therefore, $\tilde{\boldsymbol{A}}$ is an M-matrix.

□

Under the same assumptions in Lemma 4.4, we have

$$M_{ii} \geq \sum_{K\in\omega_i}\left(\frac{2|K|}{(d+1)(d+2)} - \tau_K\frac{|K|}{(d+1)h_i^K}\|b\|_{\infty,K}\right),$$

and

$$\begin{aligned}A_{ii} &= \sum_{K\in\omega_i}\int_K\left(\varphi_i\boldsymbol{b}\cdot\nabla\varphi_i + \tau_K(\boldsymbol{b}\cdot\nabla\varphi_i)^2\right)\mathrm{d}\boldsymbol{x} + |K|(\nabla\varphi_i)^T\varepsilon(\tau_K\mathbb{K}_K + \mathbb{D}_K)\nabla\varphi_i\\ &\leq \sum_{K\in\omega_i}|K|\left(\frac{\|b\|_{\infty,K}}{(d+1)h_i^K} + \frac{\|b\|^2_{\infty,K}}{(h_i^K)^2} + \varepsilon\frac{\rho_{min}(\mathbb{D}_K) + \tau_K\rho_{min}(\mathbb{K}_K)}{(h_i^K)^2}\right).\end{aligned} \tag{116}$$

Then we have the following lemma.

Lemma 4.9. *Assume*

$$M_{ii}^{low} := \sum_{K\in\omega_i}\left(\frac{2|K|}{(d+1)(d+2)} - \tau_K\frac{|K|}{(d+1)h_i^K}\|b\|_{\infty,K}\right) > 0, \tag{117}$$

and let

$$A_{ii}^{up} := \sum_{K\in\omega_i}|K|\left(\frac{\|b\|_{\infty,K}}{(d+1)h_i^K} + \frac{\|b\|^2_{\infty,K}}{(h_i^K)^2} + \varepsilon\frac{\rho_{min}(\mathbb{D}_K) + \tau_K\rho_{min}(\mathbb{K}_K)}{(h_i^K)^2}\right). \tag{118}$$

If

$$\Delta t \leq \frac{1}{1-\theta} \min_i \frac{M_{ii}^{low}}{A_{ii}^{up}}, \tag{119}$$

then

$$M_{ii} - \Delta t(1-\theta)A_{ii} \geq 0 \quad \textit{for all } i. \tag{120}$$

Proof. Since $M_{ii} \geq M_{ii}^{low}$ and $A_{ii} \leq A_{ii}^{up}$, we have

$$M_{ii} - \Delta t(1-\theta)A_{ii} \geq M_{ii}^{low} - \Delta t(1-\theta)A_{ii}^{up} \geq 0 \tag{121}$$

whenever (119) holds. □

4.4 Metric tensor for DMP

We now develop the metric tensor to control the mesh so that the numerical solution satisfies DMP. Assume the temporal estimates from the previous section. As mentioned in the proof of Theorem 3.1 of [27], $\tilde{\boldsymbol{A}}^{-1} \geq 0$ and $\tilde{\boldsymbol{A}}$ is an M-matrix. To see that $\tilde{\boldsymbol{A}}$ has nonnegative row sums, note that $\sum_j \varphi_j = 1$. Therefore, the solution to (32) satisfies

$$\max_{i=1,\ldots,(N+1)N_v} (u_h)_i = \max\left\{0, \max_{i \in S(\boldsymbol{f}^+)} (u_h)_i\right\}, \tag{122}$$

where $S(\boldsymbol{f}^+)$ is the set of indices where f_i is nonnegative. Piecewise linear functions attain their maximum value at the vertices of element K. Hence, when $f(\boldsymbol{x}, t) \leq 0$, we have

$$\max_{n=0,\ldots,N} \max_{\boldsymbol{x} \in \overline{\Omega}} u_h^n(\boldsymbol{x}) = \max\left\{0, \max_{n=0,\ldots,N} \max_{\boldsymbol{x} \in \overline{\Omega}} u_h^n(\boldsymbol{x}), \max_{\boldsymbol{x} \in \overline{\Omega}} u_h^0(\boldsymbol{x})\right\}, \tag{123}$$

where

$$u_h^n = \sum_{j=1}^{N_{vi}} u_j^n \varphi_j(\boldsymbol{x}) + \sum_{j=N_{vi}+1}^{N_v} u_j^n \varphi_j(\boldsymbol{x}), \quad n = 0, \ldots, N. \tag{124}$$

Theorem 4.10. *Consider the fully discrete MM-SUPG scheme* (26)-(28)*. Assume that, for each element $K \in \mathcal{T}_h$, the following conditions hold:*

(i) The elementwise tensors $\mathbb{D}_K$ and $\mathbb{K}_K$ are symmetric positive definite.

(ii) The meshes associated with $\mathbb{D}_K$ and $\mathbb{K}_K$ satisfy the anisotropic nonobtuse angle condition (ANAC).

(iii) The stabilization parameter τ_K satisfies (86)*,* (87)*, and* (90)*.*

(vi) The time step Δt satisfies (101) *and* (119)*.*

Then the coefficient matrices satisfy:

(a) $\boldsymbol{M} - \Delta t(1-\theta)\boldsymbol{A}$ has nonnegative entries,

(b) $\widetilde{\boldsymbol{A}} = \boldsymbol{M} + \theta \Delta t \boldsymbol{A}$ is an M-matrix.

Consequently, the fully discrete MM-SUPG scheme satisfies the discrete maximum principle (DMP), i.e., the numerical solution preserves the bounds of the initial and boundary data.

Proof. The results follows immediately from the previous lemmas. □

Corollary 4.10.1. *Under the assumptions of Theorem 4.10, the discrete maximum principle (DMP) holds if the following practical conditions are satisfied:*

(i) ***Mesh condition:*** *The mesh is sufficiently aligned with the anisotropy of the problem in the sense that it satisfies the anisotropic nonobtuse angle condition (ANAC) with respect to both $\mathbb{D}_K$ and $\mathbb{K}_K$.*

(ii) ***Stabilization condition:*** *The stabilization parameter τ_K is chosen such that*

$$\tau_K \|b\|_{\infty,K} \lesssim h_K,$$

where h_K denotes a representative element size.

(iii) ***Balance condition:*** *The combined diffusion and stabilization effects dominate the convection contribution, i.e.,*

$$\rho_{\min}(\varepsilon\mathbb{D}_K) + \tau_K \rho_{\min}(\varepsilon\mathbb{K}_K) \gtrsim h_K \|b\|_{\infty,K} + \tau_K \|b\|_{\infty,K}^2.$$

(iv) ***Time step condition:*** *The time step Δt is chosen within a suitable range such that*

$$\Delta t \gtrsim \frac{1}{\theta} \frac{M_{ij}}{|A_{ij}|} \quad \textit{and} \quad \Delta t \lesssim \frac{1}{1-\theta} \frac{M_{ii}}{A_{ii}},$$

ensuring that $\widetilde{A} = M + \theta \Delta t A$ is an M-matrix and that $M - \Delta t(1-\theta)A$ remains nonnegative.

Under these conditions, the fully discrete MM-SUPG scheme satisfies the discrete maximum principle.

Remark 1. *The corollary highlights that the DMP is achieved through a balance between mesh alignment, stabilization, and time stepping. In particular, the interaction between the mesh metric and the SUPG stabilization plays a key role in controlling oscillations and preserving monotonicity.*

Remark 2. *Since we use implicit Euler method in time for all computations (that is, $\theta = 1$), the lower bound* (101) *for time step Δt is practically removed.*

Since two different metric tensors will be used on any given node, we follow [42] and take into consideration for every time step, we use metric intersection in order to generate the anisotropic mesh. This is equivalent to finding the largest ellipsoid included in the intersection of the two ellipses $\mathcal{E}$ and $\mathcal{E}'$, the geometric representations of symmetric, positive definite $\mathcal{M}$ and $\mathcal{M}'$, respectively (see Figure 10). By our assumptions on $\mathcal{M}$ and $\mathcal{M}'$, $\mathcal{N} = \mathcal{M}^{-1}\mathcal{M}'$ is diagonalizable by the similarity transform. Denote the shared normalized eigenvectors of $\mathcal{M}$ and $\mathcal{M}'$ by $\boldsymbol{\lambda}_1$ and $\boldsymbol{\lambda}_2$ and write

$$\mathcal{M}_{int} = \mathcal{M} \cap \mathcal{M}' = \Lambda^{-T} \mathrm{diag}\left(\max\{\alpha_1, \beta_1\}, \max\{\alpha_2, \beta_2\}\right) \Lambda^{-1}. \tag{125}$$

Figure 11 shows the coordinate transformation of a typical element K to $\tilde{K}$ and $\check{K}$.

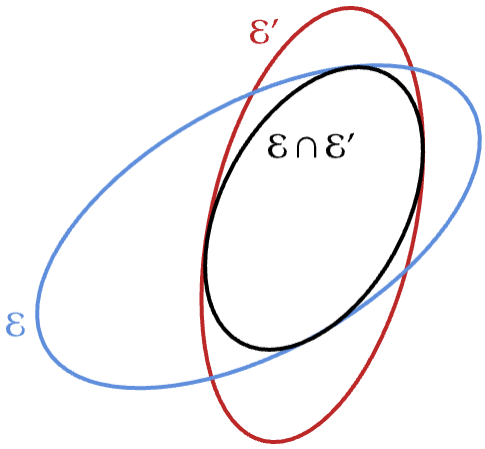


Figure 10: Geometric illustration of intersecting ellipses $\mathcal{E}$ and $\mathcal{E}'$, resulting in $\mathcal{E} \cap \mathcal{E}'$.

4.5 Numerical Results

In this section we present our numerical findings for using different metric tensors. We start with the metric tensors $\mathcal{M}_{DMP} = c_1\mathbb{D}^{-1}$ and $\mathcal{M}_{\mathbb{K}} = c_2\mathbb{K}^{-1}$, where c_1 and c_2 are arbitray constants. Define the metric tensor

$$\mathcal{M}_{int} = \mathcal{M}_{DMP} \cap \mathcal{M}_{\mathbb{K}}. \tag{126}$$

We use the implicit Euler method in time for all computations (that is, $\theta = 1$) and compare the results with the classic Galerkin method with SUPG using fixed mesh (FM-SUPG).

Example 4.1. Consider the manufactured solution on the domain $\Omega_T = [0,1]^2 \times [0, 0.3]$ to (1) given by

$$u(x, y, t) = (1 + \exp(C(x + y - t))^{-1}, \quad C > 0, \tag{127}$$

with diffusion tensor defined by

$$\mathbb{D} = \begin{pmatrix} 50.5 & 49.5 \\ 49.5 & 50.5 \end{pmatrix}, \tag{128}$$

and time-dependent flow given by

$$\boldsymbol{b}(x, y, t) = (x + t, y + t). \tag{129}$$

In the computations, we have used $\varepsilon = 10^{-6}$ and $\Delta t = 10^{-2}$ for this example. Based on the maximum principle, the solution should satisfy $\min_{\Omega_T} u = 0$.

The diffusion tensor has eigenvalues $\lambda_1 = 100$ and $\lambda_2 = 1$, which indicates that the diffusion in horizontal direction is 100 times faster than that in the vertical direction. For the given flow in (129), the tensor $\mathbb{K}$ defined in (58) is the same as $\mathbb{D}$. Therefore, the metric tensor $\mathcal{M}_{DMP}$ is the same as $\mathcal{M}_{\mathbb{K}}$ for this example. Hence, for this example, $\mathcal{M}_{int}$ is also the same as $\mathcal{M}_{DMP}$.

Figure 12 shows the side view of the numerical solutions using different metric tensors. As can be seen, the graph in Figure 12(a) shows significant artificial oscillation near the top shelf when using the fixed mesh approach. On the other hand, as shown in Figure 12(b), when using $\mathcal{M}_{int}$ or $\mathcal{M}_{DMP}$, the artificial oscillation at the top shelf is essentially eliminated.

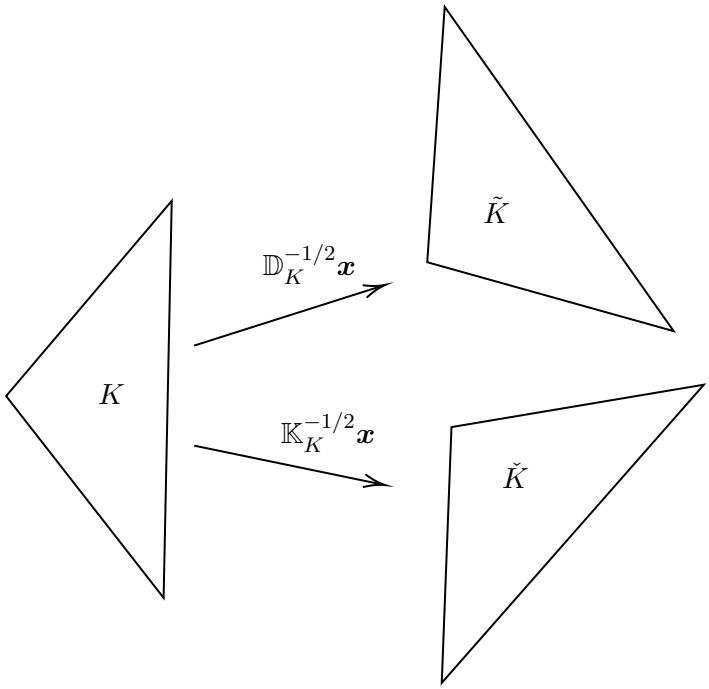


Figure 11: Coordinate transformation of a typical element K to $\tilde{K}$ and $\check{K}$.

Example 4.2. In this example, we consider the metric tensor $\mathcal{M}_{DMP+adapt}$ developed in [26] that is based on both DMP satisfaction as well as minimization of the H^1-seminorm of the interpolation error. Correspondingly, our intersection metric tensor is $\mathcal{M}_{int+adapt} = \mathcal{M}_{DMP+adap} \cap \mathcal{M}'$ where $\mathcal{M}' = \theta' \mathbb{K}^{-1}$. The anisotropic diffusion tensor is given by

$$\mathbb{D} = \begin{pmatrix} 1 & 1 \\ 1 & 3 \end{pmatrix}, \tag{130}$$

and the flow is given by

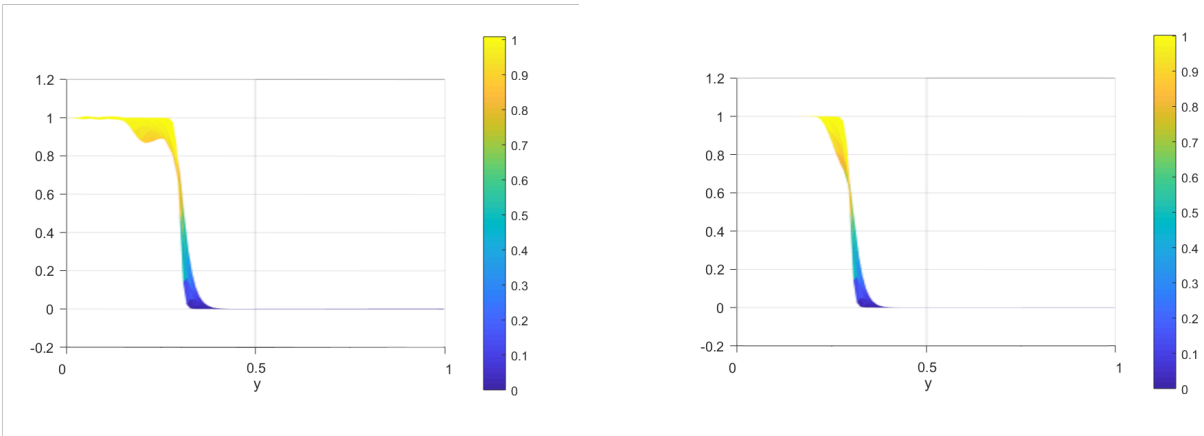
$$\boldsymbol{b}(x, y, t) = (x + y + t + 8, x + 3y + t + 16), \tag{131}$$

for $\Omega_T = [0, 1.5]^2 \times [0, 0.015]$. For the given $\boldsymbol{b}$ and $\mathbb{D}$, we have

$$\mathbb{K} = \begin{pmatrix} 2 & 4 \\ 4 & 10 \end{pmatrix}. \tag{132}$$

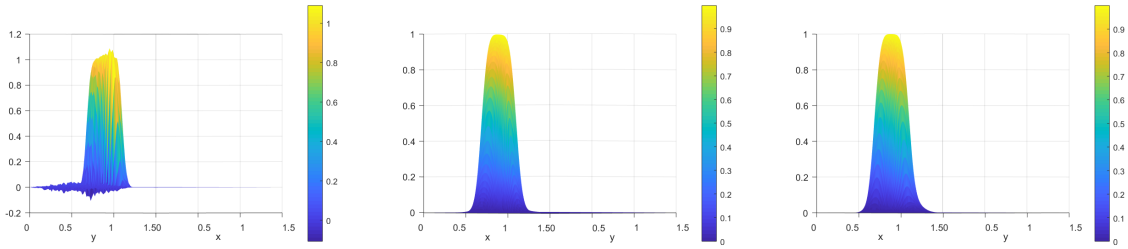
We first use $N = 20,402$ elements in the mesh and compute the numerical solutions for $\varepsilon = 10^{-3}$ with time step size $\Delta t = 10^{-4}$. Figure 13 shows the side view of the solution using fixed mesh, $\mathcal{M}_{DMP+adapt}$ mesh and $\mathcal{M}_{int+adapt}$ mesh, respectively. As can be seen, $\mathcal{M}_{DMP+adapt}$ mesh and $\mathcal{M}_{int+adapt}$ mesh provide comparable results and both perform much better than the fixed mesh SUPG approach.

To see the advantage of $\mathcal{M}_{int+adapt}$ mesh over $\mathcal{M}_{DMP+adapt}$ mesh, we compute the numerical solutions at different N values with larger time step $\Delta t = 10^{-3}$. The minimum values of the numerical solutions are summarized in Table 5. As can be seen, except on the course mesh, $\mathcal{M}_{int+adapt}$ provides the best results.



(a): Fixed mesh (b): $\mathcal{M}_{int}$

Figure 12: Example 4.1 – side view of solutions (scaled-up) for a fixed mesh and $\mathcal{M}_{int}$ with $N = 34322$, $\varepsilon = 10^{-6}$, and $\Delta t = 10^{-2}$.



(a): Fixed mesh (b): $\mathcal{M}_{DMP+adapt}$ (c): $\mathcal{M}_{int+adapt}$

Figure 13: Example 4.2 – side view with $\varepsilon = 10^{-3}$ for different metric tensors and $\Delta t = 10^{-4}$, $N = 20,402$.

5 Summary and Conclusions

In this paper, we have developed a moving mesh streamline upwind Petrov-Galerkin (MM-SUPG) method that combines the moving mesh partial differential equation (MMPDE) approach with the SUPG stabilization for convection-diffusion problems. The key feature of the proposed method is not only the combination of these two techniques, but also their interaction: the evolving mesh modifies the local element geometry and hence influences both the direction and magnitude of the SUPG stabilization, while the stabilization improves robustness in convection-dominated regimes.

The MM-SUPG method leverages the complementary roles of the two components. The moving mesh method concentrates mesh elements in regions with sharp gradients, improving resolution and computational efficiency, while the SUPG method stabilizes the numerical solution by adding streamline diffusion. The method is applicable to both time-dependent and time-independent problems, and is designed to handle both isotropic and anisotropic diffusion tensors.

For the isotropic diffusion case, the MM-SUPG method exhibits approximately $O(h^1)$ convergence in the H^1-seminorm and $O(h^{1.5})$ convergence in the L^2 norm (see Figures 6 and 7). Numerical experiments show that, for convection-dominated problems with sharp layers, moving mesh with the standard Galerkin method (MM-FEM) may still exhibit noticeable oscillations on moderately refined

Table 5: Example 4.2. Results for $\mathcal{M}_{int+adapt}$, $\mathcal{M}_{DMP+adapt}$, and fixed mesh for varying N with $\varepsilon = 10^{-3}$ and $\Delta t = 10^{-3}$.

N	Fixed Mesh	$\mathcal{M}_{DMP+adapt}$	$\mathcal{M}_{int+adapt}$
3362	-2.4050e-01	-2.3330e-03	-3.6297e-03
13122	-1.6762e-01	-1.3796e-03	-6.8634e-04
29282	-1.0677e-01	-9.3509e-04	-1.8547e-04

meshes (see Figure 1). The use of SUPG stabilization within the moving mesh framework reduces these oscillations and leads to improved stability. Comparisons in Examples 3.1 and 3.2 indicate that MM-SUPG improves robustness over fixed mesh SUPG (FM-SUPG), and provides comparable accuracy to MM-FEM in cases where mesh adaptation alone already captures the solution well.

For the anisotropic diffusion case, we have derived sufficient conditions on the mesh geometry and time step size under which the fully discrete scheme satisfies the discrete maximum principle (DMP). The analysis is based on quantitative bounds for the convection terms rather than sign arguments, and therefore avoids restrictive assumptions on the divergence of the velocity field. Based on the diffusivity tensor $\mathbb{D}$ and the flow vector $\boldsymbol{b}$, we introduce a weighted tensor $\mathbb{K}$ in (58), which serves as an analytical tool for incorporating the interaction between convection and diffusion. The tensor $\mathbb{K}$ is not guaranteed to be symmetric positive definite in general; thus, the DMP analysis is established under structural assumptions on $\boldsymbol{b}$ and $\mathbb{D}$ such that its elementwise form is of SPD type.

To further improve monotonicity properties, we construct a metric tensor $\mathcal{M}_{int}$ in (125) through metric intersection, combining information from both $\mathbb{D}$ and $\mathbb{K}$. Numerical results in Examples 4.1 and 4.2 show that meshes generated using $\mathcal{M}_{int}$ improve DMP behavior, as evidenced by reduced undershoots, while differences in standard error measures remain modest.

In summary, the MM-SUPG method provides a stable and effective approach for convection-dominated convection-diffusion problems by combining mesh adaptation and stabilization in a unified framework. Under suitable conditions on the mesh, problem data, and time step size, the method can produce solutions that satisfy the discrete maximum principle, particularly in anisotropic diffusion settings. The results demonstrate that the interaction between mesh adaptation and stabilization plays a key role in achieving both accuracy and improved monotonicity properties.

References

[1] K.W. Morton. *Revival: Numerical Solution Of Convection-Diffusion Problems (1996).* CRC Press, Boca Raton, Florida, 2019.

[2] 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. *Computer methods in applied mechanics and engineering*, 32(1-3):199–259, 1982.

[3] Y. Di, H. Xie, and X. Yin. Anisotropic meshes and stabilization parameter design of linear SUPG method for 2D convection-dominated convection–diffusion equations. *Journal of Scientific Computing*, 76(1):48–68, 2018.

[4] C. Erath and D. Praetorius. Optimal adaptivity for the SUPG finite element method. *Computer Methods in Applied Mechanics and Engineering*, 353:308–327, 2019.

[5] M. Picasso and V. Prachittham. An adaptive algorithm for the Crank–Nicolson scheme applied to a time-dependent convection–diffusion problem. *Journal of computational and applied mathematics*, 233(4):1139–1154, 2009.

[6] S. Ganesan and S. Srivastava. ALE-SUPG finite element method for convection–diffusion problems in time-dependent domains: Conservative form. *Applied Mathematics and Computation*, 303:128–145, 2017.

[7] T.P. Fries, H. M., et al. *A review of Petrov-Galerkin stabilization approaches and an extension to meshfree methods*. Citeseer, Technical University Braunschweig, 2004.

[8] L.P. Franca, S.L. Frey, and T.J.R. Hughes. Stabilized finite element methods: I. application to the advective-diffusive model. *Computer Methods in Applied Mechanics and Engineering*, 95(2):253–276, 1992.

[9] T. Gelhard, G. Lube, M.A. Olshanskii, and J.H. Starcke. Stabilized finite element schemes with LBB-stable elements for incompressible flows. *Journal of computational and applied mathematics*, 177(2):243–267, 2005.

[10] M. Augustin, A. Caiazzo, A. Fiebach, J. Fuhrmann, V. John, A. Linke, and R. Umla. An assessment of discretizations for convection-dominated convection–diffusion equations. *Computer Methods in Applied Mechanics and Engineering*, 200(47-48):3395–3409, 2011.

[11] H. Nguyen, M. Gunzburger, L. Ju, and J. Burkardt. Adaptive anisotropic meshing for steady convection-dominated problems. *Computer methods in applied mechanics and engineering*, 198(37-40):2964–2981, 2009.

[12] M. Bause and P. Knabner. Uniform error analysis for Lagrange–Galerkin approximations of convection-dominated problems. *SIAM Journal on Numerical Analysis*, 39(6):1954–1984, 2002.

[13] M. Bause and P. Knabner. On uniform convergence rates for Eulerian and Lagrangian finite element approximations of convection-dominated diffusion problems. *Calcolo*, 41(1):1–26, 2004.

[14] G.F. Carey, M. Anderson, B. Carnes, and B. Kirk. Some aspects of adaptive grid technology related to boundary and interior layers. *Journal of computational and applied mathematics*, 166(1):55–86, 2004.

[15] P.P. Chinchapatnam, K. Djidjeli, and P.B. Nair. Unsymmetric and symmetric meshless schemes for the unsteady convection–diffusion equation. *Computer methods in applied mechanics and engineering*, 195(19-22):2432–2453, 2006.

[16] V. John and E. Schmeyer. Finite element methods for time-dependent convection–diffusion–reaction equations with small diffusion. *Computer methods in applied mechanics and engineering*, 198(3-4):475–494, 2008.

[17] V. John and E. Schmeyer. On finite element methods for 3D time-dependent convection-diffusion-reaction equations with small diffusion. In *BAIL 2008-boundary and interior layers*, pages 173–181. Springer, Berlin, Heidelberg, 2009.

[18] V. John and P. Knobloch. On spurious oscillations at layers diminishing (SOLD) methods for convection–diffusion equations: Part I–a review. *Computer methods in applied mechanics and engineering*, 196(17-20):2197–2215, 2007.

[19] Y. Cheng, Y. Mei, and H.G. Roos. The local discontinuous Galerkin method on layer-adapted meshes for time-dependent singularly perturbed convection-diffusion problems. *arXiv preprint arXiv:2012.03560*, 2020.

[20] F. Wang, W. Chen, A. Tadeu, and C.G. Correia. Singular boundary method for transient convection–diffusion problems with time-dependent fundamental solution. *International Journal of Heat and Mass Transfer*, 114:1126–1134, 2017.

[21] H. von Wahl, T. Richter, and C. Lehrenfeld. An unfitted Eulerian finite element method for the time-dependent stokes problem on moving domains. *arXiv preprint arXiv:2002.02352*, 2020.

[22] P. Hansbo, M.G. Larson, and S. Zahedi. A cut finite element method for coupled bulk-surface problems on time-dependent domains. *Computer Methods in Applied Mechanics and Engineering*, 307:96–116, 2016.

[23] S. Zahedi. A space-time cut finite element method with quadrature in time. In *Geometrically Unfitted Finite Element Methods and Applications*, pages 281–306. Springer, Gewerbestrasse, Switzerland, 2017.

[24] S. Badia, F. Verdugo, and A.F. Martín. The aggregated unfitted finite element method for elliptic problems. *Computer Methods in Applied Mechanics and Engineering*, 336:533–553, 2018.

[25] X. Zhang, and X. Xu. Moving mesh method with variational multiscale finite element method for convection-diffusion-reaction equations. *Engineering with Computers*, 40:1943–1965, 2024.

[26] X. Li and W. Huang An anisotropic mesh adaptation method for the finite element solution of heterogeneous anisotropic diffusion problems. *Journal of Computational Physics*, 229(21):8072–8094, 2010.

[27] X. Li and W. Huang. Maximum principle for the finite element solution of time-dependent anisotropic diffusion problems. *Numerical Methods for Partial Differential Equations*, 29(6):1963–1985, 2013.

[28] C. Lu, W. Huang, and J. Qiu. Maximum principle in linear finite element approximations of anisotropic diffusion–convection–reaction problems. *Numerische Mathematik*, 127(3):515–537, 2014.

[29] T. Linß and M. Stynes. Numerical methods on Shishkin meshes for linear convection–diffusion problems. *Computer methods in applied mechanics and engineering*, 190(28):3527–3542, 2001.

[30] R. Codina. Stabilization of incompressibility and convection through orthogonal sub-scales in finite element methods. *Computer Methods in Applied Mechanics and Engineering*, 190(13-14):1579–1599, 2000.

[31] R. Codina. On stabilized finite element methods for linear systems of convection–diffusion-reaction equations. *Computer Methods in Applied Mechanics and Engineering*, 188(1-3):61–82, 2000.

[32] L.P. Franca and F. Valentin. On an improved unusual stabilized finite element method for the advective–reactive–diffusive equation. *Computer Methods in Applied Mechanics and Engineering*, 190(13-14):1785–1800, 2000.

[33] T. Knopp, G. Lube, and G. Rapin. Stabilized finite element methods with shock capturing for advection–diffusion problems. *Computer methods in applied mechanics and engineering*, 191(27-28):2997–3013, 2002.

[34] G. Lube and G. Rapin. Residual-based stabilized higher-order FEM for advection-dominated problems. *Computer methods in applied mechanics and engineering*, 195(33-36):4124–4138, 2006.

[35] W. Huang and R.D. Russell. *Adaptive moving mesh methods*. Springer, New York, NY, 2011.

[36] A. Kolasinski and W. Huang. A new functional for variational mesh generation and adaptation based on equidistribution and alignment conditions. *Computers & Mathematics with Applications*, 75(6):2044–2058, 2018.

[37] F. Zhang, W. Huang, X. Li, and S. Zhang. Moving mesh finite element simulation for phase-field modeling of brittle fracture and convergence of Newton's iteration. *Journal of Computational Physics*, 356:127–149, 2018.

[38] S. Xie, P. Zhu, and X. Wang. Error analysis of weak galerkin finite element methods for time-dependent convection–diffusion equations. *Applied Numerical Mathematics*, 137:19–33, 2019.

[39] L.C. Evans. Partial differential equations. *Graduate studies in mathematics*, 19(2), 1998.

[40] J.H. Brandts, S. Korotov, and M. Křížek. The discrete maximum principle for linear simplicial finite element approximations of a reaction–diffusion problem. *Linear Algebra and its Applications*, 429(10):2344–2357, 2008.

[41] G. Stoyan. On maximum principles for monotone matrices. *Linear algebra and its applications*, 78:147–161, 1986.

[42] F. Alauzet and G. Olivier. Extension of metric-based anisotropic mesh adaptation to time-dependent problems involving moving geometries. in *49th AIAA Aerospace Sciences Meeting including the New Horizons Forum and Aerospace Exposition*, p. 896, 2011.