FORMULATION AND ANALYSIS OF A PARAMETER-FREE STABILIZED FINITE ELEMENT METHOD*

P. BOCHEV†, M. PEREGO† , AND K. PETERSON†

Abstract. We present a stabilized finite element method for the scalar advection-diffusion equation, which does not require tunable mesh-dependent parameters. Stabilization is achieved by using diffusive fluxes extracted from an edge element lifting of Scharfetter-Gummel edge fluxes into the elements. Although the method is formally first-order accurate, qualitative numerical studies suggest that it occupies a middle ground between an artificial diffusion and a streamline-upwind Petrov-Galerkin formulations. The method is substantially less dissipative than the former, while having much smaller overshoots and undershoots than the latter.

Key words. Advection-diffusion, stabilization, Scharfetter-Gummel upwinding, finite elements, edge elements.

AMS subject classifications. 65N30, 65N12, 65N15, 65L11,65L11, 65L60

1. Introduction.

We consider the scalar advection-diffusion equation

$$\begin{cases} -\nabla \cdot F(\phi) = f & \text{in } \Omega \\ F(\phi) = (\varepsilon \nabla \phi - \mathbf{u}\phi) & \text{in } \Omega \end{cases} \quad \text{and} \quad \phi = g \text{ on } \Gamma \tag{1.1}$$

where ε is a diffusion coefficient, $\mathbf{u}$ is the advective velocity, and f and g are given functions. When ε is small relative to $\mathbf{u}$, solutions of (1.1) can develop internal and/or boundary layers. If the grid is not fine enough to resolve these layers, Galerkin solutions of (1.1) exhibit spurious oscillations.

To stabilize the solution one can use a variety of tools ranging from artificial diffusion [14] to consistently stabilized methods such as SUPG [12, 13], and multiscale and enriched methods [11, 18]. However, most if not all of the existing stabilized methods for (1.1) require a mesh-dependent stabilization parameter. The choice of this parameter is critical for the accuracy and stability of the corresponding finite element solution. Yet, finding the best possible stabilization parameter for a given problem is difficult and remains an open question [15, 7]. The principal reasons for this are (i) the dependence of the stabilization parameter on constants that are known exactly only in special cases [10], and (ii) the fact that different solution features, such as internal and boundary layers may require different stabilization techniques; see [15]. As a result, stabilization parameters may have to be individually tailored to different applications using heuristic arguments [5], numerical calibration or a combination thereof.

This paper presents and studies a new stabilized finite element method for (1.1), which does not require tunable mesh-dependent stabilization parameters. Stabilization is achieved by augmenting the standard Galerkin formulation of (1.1) with an artificial diffusion tensor defined by a product of two diffusive fluxes. We extract these fluxes from an approximation $F_E \in H(curl, \Omega)$ of the total flux F, defined by an edge element lifting of Scharfetter-Gummel edge fluxes into the elements [3, 4]. Previously, we used $F_E(\phi_h)$ to replace the standard nodal flux $F(\phi_h)$ in a Galerkin [3] and a control volume finite element (CVFEM) [4] formulations of the governing equations (1.1). These methods provide multidimensional extensions of the Scharfetter-Gummel scheme [17] to unstructured grids and have been implemented in Sandia's device modeling code Charon with encouraging results.

*Sandia National Laboratories is a multi-program laboratory managed and operated by Sandia Corporation, a wholly owned subsidiary of Lockheed Martin Corporation, for the U.S. Department of Energys National Nuclear Security Administration under contract DE-AC04-94AL85000

†Computational Mathematics, Sandia National Laboratories, Mail Stop 1320 Albuquerque, New Mexico, 87185-1320 ({pbboche,mperego,kjpeter}@sandia.gov)

In this paper the edge element flux F_E is a starting point for the development of a new stabilized Galerkin formulation that improves upon the robustness and accuracy of the method in [3]. A careful analysis reveals that F_E comprises a projection of the nodal flux $F(\phi_h)$ onto an edge element space, and a stabilizing diffusive flux. The pairing of the latter with the gradient of a nodal finite element function ϕ_h produces a non-symmetric diffusion tensor whose stabilizing effect may deteriorate on unstructured grids. To correct this drawback we (i) discard the part of the F_E that corresponds to the projection of $F(\phi_h)$ onto the edge element space, and (ii) use the remaining diffusive flux to define a symmetric diffusion tensor. In a nutshell, the new method comprises a standard nodal Galerkin formulation of (1.1), augmented by this symmetric artificial diffusion tensor. Because edge elements are used only locally, assembly of the diffusion tensor does not require global edge data structured and can be easily incorporated into an existing Galerkin code for (1.1).

The rest of this section introduces the notation. Sections 2–3 present the formulation and the analysis of the method, respectively. Section 4 illustrates the method numerically, including a demonstration of its improved stability on unstructured grids.

1.1. Notation. In this paper $\Omega \subset \Re^n$, $n = 2, 3$ is a bounded domain with Lipschitz-continuous boundary $\Gamma = \partial\Omega$, $W^{k,p}(\Omega)$ is a Sobolev space of order k, $W_0^{k,p}(\Omega)$ is the subspace of its functions with vanishing trace, $L^p(\Omega) = W^{0,p}(\Omega)$, and $H^k(\Omega) = W^{k,2}(\Omega)$. The space $H(curl, \Omega)$ contains vector fields in $L^2(\Omega)^n$ whose curl belongs in $L^2(\Omega)^{2n-3}$. The meaning of the symbol $|\cdot|$ depends on the context and can be Euclidean length, semi-norm, domain measure, or cardinality of a finite set. We use lower case bold face for vectors, upper case Roman for matrices, $\underline{\lambda}(A)$ and $\overline{\lambda}(A)$ are the smallest and largest eigenvalues of A, respectively, and $\kappa(A)$ is the condition number of A. The versors of a Cartesian coordinate system are $\mathbf{i} = (1, 0, 0)$, $\mathbf{j} = (0, 1, 0)$, and $\mathbf{k} = (0, 0, 1)$ in $\Re^3$.

We use shape-regular finite element partitions $K(\Omega)$ of Ω into elements $\boldsymbol{k}$ with size $h_{\boldsymbol{k}}$ The average element size is denoted by h. For brevity we restrict attention to triangles and quadrilaterals in 2D and tetrahedrons and hexahedrons in 3D. We assume that the simplicial elements are affine, the quadrilaterals are bilinear images of the unit square, and the hexahedrals are trilinear images of the unit cube. The symbols $V(\star)$ and $E(\star)$ stand for the sets of all vertices and edges in $K(\Omega)$ belonging to an entity $\star$. For instance, $V(\Omega)$ is the set of all vertices, $E(\Omega)$ is the set of all edges, $V(\boldsymbol{k})$ are the vertices of element $\boldsymbol{k}$, $E(\boldsymbol{v}_i)$ are all edges having $\boldsymbol{v}_i$ as a vertex, and so on. We label edges by a multi-index $\alpha = (\alpha_1, \alpha_2)$ comprising the indices of their endpoints, i.e., $\boldsymbol{e}_\alpha$ is an edge with endpoints $\boldsymbol{v}_{\alpha_1}$ and $\boldsymbol{v}_{\alpha_2}$. We also need the node-to-edge incidence matrix G with element $g_{\alpha,j}$ and dimension $|E(\Omega)| \times |V(\Omega)|$. Assuming that every edge $\boldsymbol{e}_\alpha$ in the mesh is oriented by choosing the order of its vertices, $g_{\alpha,j} = 0$ if $\boldsymbol{v}_j$ is not a vertex of $\boldsymbol{e}_\alpha$ and

$$g_{\alpha,j} = \begin{cases} -1 & \text{if } \boldsymbol{v}_j \text{ is the first vertex of } \boldsymbol{e}_\alpha \\ 1 & \text{if } \boldsymbol{v}_j \text{ is the second vertex of } \boldsymbol{e}_\alpha \end{cases} \tag{1.2}$$

For most "reasonable" finite element partitions, e.g., meshes satisfying the assumptions of [6, p.51] the matrix G has a one dimensional kernel comprising the constant vector. Thus, G can be interpreted as a "topological" gradient operator that depends only on the mesh connectivity.

The midpoint, the length, and the unit tangent of an oriented edge $\boldsymbol{e}_\alpha$ are

$$\boldsymbol{m}_\alpha = \frac{\boldsymbol{v}_{\alpha_1} + \boldsymbol{v}_{\alpha_2}}{2}\,, \quad h_\alpha = |\boldsymbol{v}_{\alpha_1} - \boldsymbol{v}_{\alpha_2}|\,, \quad \text{and} \quad \mathbf{t}_\alpha = g_{\alpha,\alpha_1} \frac{\boldsymbol{v}_{\alpha_1} - \boldsymbol{v}_{\alpha_2}}{|\boldsymbol{v}_{\alpha_1} - \boldsymbol{v}_{\alpha_2}|}\,,$$

respectively. Definition of $\mathbf{t}_\alpha$ implies that it always points towards the second vertex of $\boldsymbol{e}_\alpha$.

In what follows, $N^h(\Omega)$ is the isoparametric C^0 piecewise linear, bilinear or trilinear *nodal* finite element space and $E^h(\Omega)$ is the lowest-order Nedelec *edge element* space [16]. $N_0^h(\Omega)$ is the

subspace of all functions in $N^h(\Omega)$ that vanish on Γ. The standard nodal basis $\{N_i\}$, $\boldsymbol{v}_i \in V(\Omega)$ of $N^h(\Omega)$ has the property $N_i(\boldsymbol{v}_j) = \delta_i^j$. The coefficient vector of $\phi_h \in N^h(\Omega)$ is $\boldsymbol{\phi} = (\boldsymbol{\phi}_i, \ldots, \boldsymbol{\phi}_m)$, $m = |V(\Omega)|$. The standard basis of $E^h(\Omega)$ is $\{\vec{W}_\xi\}$, $\boldsymbol{e}_\xi \in E(\Omega)$ such that

$$\int_{\boldsymbol{e}_\eta} \vec{W}_\xi \cdot \mathbf{t}_\eta d\ell = \delta_\xi^\eta \quad \text{and} \quad \vec{W}_\xi \cdot \mathbf{t}_\eta \Big|_{\boldsymbol{e}_\eta} = const = \frac{\delta_\xi^\eta}{h_\eta}\,. \tag{1.3}$$

Since $\vec{W}_\xi \cdot \mathbf{t}_\xi > 0$ the orientation of $\vec{W}_\xi$ matches the orientation of its associated edge $\boldsymbol{e}_\xi$. $N^h(\Omega)$ and $E^h(\Omega)$ belong to an exact sequence [1] and $\nabla N^h(\Omega) \subset E^h(\Omega)$. In particular, there holds

$$\nabla N_i = \sum_{\boldsymbol{e}_\xi \in E(\boldsymbol{v}_i)} g_{\xi,i} \vec{W}_\xi \quad \text{and} \quad \nabla \phi_h = \sum_{\boldsymbol{e}_\xi \in E(\Omega)} (g_{\xi,\xi_1} \boldsymbol{\phi}_{\xi_1} + g_{\xi,\xi_2} \boldsymbol{\phi}_{\xi_2}) \vec{W}_\xi = \sum_{\boldsymbol{e}_\xi \in E(\Omega)} (\boldsymbol{g}_\xi \boldsymbol{\phi}) \vec{W}_\xi\,, \tag{1.4}$$

where $\boldsymbol{g}_\xi$ is the row of G corresponding to edge $\boldsymbol{e}_\xi$. The last identity implies the following factorization of the weak Laplace operator; see e.g. [2, p.108]:

$$\int_\Omega \nabla \phi_h \cdot \nabla \varphi_h \, dx = \boldsymbol{\phi}^T (G^T M G) \boldsymbol{\varphi}\,, \tag{1.5}$$

where M is the Gramm matrix of the edge element basis $\{\vec{W}_\xi\}$. Since G has one-dimensional nullspace spanned by the constant vector, it follows that for $\phi_h \in N_0^h(\Omega)$ with coefficient vector $\boldsymbol{\phi}$,

$$G\boldsymbol{\phi} = 0 \quad \text{if and only if} \quad \phi_h = 0\,. \tag{1.6}$$

We recall the nodal interpolation operator $\mathcal{I}_N : H^1(\Omega) \cap C^0(\Omega) \mapsto N^h(\Omega)$,

$$\mathcal{I}_N(\phi) = \sum_{\boldsymbol{v}_i \in V(\Omega)} \phi(\boldsymbol{v}_i) N_i(\boldsymbol{x})\,, \tag{1.7}$$

and, for $p > 2$, the edge interpolation operator $\mathcal{I}_E : H(curl, \Omega) \cap (W^{1,p}(\Omega))^n \mapsto E^h(\Omega)$,

$$\mathcal{I}_E(\mathbf{u}) = \sum_{\boldsymbol{e}_\xi \in E(\Omega)} \vec{W}_\xi \int_{\boldsymbol{e}_\xi} \mathbf{u} \cdot \mathbf{t}_\xi d\ell\,. \tag{1.8}$$

We will also need the local interpolation result [9, Theorem 1.103]

$$|\mathcal{I}_N(\phi) - \phi|_{m,p,\boldsymbol{k}} \le C h_{\boldsymbol{k}}^{l+1-m} |\phi|_{l+1,p,\boldsymbol{k}}\,, \tag{1.9}$$

which holds for all $\phi \in W^{l+1,p}(\boldsymbol{k})$ with $\boldsymbol{k} \in K(\Omega)$ and $1 \le p \le \infty$, $0 \le l \le 1$, and the fact that

$$\|\mathcal{I}_N(\phi)\|_1 \le C \|\phi\|_2 \quad \forall \phi \in H^2(\Omega)\,. \tag{1.10}$$

2. Formulation. Consider the Galerkin method for (1.1): find $\phi_h \in N^h(\Omega)$ such that

$$a(\phi_h, \varphi_h) = b(\varphi_h) \quad \forall \varphi_h \in N_0^h(\Omega)\,, \tag{2.1}$$

where the bilinear form $a(\cdot,\cdot)$ and the linear functional $b(\cdot)$ are given by

$$a(\phi_h, \varphi_h) = \int_\Omega F(\phi_h) \cdot \nabla \varphi_h \, dV \quad \text{and} \quad b(\varphi_h) = \int_\Omega f \varphi_h dV\,, \tag{2.2}$$

respectively. We will stabilize (2.1) using a diffusive flux $\widetilde{\Theta}(\phi_h)$ extracted from an edge element approximation $F_E(\phi_h)$ of the total flux $F(\phi_h)$. We define the former by an edge element lifting of one-dimensional, Scharfetter-Gummel edge fluxes F_α into the elements.

2.1. The $H(curl,\Omega)$ edge element flux. To define the edge fluxes F_α and their $H(\mathrm{curl}\,,\Omega)$ lifting F_E, we temporarily adopt the convention that the first vertex of $\boldsymbol{e}_\alpha$ is always $\boldsymbol{v}_{\alpha_1}$ and so, $g_{\alpha,\alpha_1}=-1$ and $g_{\alpha,\alpha_2}=1$. Then, along $\boldsymbol{e}_\alpha$ we consider the following problem:

$$\left\{\begin{array}{c} -\dfrac{d}{ds}\left(\bar{\varepsilon}_\alpha\dfrac{d\phi(s)}{ds}-\bar{u}_\alpha\phi(s)\right)=0 \quad \text{for } 0<s<h_\alpha \\ \phi(0)=\phi_{\alpha_1} \quad \text{and} \quad \phi(h_\alpha)=\phi_{\alpha_2}\end{array}\right. \tag{2.3}$$

where the unknown nodal coefficients ϕ_{α_1} and ϕ_{α_2} of ϕ_h specify the boundary data in (2.3), and

$$\bar{u}_\alpha=\frac{1}{h_\alpha}\int_{\boldsymbol{e}_\alpha}\mathbf{u}\cdot\mathbf{t}_\alpha d\ell \quad \text{and} \quad \bar{\varepsilon}_\alpha=\frac{1}{h_\alpha}\int_{\boldsymbol{e}_\alpha}\varepsilon d\ell$$

are the mean edge velocity and diffusion, respectively. The exact solution of (2.3) is

$$\phi(s)=\frac{\exp(2p_\alpha)\boldsymbol{\phi}_{\alpha_1}-\boldsymbol{\phi}_{\alpha_2}}{\exp(2p_\alpha)-1}+\frac{\boldsymbol{\phi}_{\alpha_2}-\boldsymbol{\phi}_{\alpha_1}}{\exp(2p_\alpha)-1}\exp(s\bar{u}_\alpha/\bar{\varepsilon}_\alpha)\,; \qquad p_\alpha=\frac{\bar{u}_\alpha h_\alpha}{2\bar{\varepsilon}_\alpha}\,. \tag{2.4}$$

The value of p_α relates the rates of advection and diffusion along $\boldsymbol{e}_\alpha$ and so we call it *the edge Péclet number*. The edge flux corresponding to the exact solution (2.4) is given by

$$F_\alpha=h_\alpha\left(\bar{\varepsilon}_\alpha\frac{d\phi(s)}{ds}-\bar{u}_\alpha\phi(s)\right)=h_\alpha\bar{u}_\alpha\frac{\boldsymbol{\phi}_{\alpha_2}-\exp(2p_\alpha)\boldsymbol{\phi}_{\alpha_1}}{\exp(2p_\alpha)-1}\,. \tag{2.5}$$

After discarding the temporary orientation convention, multiplying and dividing (2.5) by $\exp(-p_\alpha)$, and performing some simple algebraic manipulations, the edge fluxes assume the form:

$$F_\alpha=\sigma_\alpha\boldsymbol{g}_\alpha\boldsymbol{\phi}-\frac{h_\alpha\bar{u}_\alpha}{2}(\phi_{\alpha_1}+\phi_{\alpha_2}); \quad \sigma_\alpha=\frac{h_\alpha\bar{u}_\alpha}{2}\coth(p_\alpha)=\bar{\varepsilon}_\alpha p_\alpha\coth(p_\alpha)\,. \tag{2.6}$$

The lifting of the edge fluxes (2.6) into the edge element space $E^h(\Omega)$ defines the $H(curl,\Omega)$ flux

$$F_E(\phi_h)=\sum_{\boldsymbol{e}_\xi\in E(\Omega)}F_\xi\vec{W}_\xi=\sum_{\boldsymbol{e}_\xi\in E(\Omega)}\Big[\sigma_\xi\boldsymbol{g}_\xi\boldsymbol{\phi}-\frac{h_\xi\bar{u}_\xi}{2}(\phi_{\xi_1}+\phi_{\xi_2})\Big]\vec{W}_\xi\,. \tag{2.7}$$

2.2. The stabilizing diffusive flux. We start by examining the structure of F_E.

LEMMA 2.1. *Let $\phi_h\in N^h(\Omega)$. Then,*

$$F_E(\phi_h)=\mathcal{I}_E(\varepsilon\nabla\phi_h-\mathbf{u}_E\phi_h)+\Theta(\phi_h)\,, \tag{2.8}$$

where $\mathbf{u}_E=\mathcal{I}_E\mathbf{u}$ and

$$\Theta(\phi_h)=\sum_{\boldsymbol{e}_\xi\in E(\Omega)}\theta_\xi(\boldsymbol{g}_\xi\boldsymbol{\phi})\vec{W}_\xi \quad \textit{with} \quad \theta_\xi=\sigma_\xi-\bar{\varepsilon}_\xi=\bar{\varepsilon}_\xi(p_\xi\coth(p_\xi)-1) \tag{2.9}$$

is a stabilizing diffusive flux.

Proof. Adding and subtracting $\bar{\varepsilon}_\xi\boldsymbol{g}_\xi\boldsymbol{\phi}$ to (2.6) allows us to write the $H(\mathrm{curl}\,,\Omega)$ flux as

$$F_E=\sum_{\boldsymbol{e}_\xi\in E(\Omega)}\Big[\bar{\varepsilon}_\xi\boldsymbol{g}_\xi\boldsymbol{\phi}-\frac{h_\xi\bar{u}_\xi}{2}(\phi_{\xi_1}+\phi_{\xi_2})\Big]\vec{W}_\xi+\sum_{\boldsymbol{e}_\xi\in E(\Omega)}(\sigma_\xi-\bar{\varepsilon}_\xi)(\boldsymbol{g}_\xi\boldsymbol{\phi})\vec{W}_\xi\,.$$

The second term is the stabilizing diffusive flux Θ. To complete the proof it remains to show that the first term equals $\mathcal{I}_E(\varepsilon\nabla\phi_h - \mathbf{u}_E\phi_h)$. Using (1.8) and the exactness (1.4) yields

$$\begin{aligned}\mathcal{I}_E(\varepsilon\nabla\phi_h) &= \sum_{\boldsymbol{e}_\xi\in E(\Omega)}\Big[\int_{\boldsymbol{e}_\xi}\varepsilon\nabla\phi_h\cdot\mathbf{t}_\xi d\ell\Big]\vec{W}_\xi = \sum_{\boldsymbol{e}_\xi\in E(\Omega)}\Big[\int_{\boldsymbol{e}_\xi}\varepsilon\Big(\sum_{\boldsymbol{e}_\eta\in E(\Omega)}(\boldsymbol{g}_\eta\boldsymbol{\phi})\vec{W}_\eta\Big)\cdot\mathbf{t}_\xi d\ell\Big]\vec{W}_\xi\\ &= \sum_{\boldsymbol{e}_\xi\in E(\Omega)}\Big[\int_{\boldsymbol{e}_\xi}\frac{\varepsilon}{h_\xi}(\boldsymbol{g}_\xi\boldsymbol{\phi})d\ell\Big]\vec{W}_\xi = \sum_{\boldsymbol{e}_\xi\in E(\Omega)}\bar{\varepsilon}_\xi(\boldsymbol{g}_\xi\boldsymbol{\phi})\vec{W}_\xi\,.\end{aligned}$$

From the definition of $\bar{u}_\alpha$ and the linearity of ϕ_h along any edge, it follows that

$$\mathbf{u}_E = \sum_{\boldsymbol{e}_\eta\in E(\Omega)}\Big[\int_{\boldsymbol{e}_\eta}\mathbf{u}\cdot\mathbf{t}_\eta d\ell\Big]\vec{W}_\eta = \sum_{\boldsymbol{e}_\eta\in E(\Omega)}h_\eta\bar{u}_\eta\vec{W}_\eta \quad\text{and}\quad \int_{\boldsymbol{e}_\xi}\phi_h d\ell = \frac{h_\xi}{2}(\boldsymbol{\phi}_{\xi_1}+\boldsymbol{\phi}_{\xi_2})\,,$$

respectively. As a result, using (1.3)

$$\begin{aligned}\mathcal{I}_E(\mathbf{u}_E\phi_h) &= \sum_{\boldsymbol{e}_\xi\in E(\Omega)}\Big[\int_{\boldsymbol{e}_\xi}(\mathbf{u}_E\phi_h)\cdot\mathbf{t}_\xi d\ell\Big]\vec{W}_\xi = \sum_{\boldsymbol{e}_\xi\in E(\Omega)}\left[\int_{\boldsymbol{e}_\xi}\Big[\sum_{\boldsymbol{e}_\eta\in E(\Omega)}h_\eta\bar{u}_\eta\vec{W}_\eta\Big]\cdot\mathbf{t}_\xi\phi_h\,d\ell\right]\vec{W}_\xi\\ &= \sum_{\boldsymbol{e}_\xi\in E(\Omega)}\Big[\bar{u}_\xi\int_{\boldsymbol{e}_\xi}\phi_h\,d\ell\Big]\vec{W}_\xi = \sum_{\boldsymbol{e}_\xi\in E(\Omega)}\frac{h_\xi\bar{u}_\xi}{2}(\boldsymbol{\phi}_{\xi_1}+\boldsymbol{\phi}_{\xi_2})\vec{W}_\xi\,,\end{aligned}$$

which completes the proof. ∎

The following result confirms that Θ is responsible for the stabilizing effect of F_E.

LEMMA 2.2. *For $\varepsilon>0$ and any $\mathbf{u}$ there holds*

$$\max\Big\{\bar{\varepsilon}_\alpha,\frac{h_\alpha|u_\alpha|}{2}\Big\}\le\sigma_\alpha\le\bar{\varepsilon}_\alpha+\frac{h_\alpha|u_\alpha|}{2}\quad\textit{and}\quad\max\Big\{0,\frac{h_\alpha|u_\alpha|}{2}-\bar{\varepsilon}_\alpha\Big\}\le\theta_\alpha\le\frac{h_\alpha|u_\alpha|}{2}\,.\tag{2.10}$$

In particular, $\theta_\alpha\ge 0$, $\theta_\alpha\to 0$ in the diffusive limit, and $\theta_\alpha\to h_\alpha\bar{u}_\alpha/2$ in the advective limit.

Proof. The lemma follows from the inequality

$$\max\{1,|x|\}\le x\coth(x)\le 1+|x|\quad\forall x\in\Re\,.$$

∎

2.3. The stabilized Galerkin formulation. In [3] we considered the formulation

$$\int_\Omega F_E(\phi_h)\cdot\nabla\varphi_h\,dV = b(\varphi_h)\quad\forall\varphi_h\in N_0^h(\Omega)\,,\tag{2.11}$$

which stabilizes (2.1) by replacing the nodal flux $F(\phi_h)$ with the edge element flux $F_E(\phi_h)$. The decomposition of F_E in Lemma 2.1 prompts a better stabilization strategy. Specifically, Lemma 2.2 asserts that in the advective limit the stabilizing effect of F_E is solely due to the diffusive flux Θ, whereas the first part of F_E is just an approximation of the (unstable) nodal flux $F(\phi_h)$.

Since this part is not important for the stabilization we may as well switch back to $F(\phi_h)$, i.e., the standard Galerkin formulation (2.1), and use Θ to stabilize the latter. This would result in the following modification of (2.11): seek $\phi_h\in N^h(\Omega)$ such that

$$a(\phi_h,\varphi_h)+\int_\Omega\Theta(\phi_h)\cdot\nabla\varphi_h\,dV = b(\varphi_h)\quad\forall\varphi_h\in N_0^h(\Omega)\,.\tag{2.12}$$

However, the pairing of Θ and $\nabla\varphi_h$ produces a non-symmetric artificial diffusion. Numerical examples in Section 4 show that on unstructured grids this may reduce the stability of (2.12) and (2.11). Thus, we consider instead the following symmetric diffusion form:

$$Q(\phi_h,\varphi_h)=\int_\Omega \widetilde{\Theta}_h(\phi_h)\cdot\widetilde{\Theta}_h(\varphi_h)\,dV\,, \tag{2.13}$$

where

$$\widetilde{\Theta}_h(\phi_h)=\sum_{\boldsymbol{e}_\xi\in E(\Omega)}\widetilde{\theta}_\xi\,(\boldsymbol{g}_\xi\boldsymbol{\phi})\vec{W}_\xi\quad\text{and}\quad\widetilde{\theta}_\xi=\sqrt{\theta_\xi}\,.$$

Using (2.13) to stabilize (2.1) yields the new method, which is a "symmetrized" version of (2.12) and reads: seek $\phi_h\in N^h(\Omega)$ such that

$$a(\phi_h,\varphi_h)+Q(\phi_h,\varphi_h)=b(\varphi_h)\quad\forall\varphi_h\in N_0^h(\Omega)\,. \tag{2.14}$$

3. Analysis. Let $a^h(\cdot,\cdot):=a(\cdot,\cdot)+Q(\cdot,\cdot)$ denote the bilinear form of the stabilized method (2.14). The analysis in this section utilizes a non-standard approach based on the algebraic representation of this form. In order to avoid non-essential technical details and focus instead on the key junctures of the proofs we assume that (i) $g=0$, i.e., the solution of (2.14) $\phi_h\in N_0^h(\Omega)$, and (ii) ε and $\mathbf{u}$ are constant on Ω. Appendix A extends the analysis to non-constant diffusion and velocity. We first establish some useful matrix representations of the forms involved in (2.14).

LEMMA 3.1. *Let $\boldsymbol{\phi}$ and $\boldsymbol{\varphi}$ denote the coefficient vectors of $\phi_h,\varphi_h\in N_0^h(\Omega)$. There holds*

$$Q(\phi_h,\varphi_h)=\boldsymbol{\varphi}^T\big(G^TDMDG\big)\boldsymbol{\phi} \tag{3.1}$$

where M is the Gramm matrix of $\{\vec{W}_\xi\}$ and D is diagonal matrix with element $d_{\xi,\xi}=\widetilde{\theta}_\xi$.

Proof. The formula follows by factoring out the finite element coefficient vectors in (2.13)

$$Q(\phi_h,\varphi_h)=\boldsymbol{\varphi}^T\left(\sum_{\boldsymbol{e}_\xi,\boldsymbol{e}_\eta}\boldsymbol{g}_\xi^T\widetilde{\theta}_\xi\Big(\int_\Omega\vec{W}_\xi\cdot\vec{W}_\eta dV\Big)\widetilde{\theta}_\eta\,\boldsymbol{g}_\eta\right)\boldsymbol{\phi}$$

and noting that the integral above is element $M_{\xi,\eta}$ of the Gramm matrix and $G^TDMDG=\sum_{\xi,\eta}\boldsymbol{g}_\xi^T\widetilde{\theta}_\xi\,M_{\xi,\eta}\,\widetilde{\theta}_\eta\boldsymbol{g}_\eta$, where $\boldsymbol{g}_\xi$ and $\boldsymbol{g}_\eta$ are the rows of G corresponding to edges $\boldsymbol{e}_\xi$ and $\boldsymbol{e}_\eta$. ∎

3.1. Stability. We establish a baseline property of the stabilized formulation.

LEMMA 3.2. *For $\varepsilon>0$ the form $a^h(\cdot,\cdot)$ defines an inner product on $N_0^h(\Omega)\times N_0^h(\Omega)$.*

Proof. Obviously, $a^h(\cdot,\cdot)$ is a symmetric bilinear form. Therefore, we only need to show that this form is positive definite. Since $\phi_h\in N_0^h(\Omega)$ and $\nabla\cdot\mathbf{u}=0$,

$$a^h(\phi_h,\,\phi_h)=\|\sqrt{\varepsilon}\nabla\phi_h\|_0^2+Q(\phi_h,\,\phi_h).$$

The lemma follows from the fact that $\nabla\phi_h=0$ for $\phi_h\in N_0^h(\Omega)$ iff $\phi_h=0$, and

$$Q(\phi_h,\phi_h)=(M^{1/2}DG\boldsymbol{\phi})^T(M^{1/2}DG\boldsymbol{\phi})=|M^{1/2}DG\boldsymbol{\phi}|^2\geq 0\,,$$

i.e., the stabilizing term is non-negative. ∎

Lemma 3.2 implies that

$$|||\phi_h||| := \sqrt{a^h(\phi_h,\, \phi_h)} \tag{3.2}$$

is a mesh-dependent "energy" norm on $N_0^h(\Omega)$. We examine its properties in the following theorem.

THEOREM 3.3. *Assume that $K(\Omega)$ is shape-regular finite element partition of Ω, and that ε and $\mathbf{u}$ are constant in Ω. There exists a positive constant, ρ, independent of ε, $\mathbf{u}$, and h, such that*

(i) for linear finite elements (on triangular or tetrahedral meshes)

$$a^h(\phi_h\, \phi_h) \geq \rho\, h \sum_{\boldsymbol{k}\in K(\Omega)} |\mathbf{u}|\ \|\nabla\phi_h\|^2_{0,\boldsymbol{k}}\ , \tag{3.3}$$

(ii) for bilinear and trilinear elements (on quadrilaterals and hexahedrals)

$$a^h(\phi_h, \phi_h) \geq \rho\, h \sum_{\boldsymbol{k}\in K(\Omega)} \|\mathbf{u}_{\boldsymbol{k}} \cdot \nabla\phi_h\|^2_{0,\boldsymbol{k}}\ , \tag{3.4}$$

where $\mathbf{u}_{\boldsymbol{k}}$ is the effective advective velocity *field on element $\boldsymbol{k}$, defined in (B.4).*

Proof. The stabilized bilinear form can be written as a sum of element forms

$$a^h(\phi_h, \varphi_h) = \sum_{\boldsymbol{k}\in K(\Omega)} a^h_{\boldsymbol{k}}(\phi_h, \varphi_h) \quad \text{where} \quad a^h_{\boldsymbol{k}}(\phi_h, \varphi_h) := a^h(\phi_h, \varphi_h)|_{\boldsymbol{k}}\,.$$

We will prove that for every element $\boldsymbol{k}$ there is $\rho_{\boldsymbol{k}} > 0$ such that

$$a^h_{\boldsymbol{k}}(\phi_h, \phi_h) \geq \rho_{\boldsymbol{k}}\, h_{\boldsymbol{k}} \begin{cases} |\mathbf{u}|\ \|\nabla\phi_h\|^2_{0,\boldsymbol{k}} & \text{on simplices} \\ \|\mathbf{u}_{\boldsymbol{k}} \cdot \nabla\phi_h\|^2_{0,\boldsymbol{k}} & \text{on quads and hexes} \end{cases} \tag{3.5}$$

Then, the theorem will follow with $\rho = \min_{\boldsymbol{k}\in K(\Omega)} \rho_{\boldsymbol{k}}$ and $h = \min_{\boldsymbol{k}\in K(\Omega)} h_{\boldsymbol{k}}$.

Using the same arguments as in Lemmae 3.1 and 3.2 one can show that

$$a^h_{\boldsymbol{k}}(\phi_h, \phi_h) = \boldsymbol{\phi}^T_{\boldsymbol{k}} \big(\varepsilon G^T_{\boldsymbol{k}} M_{\boldsymbol{k}} G_{\boldsymbol{k}} + G^T_{\boldsymbol{k}}\, D_{\boldsymbol{k}}\, M_{\boldsymbol{k}} D_{\boldsymbol{k}}\, G_{\boldsymbol{k}}\big) \boldsymbol{\phi}_{\boldsymbol{k}}\,, \tag{3.6}$$

where $\boldsymbol{\phi}_{\boldsymbol{k}}$ is the restriction of $\boldsymbol{\phi}$ to the nodes of $\boldsymbol{k}$, $M_{\boldsymbol{k}}$ is the Gramm matrix of the edge element basis on $\boldsymbol{k}$, and $G_{\boldsymbol{k}}$ and $D_{\boldsymbol{k}}$ are matrices containing the rows of G and D, respectively, corresponding to the edges of $\boldsymbol{k}$. Similar to G, the "topological" element gradient matrix $G_{\boldsymbol{k}}$ has a one-dimensional nullspace spanned by the constant vector. As a result, if $\phi_h\big|_{\boldsymbol{k}} = const$, then $a^h_{\boldsymbol{k}}(\phi_h, \phi_h) = 0$, $\nabla\phi_h\big|_{\boldsymbol{k}} = 0$, and (3.5) is trivially satisfied on $\boldsymbol{k}$. Thus, without a loss of generality we may assume that $\phi_h\big|_{\boldsymbol{k}} \neq const$, or what is the same – $G_{\boldsymbol{k}}\boldsymbol{\phi}_{\boldsymbol{k}} \neq 0$.

Algebraic bounds for element forms. Let $\mathbf{z}_{\boldsymbol{k}} = G_{\boldsymbol{k}}\boldsymbol{\phi}_{\boldsymbol{k}}$ and $\mathbf{y}_{\boldsymbol{k}} = D_{\boldsymbol{k}} G_{\boldsymbol{k}} \boldsymbol{\phi}_{\boldsymbol{k}}$. Using (3.6)

$$\begin{aligned} a^h_{\boldsymbol{k}}(\phi_h, \phi_h) &= \varepsilon \mathbf{z}^T_{\boldsymbol{k}}\, M_{\boldsymbol{k}}\, \mathbf{z}_{\boldsymbol{k}} + \mathbf{y}^T_{\boldsymbol{k}}\, M_{\boldsymbol{k}} \mathbf{y}_{\boldsymbol{k}} \geq \underline{\lambda}(M_{\boldsymbol{k}}) \Big(\varepsilon \mathbf{z}^T_{\boldsymbol{k}}\, \mathbf{z}_{\boldsymbol{k}} + \mathbf{y}^T_{\boldsymbol{k}}\, \mathbf{y}_{\boldsymbol{k}}\Big) \\ &= \underline{\lambda}(M_{\boldsymbol{k}}) \Big(\varepsilon \mathbf{z}^T_{\boldsymbol{k}}\, \mathbf{z}_{\boldsymbol{k}} + \mathbf{z}^T_{\boldsymbol{k}}\, D^2_{\boldsymbol{k}} \mathbf{z}_{\boldsymbol{k}}\Big) = \underline{\lambda}(M_{\boldsymbol{k}}) \sum_{\boldsymbol{e}_\xi \in E(\boldsymbol{k})} \big(\bar{\varepsilon}_\xi + \theta_\xi\big) z^2_\xi = \underline{\lambda}(M_{\boldsymbol{k}}) \sum_{\boldsymbol{e}_\xi \in E(\boldsymbol{k})} \sigma_\xi z^2_\xi\,, \end{aligned}$$

where σ_ξ is the number defined in (2.6). From the lower bound (2.10) in Lemma 2.2

$$a^h_{\boldsymbol{k}}(\phi_h,\phi_h) \geq \underline{\lambda}(M_{\boldsymbol{k}}) \sum_{\boldsymbol{e}_\xi \in E(\boldsymbol{k})} \frac{h_\xi |\bar{u}_\xi|}{2} z_\xi^2 \,.$$

Because $\mathbf{u}$ is constant $\bar{u}_\xi = |\mathbf{u}| v_\xi$, where $v_\xi = (\mathbf{u}/|\mathbf{u}|)\cdot \mathbf{t}_\xi$. Furthermore, there exists a constant $C_{\boldsymbol{k}}$ such that $0 < C_{\boldsymbol{k}} h_{\boldsymbol{k}} \leq h_\xi$ for every edge $\boldsymbol{e}_\xi \in E(\boldsymbol{k})$. It follows that

$$a^h_{\boldsymbol{k}}(\phi_h\,\phi_h) \geq \underline{\lambda}(M_{\boldsymbol{k}}) C_{\boldsymbol{k}} \frac{h_{\boldsymbol{k}}|\mathbf{u}|}{2} \sum_{\boldsymbol{e}_\xi \in E(\boldsymbol{k})} z_\xi^2 |v_\xi| \,. \tag{3.7}$$

We estimate the sum on the right hand side of (3.7) separately for simplices (triangles and tets) and tensor product elements (quads and hexes).

Simplicial elements. On triangles $\mathbf{u}\cdot\mathbf{t}_\xi = 0$ for at most one edge $\boldsymbol{e}_\alpha \in E(\boldsymbol{k})$, in which case $v_\alpha = 0$. On tetrahedrons $\mathbf{u}$ can be orthogonal to either 3 planar edges $\boldsymbol{e}_\alpha$, $\boldsymbol{e}_\beta$ and $\boldsymbol{e}_\gamma$, or at most two non-coplanar edges $\boldsymbol{e}_\alpha$ and $\boldsymbol{e}_\beta$. In the first case $v_\alpha = v_\beta = v_\gamma = 0$ and in the second case $v_\alpha = v_\beta = 0$. In general, let $\boldsymbol{e}_\alpha, \boldsymbol{e}_\beta \in E(\boldsymbol{k})$ be the edges corresponding to the smallest two values $0 \leq |v_\alpha| \leq |v_\beta|$ on element $\boldsymbol{k}$. If $\boldsymbol{k}$ is a triangle, define the set $E(\boldsymbol{k}/\mathbf{u})$ to be the set containing the edge $\boldsymbol{e}_\alpha$. On tetrahedrons, define this set as follows. If $\boldsymbol{e}_\alpha$ and $\boldsymbol{e}_\beta$ are coplanar, set $E(\boldsymbol{k}/\mathbf{u}) = \{\boldsymbol{e}_\alpha, \boldsymbol{e}_\beta, \boldsymbol{e}_\gamma\}$, where $\boldsymbol{e}_\gamma$ is the edge coplanar to $\boldsymbol{e}_\alpha$ and $\boldsymbol{e}_\beta$. If $\boldsymbol{e}_\alpha$ and $\boldsymbol{e}_\beta$ are not coplanar, then define $E(\boldsymbol{k}/\mathbf{u}) = \{\boldsymbol{e}_\alpha, \boldsymbol{e}_\beta\}$. Finally, let $v_{\boldsymbol{k}} = \min_{\boldsymbol{e}_\xi \notin E(\boldsymbol{k}/\mathbf{u})} |v_\xi|$. Then[1],

$$\sum_{\boldsymbol{e}_\xi \in E(\boldsymbol{k})} z_\xi^2 |v_\xi| \geq v_{\boldsymbol{k}} \sum_{\boldsymbol{e}_\xi \notin E(\boldsymbol{k}/\mathbf{u})} z_\xi^2 = v_{\boldsymbol{k}}\, \boldsymbol{\phi}_{\boldsymbol{k}}^T \left(\sum_{\boldsymbol{e}_\xi \notin E(\boldsymbol{k}/\mathbf{u})} \boldsymbol{g}_\xi^T \boldsymbol{g}_\xi \right) \boldsymbol{\phi}_{\boldsymbol{k}}.$$

It is easy to check that for all possible configurations of $E(\boldsymbol{k}/\mathbf{u})$ the matrix $\sum_{\boldsymbol{e}_\xi \notin E(\boldsymbol{k}/\mathbf{u})} \boldsymbol{g}_\xi^T \boldsymbol{g}_\xi$ is positive semidefinite with a single zero eigenvalue corresponding to the constant eigenvector, a case that is ruled out by the assumption $\boldsymbol{\phi}_{\boldsymbol{k}} \neq const$. Its smallest eigenvalue is either 2, when $\boldsymbol{k}$ is a tetrahedron and $E(\boldsymbol{k}/\mathbf{u})$ contains two non coplanar edges, or 1 in all other cases. As a result,

$$\sum_{\boldsymbol{e}_\xi \in E(\boldsymbol{k})} z_\xi^2 |v_\xi| \geq v_{\boldsymbol{k}} |\boldsymbol{\phi}_{\boldsymbol{k}}|^2 \quad \text{and} \quad a^h_{\boldsymbol{k}}(\phi_h\,\phi_h) \geq \underline{\lambda}(M_{\boldsymbol{k}}) C_{\boldsymbol{k}} \frac{h_{\boldsymbol{k}}|\mathbf{u}|}{2} v_{\boldsymbol{k}} |\boldsymbol{\phi}_{\boldsymbol{k}}|^2. \tag{3.8}$$

The next step is to use (1.5) to bound the norm of $\nabla\phi_h$ by $|\boldsymbol{\phi}_{\boldsymbol{k}}|^2$:

$$\|\nabla\phi_h\|^2_{0,\boldsymbol{k}} = \mathbf{z}_{\boldsymbol{k}}^T M_{\boldsymbol{k}}\, \mathbf{z}_{\boldsymbol{k}} \leq \bar{\lambda}(M_{\boldsymbol{k}}) |\mathbf{z}_{\boldsymbol{k}}|^2 = \bar{\lambda}(M_{\boldsymbol{k}}) \boldsymbol{\phi}_{\boldsymbol{k}}^T\, (G_{\boldsymbol{k}}^T G_{\boldsymbol{k}})\, \boldsymbol{\phi}_{\boldsymbol{k}} \leq \bar{\lambda}(M_{\boldsymbol{k}}) \bar{\lambda}(G_{\boldsymbol{k}}^T G_{\boldsymbol{k}}) |\boldsymbol{\phi}_{\boldsymbol{k}}|^2. \tag{3.9}$$

Taking into account that $\bar{\lambda}(G_{\boldsymbol{k}}^T G_{\boldsymbol{k}})$ equals 3 for triangles and 4 for tets, (3.9) yields

$$|\boldsymbol{\phi}_{\boldsymbol{k}}|^2 \geq \frac{1}{4\bar{\lambda}(M_{\boldsymbol{k}})} \|\nabla\phi_h\|^2_{0,\boldsymbol{k}}.$$

Combining this estimate with (3.8) shows that on simplicial elements

$$a^h_{\boldsymbol{k}}(\phi_h\,\phi_h) \geq \rho_{\boldsymbol{k}}\, h_{\boldsymbol{k}}\, |\mathbf{u}|\, \|\nabla\phi_h\|^2_{0,\boldsymbol{k}} \,, \quad \text{with} \quad \rho_{\boldsymbol{k}} = \frac{1}{8} \kappa^{-1}(M_{\boldsymbol{k}}) C_{\boldsymbol{k}} v_{\boldsymbol{k}} \,.$$

REMARK 1. *The value of $v_{\boldsymbol{k}}$ depends on the direction of $\mathbf{u}$ with respect to the element edges. For example, on a triangular grid $v_{\boldsymbol{k}} \geq \sin\left(\frac{\zeta_{\boldsymbol{k}}}{2}\right)$, where $\zeta_{\boldsymbol{k}}$ is the smallest angle in $K(\Omega)$.*

[1] We remind that $\boldsymbol{g}_\xi$ is the row of $G_{\boldsymbol{k}}$ corresponding to edge $\boldsymbol{e}_\xi \in E(\boldsymbol{k})$.

Quadrilateral elements. We refer to Appendix B for the relevant notations. According to (B.3) $|v_{\mathbf{i}}| \le |v_{\alpha^i}|$ and $|v_{\mathbf{j}}| \le |v_{\beta^i}|$. It follows that

$$\sum_{\boldsymbol{e}_\xi \in E(\boldsymbol{k})} z_\xi^2 |v_\xi| \ge |v_{\mathbf{i}}| \left(z_{\alpha^1}^2 + z_{\alpha^2}^2\right) + |v_{\mathbf{j}}| \left(z_{\beta^1}^2 + z_{\beta^2}^2\right) = |v_{\mathbf{i}}| \boldsymbol{\phi}_{\boldsymbol{k}}^T G_\alpha \boldsymbol{\phi}_{\boldsymbol{k}} + |v_{\mathbf{j}}| \boldsymbol{\phi}_{\boldsymbol{k}}^T G_\beta \boldsymbol{\phi}_{\boldsymbol{k}} \tag{3.10}$$

where $G_\alpha = \boldsymbol{g}_{\alpha^1}^T \boldsymbol{g}_{\alpha^1} + \boldsymbol{g}_{\alpha^2}^T \boldsymbol{g}_{\alpha^2}$ and $G_\beta = \boldsymbol{g}_{\beta^1}^T \boldsymbol{g}_{\beta^1} + \boldsymbol{g}_{\beta^2}^T \boldsymbol{g}_{\beta^2}$. We bound (3.10) from below by the norm of the streamline derivative $\mathbf{u}_{\boldsymbol{k}} \cdot \nabla \phi_h$ along the *effective advective velocity*. Since the terms involving G_α and G_β have the same structure we give the details for the first term only.

Step 1. Algebraic lower bound for $\boldsymbol{\phi}_{\boldsymbol{k}}^T G_\alpha \boldsymbol{\phi}_{\boldsymbol{k}}$. Let $\{\mathbf{q}_i\}$ be an orthonormal basis for the (two-dimensional) kernel of G_α and consider the orthogonal decomposition of the coefficient vector $\boldsymbol{\phi}_{\boldsymbol{k}}$:

$$\boldsymbol{\phi}_{\boldsymbol{k}} = \boldsymbol{\phi}_{\mathbf{q}}^\perp + \boldsymbol{\phi}_{\mathbf{q}} \quad \text{with} \quad \boldsymbol{\phi}_{\mathbf{q}}^\perp = \left(I - \sum_{i=1}^2 \mathbf{q}_i \mathbf{q}_i^T \right) \boldsymbol{\phi}_{\boldsymbol{k}} \,.$$

It is easy to check that all non-zero eigenvalues of G_α equal 2, and so,

$$\boldsymbol{\phi}_{\boldsymbol{k}}^T G_\alpha \boldsymbol{\phi}_{\boldsymbol{k}} = (\boldsymbol{\phi}_{\mathbf{q}}^\perp + \boldsymbol{\phi}_{\mathbf{q}})^T G_\alpha (\boldsymbol{\phi}_{\mathbf{q}}^\perp + \boldsymbol{\phi}_{\mathbf{q}}) = (\boldsymbol{\phi}_{\mathbf{q}}^\perp)^T G_\alpha \boldsymbol{\phi}_{\mathbf{q}}^\perp \ge 2 \left|\boldsymbol{\phi}_{\mathbf{q}}^\perp\right|^2 . \tag{3.11}$$

Step 2. Lower bound for $|\boldsymbol{\phi}_{\mathbf{q}}^\perp|$. We obtain an analogue of (3.9) for tensor product elements. Let $\phi_{\mathbf{q},h}^\perp$ be the finite element function with coefficient vector $\boldsymbol{\phi}_{\mathbf{q}}^\perp$. Then,

$$\|\nabla \phi_{\mathbf{q},h}^\perp\|_{0,\boldsymbol{k}}^2 = (\boldsymbol{\phi}_{\mathbf{q}}^\perp)^T G_{\boldsymbol{k}}^T M_{\boldsymbol{k}} G_{\boldsymbol{k}} \, \boldsymbol{\phi}_{\mathbf{q}}^\perp \le \bar{\lambda}(M_{\boldsymbol{k}}) \left| G_{\boldsymbol{k}} \, \boldsymbol{\phi}_{\mathbf{q}}^\perp \right|^2 \le \bar{\lambda}(M_{\boldsymbol{k}}) \bar{\lambda}(G_{\boldsymbol{k}}^T G_{\boldsymbol{k}}) |\boldsymbol{\phi}_{\mathbf{q}}^\perp|^2 . \tag{3.12}$$

Using (3.12) in conjunction with the fact that $\bar{\lambda}(G_{\boldsymbol{k}}^T G_{\boldsymbol{k}}) = 4$ for quads yields

$$|\boldsymbol{\phi}_{\mathbf{q}}^\perp|^2 \ge \frac{1}{4\bar{\lambda}(M_{\boldsymbol{k}})} \|\nabla \phi_{\mathbf{q},h}^\perp\|_{0,\boldsymbol{k}}^2 . \tag{3.13}$$

Step 3. Lower bound for $\|\nabla \phi_{\mathbf{q},h}^\perp\|_{0,\boldsymbol{k}}$. Bounding of the element forms for tensor product elements requires one more step that was not necessary for simplicial elements. This step estimates $\|\nabla \phi_{\mathbf{q},h}^\perp\|_{0,\boldsymbol{k}}$ by a norm of the directional derivative $\mathbf{t}_{\mathbf{i}} \cdot \nabla \phi_h$ of the finite element function along the vector field aligned with the edges $\boldsymbol{e}_{\alpha^i}$. Using (1.4) and the formula for $\boldsymbol{\phi}_{\mathbf{q}}^\perp$ yields

$$\nabla \phi_{\mathbf{q},h}^\perp = \sum_{\boldsymbol{e}_\xi \in E(\boldsymbol{k})} \left(\boldsymbol{g}_\xi \boldsymbol{\phi}_{\boldsymbol{k}} - \boldsymbol{g}_\xi \sum_{i=1}^2 (\mathbf{q}_i^T \boldsymbol{\phi}_{\boldsymbol{k}}) \mathbf{q}_i \right) \vec{W}_\xi = \nabla \phi_h - \sum_{\boldsymbol{e}_\xi \in E(\boldsymbol{k})} \boldsymbol{g}_\xi \left(\sum_{i=1}^2 (\mathbf{q}_i^T \boldsymbol{\phi}_{\boldsymbol{k}}) \mathbf{q}_i \right) \vec{W}_\xi$$

Since $\mathbf{q}_i \in \ker G_\alpha$, $\boldsymbol{g}_{\alpha^j} \mathbf{q}_i = 0$ and summation over $E(\boldsymbol{k})$ reduces to a sum over $\boldsymbol{e}_{\beta^j}$ only:

$$\nabla \phi_{\mathbf{q},h}^\perp = \nabla \phi_h - \sum_{j=1}^2 \boldsymbol{g}_{\beta^j} \left(\sum_{i=1}^2 (\mathbf{q}_i^T \boldsymbol{\phi}_{\boldsymbol{k}}) \mathbf{q}_i \right) \vec{W}_{\beta^j} \,.$$

Using the expression (B.5) for the edge element basis functions we find that

$$\nabla \phi_{\mathbf{q},h}^\perp = \nabla \phi_h - \left[\sum_{j=1}^2 \boldsymbol{g}_{\beta^j} \left(\sum_{i=1}^2 (\mathbf{q}_i^T \boldsymbol{\phi}_{\boldsymbol{k}}) \mathbf{q}_i \right) B_j(\boldsymbol{x}) \right] \mathbf{J}_{\boldsymbol{k}}^{-T}(\boldsymbol{x}) \, \mathbf{j} = \nabla \phi_h - r(\boldsymbol{x}) \frac{\mathbf{J}_{\boldsymbol{k}}^{-T}(\boldsymbol{x}) \, \mathbf{j}}{|\mathbf{J}_{\boldsymbol{k}}^{-T}(\boldsymbol{x}) \, \mathbf{j}|} \,,$$

where $r(\boldsymbol{x})$ is a scalar function corresponding to the expression in the square brackets scaled by $\left|\mathbf{J}_{\boldsymbol{k}}^{-T}(\boldsymbol{x})\,\mathbf{j}\right|$. On the other hand, from (B.1) it follows that

$$\frac{(\mathbf{J}_{\boldsymbol{k}}^{-T}(\boldsymbol{x})\,\mathbf{j})^T}{\left|\mathbf{J}_{\boldsymbol{k}}^{-T}(\boldsymbol{x})\,\mathbf{j}\right|}\,\mathbf{t_i} = \frac{\mathbf{j}^T\,\left(\mathbf{J}_{\boldsymbol{k}}^{-1}(\boldsymbol{x})\mathbf{J}_{\boldsymbol{k}}(\boldsymbol{x})\right)\,\mathbf{i}}{\left|\mathbf{J}_{\boldsymbol{k}}^{-T}(\boldsymbol{x})\,\mathbf{j}\right|\,|\mathbf{J}_{\boldsymbol{k}}(\boldsymbol{x})\,\mathbf{i}|} = \frac{\mathbf{j}^T\,\mathbf{i}}{\left|\mathbf{J}_{\boldsymbol{k}}^{-T}(\boldsymbol{x})\,\mathbf{j}\right|\,|\mathbf{J}_{\boldsymbol{k}}(\boldsymbol{x})\,\mathbf{i}|} = 0 \qquad \forall \boldsymbol{x} \in \boldsymbol{k}\,.$$

i.e., $\mathbf{J}_{\boldsymbol{k}}^{-T}(\boldsymbol{x})\,\mathbf{j}/\left|\mathbf{J}_{\boldsymbol{k}}^{-T}(\boldsymbol{x})\,\mathbf{j}\right|$ is perpendicular to $\mathbf{t_i}$. Accordingly, we denote the former by $\mathbf{t}_\mathbf{i}^\perp$ and so,

$$\nabla\phi_{\mathbf{q},h}^\perp = \nabla\phi_h - r(\boldsymbol{x})\mathbf{t}_\mathbf{i}^\perp\,.$$

Since $\mathbf{t_i}$ and $\mathbf{t}_\mathbf{i}^\perp$ are orthonormal we have the orthogonal decomposition

$$\nabla\phi_h = (\mathbf{t_i}\cdot\nabla\phi_h)\mathbf{t_i} + (\mathbf{t}_\mathbf{i}^\perp\cdot\nabla\phi_h)\mathbf{t}_\mathbf{i}^\perp\,. \tag{3.14}$$

The second term is the orthogonal projection of the gradient onto $\mathbf{t}_\mathbf{i}^\perp$. From the properties of orthogonal projections and (3.14) it follows that

$$\left|\nabla\phi_{\mathbf{q},h}^\perp\right| = \left|\nabla\phi_h - r(\boldsymbol{x})\mathbf{t}_\mathbf{i}^\perp\right| \geq \left|\nabla\phi_h - (\mathbf{t}_\mathbf{i}^\perp\cdot\nabla\phi_h)\mathbf{t}_\mathbf{i}^\perp\right| = |(\mathbf{t_i}\cdot\nabla\phi_h)\mathbf{t_i}| \qquad \forall \boldsymbol{x} \in \boldsymbol{k}\,. \tag{3.15}$$

Combining this result with the lower bounds in (3.11) and (3.13) yields

$$\boldsymbol{\phi}_{\boldsymbol{k}}^T G_\alpha \boldsymbol{\phi}_{\boldsymbol{k}} \geq \frac{1}{2\bar{\lambda}(M_{\boldsymbol{k}})}\|(\mathbf{t_i}\cdot\nabla\phi_h)\,\mathbf{t_i}\|_{0,\boldsymbol{k}}^2\,.$$

Repeating the same steps for the second term in (3.10) leads to

$$\boldsymbol{\phi}_{\boldsymbol{k}}^T G_\beta \boldsymbol{\phi}_{\boldsymbol{k}} \geq \frac{1}{2\bar{\lambda}(M_{\boldsymbol{k}})}\|(\mathbf{t_j}\cdot\nabla\phi_h)\,\mathbf{t_j}\|_{0,\boldsymbol{k}}^2\,.$$

Recalling that $\mathbf{t_i}$ and $\mathbf{t_j}$ are unit vectors for every $\boldsymbol{x} \in \boldsymbol{k}$ gives the intermediate lower bound

$$\sum_{\boldsymbol{e}_\xi \in E(\boldsymbol{k})} z_\xi^2 |v_\xi| \geq \frac{1}{2\bar{\lambda}(M_{\boldsymbol{k}})}\Big(|v_\mathbf{i}|\,\|\mathbf{t_i}\cdot\nabla\phi_h\|_{0,\boldsymbol{k}}^2 + |v_\mathbf{j}|\,\|\mathbf{t_j}\cdot\nabla\phi_h\|_{0,\boldsymbol{k}}^2\Big)\,.$$

Since $|v_\mathbf{i}| \leq 1$, we have that $|v_\mathbf{i}|^2 \leq |v_\mathbf{i}|$ and the same for $v_\mathbf{j}$. Using this, and the triangle inequality

$$\begin{aligned}|v_\mathbf{i}|\,\|\mathbf{t_i}\cdot\nabla\phi_h\|_{0,\boldsymbol{k}}^2 + |v_\mathbf{j}|\,\|\mathbf{t_j}\cdot\nabla\phi_h\|_{0,\boldsymbol{k}}^2 &\geq \|v_\mathbf{i}\,\mathbf{t_i}\cdot\nabla\phi_h\|_{0,\boldsymbol{k}}^2 + \|v_\mathbf{j}\,\mathbf{t_j}\cdot\nabla\phi_h\|_{0,\boldsymbol{k}}^2 \\ &\geq \|v_\mathbf{i}\,\mathbf{t_i}\cdot\nabla\phi_h + v_\mathbf{j}\,\mathbf{t_j}\cdot\nabla\phi_h\|_{0,\boldsymbol{k}}^2 = \|(v_\mathbf{i}\,\mathbf{t_i} + v_\mathbf{j}\,\mathbf{t_j})\cdot\nabla\phi_h\|_{0,\boldsymbol{k}}^2 = \frac{1}{|\mathbf{u}|}\|\mathbf{u_k}\cdot\nabla\phi_h\|_{0,\boldsymbol{k}}^2\,.\end{aligned}$$

As a result,

$$\sum_{\boldsymbol{e}_\xi \in E(\boldsymbol{k})} z_\xi^2 |v_\xi| \geq \frac{1}{2\bar{\lambda}(M_{\boldsymbol{k}})|\mathbf{u}|}\|\mathbf{u_k}\cdot\nabla\phi_h\|_{0,\boldsymbol{k}}^2\,.$$

Together with (3.7) this inequality allows us to conclude that on quads

$$a_{\boldsymbol{k}}^h(\phi_h\,\phi_h) \geq \rho_{\boldsymbol{k}} h_{\boldsymbol{k}} \|\mathbf{u_k}\cdot\nabla\phi_h\|_{0,\boldsymbol{k}}^2 \quad \text{with} \quad \rho_{\boldsymbol{k}} = \frac{1}{4}C_{\boldsymbol{k}}\kappa^{-1}(M_{\boldsymbol{k}})\,.$$

Hexahedral elements. Estimation of the element forms on hexahedra follows the same key steps. The analogue of (3.10) is given by

$$\sum_{\boldsymbol{e}_\xi \in E(\boldsymbol{k})} z_\xi^2 |v_\xi| \geq |v_{\mathbf{i}}| \boldsymbol{\phi}_{\boldsymbol{k}}^T G_\alpha \boldsymbol{\phi}_{\boldsymbol{k}} + |v_{\mathbf{j}}| \boldsymbol{\phi}_{\boldsymbol{k}}^T G_\beta \boldsymbol{\phi}_{\boldsymbol{k}} + |v_{\mathbf{k}}| \boldsymbol{\phi}_{\boldsymbol{k}}^T G_\gamma \boldsymbol{\phi}_{\boldsymbol{k}}. \tag{3.16}$$

where $G_\alpha = \sum_{i=1}^{4} \boldsymbol{g}_{\alpha^i}^T \boldsymbol{g}_{\alpha^i}$, $G_\beta = \sum_{i=1}^{4} \boldsymbol{g}_{\beta^i}^T \boldsymbol{g}_{\beta^i}$, and $G_\gamma = \sum_{i=1}^{4} G_{\gamma^i}^T G_{\gamma^i}$ have four-dimensional nullspaces. We briefly discuss bounding the first term on the right hand side of (3.16).

Step 1. Algebraic lower bound for $\boldsymbol{\phi}_{\boldsymbol{k}}^T G_\alpha \boldsymbol{\phi}_{\boldsymbol{k}}$. Let $\{\mathbf{q}_i\}$ denote an orthonormal basis of $\ker G_\alpha$. For hexahedrons all nonzero eigenvalues of G_α equal 2 and so, (3.11) continues to hold:

$$\boldsymbol{\phi}_{\boldsymbol{k}}^T G_\alpha \boldsymbol{\phi}_{\boldsymbol{k}} \geq 2 \left| \boldsymbol{\phi}_{\mathbf{q}}^{\perp} \right|^2 . \tag{3.17}$$

Step 2. Lower bound for $|\boldsymbol{\phi}_{\mathbf{q}}^{\perp}|$. On hexahedra $\bar{\lambda}(G_{\boldsymbol{k}}^T G_{\boldsymbol{k}}) = 6$ and the analogue of (3.13) is

$$|\boldsymbol{\phi}_{\mathbf{q}}^{\perp}|^2 \geq \frac{1}{6\bar{\lambda}(M_{\boldsymbol{k}})} \|\nabla \phi_{\mathbf{q},h}^{\perp}\|_{0,\boldsymbol{k}}^2. \tag{3.18}$$

Step 3. Lower bound for $\|\nabla \phi_{\mathbf{q},h}^{\perp}\|_{0,\boldsymbol{k}}$. As for quads, we have that

$$\nabla \phi_{\mathbf{q},h}^{\perp} = \nabla \phi_h - \sum_{\boldsymbol{e}_\xi \in E(\boldsymbol{k})} \boldsymbol{g}_\xi \left(\sum_{i=1}^{4} (\mathbf{q}_i^T \boldsymbol{\phi}_{\boldsymbol{k}}) \mathbf{q}_i \right) \vec{W}_\xi \, .$$

Likewise, since $\mathbf{q}_i \in \ker G_\alpha$, $\boldsymbol{g}_{\alpha^j} \mathbf{q}_i = 0$, summation over $E(\boldsymbol{k})$ reduces to a sum over $\boldsymbol{e}_{\beta^j}$ and $\boldsymbol{e}_{\gamma^j}$

$$\nabla \phi_{\mathbf{q},h}^{\perp} = \nabla \phi_h - \sum_{j=1}^{4} \boldsymbol{g}_{\beta^j} \left(\sum_{i=1}^{4} (\mathbf{q}_i^T \boldsymbol{\phi}_{\boldsymbol{k}}) \mathbf{q}_i \right) \vec{W}_{\beta^j} - \sum_{j=1}^{4} \boldsymbol{g}_{\gamma_j} \left(\sum_{i=1}^{4} (\mathbf{q}_i^T \boldsymbol{\phi}_{\boldsymbol{k}}) \mathbf{q}_i \right) \vec{W}_{\gamma^j}$$

Using (B.6)

$$\nabla \phi_{\mathbf{q},h}^{\perp} = \nabla \phi_h - \left(r_{\mathbf{j}}(\boldsymbol{x}) \frac{\mathbf{J}_{\boldsymbol{k}}^{-T}(\boldsymbol{x})\, \mathbf{j}}{|\mathbf{J}_{\boldsymbol{k}}^{-T}(\boldsymbol{x})\, \mathbf{j}|} + r_{\mathbf{k}}(\boldsymbol{x}) \frac{\mathbf{J}_{\boldsymbol{k}}^{-T}(\boldsymbol{x})\, \mathbf{k}}{|\mathbf{J}_{\boldsymbol{k}}^{-T}(\boldsymbol{x})\, \mathbf{k}|} \right)$$

For every $\boldsymbol{x} \in \boldsymbol{k}$ the vector field in the parenthesis belongs in the orthogonal complement of the vector field $\mathbf{t}_{\mathbf{i}}$. Using properties of orthogonal projections

$$\left| \nabla \phi_{\mathbf{q},h}^{\perp} \right| \geq |\nabla \phi_h - (\mathbf{t}_{\mathbf{j}} \cdot \nabla \phi_h) \mathbf{t}_{\mathbf{j}} - (\mathbf{t}_{\boldsymbol{k}} \cdot \nabla \phi_h) \mathbf{t}_{\boldsymbol{k}}| = |(\mathbf{t}_{\mathbf{i}} \cdot \nabla \phi_h) \mathbf{t}_{\mathbf{i}}| \qquad \forall \boldsymbol{x} \in \boldsymbol{k} \, .$$

Repeating the same steps for all three terms leads to

$$\boldsymbol{\phi}_{\boldsymbol{k}}^T G_\alpha \boldsymbol{\phi}_{\boldsymbol{k}} \geq \frac{1}{3\bar{\lambda}(M_{\boldsymbol{k}})} \|(\mathbf{t}_{\mathbf{i}} \cdot \nabla \phi_h)\, \mathbf{t}_{\mathbf{i}}\|_{0,\boldsymbol{k}}^2 ; \quad \boldsymbol{\phi}_{\boldsymbol{k}}^T G_\beta \boldsymbol{\phi}_{\boldsymbol{k}} \geq \frac{1}{3\bar{\lambda}(M_{\boldsymbol{k}})} \|(\mathbf{t}_{\mathbf{j}} \cdot \nabla \phi_h)\, \mathbf{t}_{\mathbf{j}}\|_{0,\boldsymbol{k}}^2$$

$$\text{and} \quad \boldsymbol{\phi}_{\boldsymbol{k}}^T G_\gamma \boldsymbol{\phi}_{\boldsymbol{k}} \geq \frac{1}{3\bar{\lambda}(M_{\boldsymbol{k}})} \|(\mathbf{t}_{\mathbf{k}} \cdot \nabla \phi_h)\, \mathbf{t}_{\mathbf{k}}\|_{0,\boldsymbol{k}}^2 \, .$$

Proceeding as in the case of quadrilaterals we find that

$$\sum_{\boldsymbol{e}_\xi \in E(\boldsymbol{k})} z_\xi^2 |v_\xi| \geq \frac{1}{3\bar{\lambda}(M_{\boldsymbol{k}})} \|(v_{\mathbf{i}}\,\mathbf{t}_{\mathbf{i}} + v_{\mathbf{j}}\,\mathbf{t}_{\mathbf{j}} + v_{\mathbf{k}}\,\mathbf{t}_{\mathbf{k}}) \cdot \nabla\phi_h\|^2_{0,\boldsymbol{k}} = \frac{1}{3\bar{\lambda}(M_{\boldsymbol{k}})|\mathbf{u}|} \|\mathbf{u}_{\boldsymbol{k}} \cdot \nabla\phi_h\|^2_{0,\boldsymbol{k}}\,,$$

which produces an analogous lower bound for the element form:

$$a^h_{\boldsymbol{k}}(\phi_h\,\phi_h) \geq \rho_{\boldsymbol{k}} h_{\boldsymbol{k}} \|\mathbf{u}_{\boldsymbol{k}} \cdot \nabla\phi_h\|^2_{0,\boldsymbol{k}} \quad \text{with} \quad \rho_{\boldsymbol{k}} = \frac{1}{6} C_{\boldsymbol{k}} \kappa^{-1}(M_{\boldsymbol{k}})\,.$$

□

Theorem 3.3 implies that (2.14) remains stable for $\varepsilon \ll h$. In conjunction with the Lax-Milgram lemma, the theorem allows us to conclude that the stabilized method (2.14) has a unique solution.

3.2. Error estimates. We start with an upper bound on the artificial diffusion term $Q(\cdot,\cdot)$.

LEMMA 3.4. *There exists a constant C, independent of h, such that for all $\phi_h \in N^h(\Omega)$*

$$Q(\phi_h, \phi_h) \leq C\, h\, |\mathbf{u}|\, \|\nabla\phi_h\|^2_0\,. \tag{3.19}$$

Proof. The stabilizing term $Q(\cdot,\cdot)$ can be written as a sum of element forms $Q_{\boldsymbol{k}}(\phi_h,\varphi_h) := Q(\phi_h,\varphi_h)|_{\boldsymbol{k}}$. We use the techniques of Theorem 3.3 to bound these forms by $\|\nabla\phi_h\|^2_{0,\boldsymbol{k}}$. As in the proof of this theorem it suffices to consider only those ϕ_h for which $G_{\boldsymbol{k}}\boldsymbol{\phi}_{\boldsymbol{k}} \neq 0$. Using the upper bound for θ_α in (2.10) and the fact that $\overline{\lambda}(G_{\boldsymbol{k}}^T G_{\boldsymbol{k}}) \leq 6$ for all elements under consideration

$$Q_{\boldsymbol{k}}(\phi_h,\phi_h) = \boldsymbol{\phi}_{\boldsymbol{k}}^T \left(G_{\boldsymbol{k}}^T\, D_{\boldsymbol{k}}\, M_{\boldsymbol{k}} D_{\boldsymbol{k}}\, G_{\boldsymbol{k}}\right) \boldsymbol{\phi}_{\boldsymbol{k}}^T \leq \overline{\lambda}(M_{\boldsymbol{k}})\overline{\lambda}(G_{\boldsymbol{k}}^T G_{\boldsymbol{k}}) \frac{h_{\boldsymbol{k}}|\mathbf{u}|}{2} |\boldsymbol{\phi}_{\boldsymbol{k}}|^2 \leq 3\overline{\lambda}(M_{\boldsymbol{k}}) h_{\boldsymbol{k}} |\mathbf{u}|\, |\boldsymbol{\phi}_{\boldsymbol{k}}|^2\,.$$

On the other hand, since the smallest non-zero eigenvalue $\underline{\lambda}(G_{\boldsymbol{k}}^T G_{\boldsymbol{k}}) \geq 1$ for all element shapes

$$\|\nabla\phi_h\|^2_{0,\boldsymbol{k}} = \boldsymbol{\phi}_{\boldsymbol{k}}^T \left(G_{\boldsymbol{k}}^T\, M_{\boldsymbol{k}}\, G_{\boldsymbol{k}}\right) \boldsymbol{\phi}_{\boldsymbol{k}}^T \geq \underline{\lambda}(M_{\boldsymbol{k}})\underline{\lambda}(G_{\boldsymbol{k}}^T G_{\boldsymbol{k}}) |\boldsymbol{\phi}_{\boldsymbol{k}}|^2 \geq \underline{\lambda}(M_{\boldsymbol{k}}) |\boldsymbol{\phi}_{\boldsymbol{k}}|^2\,.$$

Therefore, for every element in $K(\Omega)$

$$Q_{\boldsymbol{k}}(\phi_h,\phi_h) \leq C_{\boldsymbol{k}} h_{\boldsymbol{k}} |\mathbf{u}|\, \|\nabla\phi_h\|^2_{0,\boldsymbol{k}} \quad \text{with} \quad C_{\boldsymbol{k}} = 3\kappa(M_{\boldsymbol{k}})\,,$$

and the Lemma follows with $C = \max_{\boldsymbol{k}} C_{\boldsymbol{k}}$. □

We proceed to estimate the discrete error of the solution of (2.14).

THEOREM 3.5. *Assume that the exact solution of (1.1) is in $H^2(\Omega)$, ϕ_h is the solution of (2.14), and $\phi_{\mathcal{I}} := \mathcal{I}_N(\phi)$. There exists a constant C, independent of h and ε, such that*

$$|||\phi_h - \phi_{\mathcal{I}}||| \leq \gamma_1 \sqrt{h} \|\phi\|_2\,. \tag{3.20}$$

Proof. We have the following "error orthogonality" relation

$$a(\phi,\varphi_h) = b(\varphi_h) = a^h(\phi_h,\varphi_h) \quad \forall \varphi_h \in N^h_0(\Omega)\,.$$

This identity, the fact that $a^h(\cdot,\cdot) = a(\cdot,\cdot) + Q(\cdot,\cdot)$, and the Cauchy-Schwartz inequality imply that

$$\begin{aligned}
|||\phi_h - \phi_{\mathcal{I}}|||^2 &= a^h(\phi_h - \phi_{\mathcal{I}}, \phi_h - \phi_{\mathcal{I}}) \pm a(\phi - \phi_{\mathcal{I}}, \phi_h - \phi_{\mathcal{I}}) \\
&= a(\phi - \phi_{\mathcal{I}}, \phi_h - \phi_{\mathcal{I}}) + \left[a(\phi_{\mathcal{I}}, \phi_h - \phi_{\mathcal{I}}) - a^h(\phi_{\mathcal{I}}, \phi_h - \phi_{\mathcal{I}})\right] \\
&= a(\phi - \phi_{\mathcal{I}}, \phi_h - \phi_{\mathcal{I}}) + Q(\phi_{\mathcal{I}}, \phi_h - \phi_{\mathcal{I}}) \\
&\leq a(\phi - \phi_{\mathcal{I}}, \phi - \phi_{\mathcal{I}})^{1/2} a(\phi_h - \phi_{\mathcal{I}}, \phi_h - \phi_{\mathcal{I}})^{1/2} + Q(\phi_{\mathcal{I}}, \phi_{\mathcal{I}})^{1/2} Q(\phi_h - \phi_{\mathcal{I}}, \phi_h - \phi_{\mathcal{I}})^{1/2}\,.
\end{aligned} \tag{3.21}$$

Using the interpolation estimate (1.9) and the skew-symmetry of the advection term

$$a(\phi-\phi_{\mathcal{I}},\phi-\phi_{\mathcal{I}})^{1/2}\leq\sqrt{\varepsilon}\|\nabla\phi-\nabla\phi_{\mathcal{I}}\|_0\leq\sqrt{\varepsilon}h\|\phi\|_2\,,$$

while from identity $a^h(\cdot,\cdot)=a(\cdot,\cdot)+Q(\cdot,\cdot)$ it follows that

$$a(\phi_h-\phi_{\mathcal{I}},\phi_h-\phi_{\mathcal{I}})^{1/2}\leq|\!|\!|\phi_h-\phi_{\mathcal{I}}|\!|\!| \quad\text{and}\quad Q(\phi_h-\phi_{\mathcal{I}},\phi_h-\phi_{\mathcal{I}})^{1/2}\leq|\!|\!|\phi_h-\phi_{\mathcal{I}}|\!|\!|\ .$$

Finally, using (3.19) in Lemma 3.4 and the boundedness (1.10) of $\mathcal{I}_N$ we find that

$$Q(\phi_{\mathcal{I}},\phi_{\mathcal{I}})^{1/2}\leq C\,\sqrt{h}\,|\mathbf{u}|\,\|\nabla\phi_{\mathcal{I}}\|_0\leq C\,\sqrt{h}\,|\mathbf{u}|\,\|\phi\|_2\,.$$

Inserting these bounds into (3.21) and dividing through by the energy norm of $\phi_h-\phi_{\mathcal{I}}$ yields

$$|\!|\!|\phi_h-\phi_{\mathcal{I}}|\!|\!|\leq\sqrt{\varepsilon}h\|\phi\|_2+C\,\sqrt{h}\,|\mathbf{u}|\,\|\phi\|_2=C\,\sqrt{h}\,\Big(\sqrt{\varepsilon\,h}+|\mathbf{u}|\Big)\,\|\phi\|_2\,,$$

which completes the proof. □

COROLLARY 3.6. *Under the assumptions of Theorem 3.5 there holds*

$$\sqrt{\varepsilon}\|\ \nabla\phi-\nabla\phi_h\|_0\leq C\sqrt{h}\|\phi\|_2 \tag{3.22}$$

with a constant C that does not depend on h and ε.

Proof. We use triangle inequality to split the error

$$\sqrt{\varepsilon}\|\ \nabla\phi-\nabla\phi_h\|_0\leq\sqrt{\varepsilon}\|\ \nabla\phi-\nabla\phi_{\mathcal{I}}\|_0+\sqrt{\varepsilon}\|\ \nabla\phi_{\mathcal{I}}-\nabla\phi_h\|_0\,.$$

The assertion follows by noting that

$$\varepsilon\|\ \nabla\phi_{\mathcal{I}}-\nabla\phi_h\|_0^2=a(\phi_{\mathcal{I}}-\phi_h,\phi_{\mathcal{I}}-\phi_h)\leq a^h(\phi_{\mathcal{I}}-\phi_h,\phi_{\mathcal{I}}-\phi_h)=|\!|\!|\phi_{\mathcal{I}}-\phi_h,\phi_{\mathcal{I}}-\phi_h|\!|\!|^2$$

and applying the result of Theorem 3.5. □

Let $e_h:=\phi-\phi_h$. To estimate the L^2 error of the stabilized solution we assume that the adjoint problem: find $\psi\in H_0^1(\Omega)$ such that

$$a(\varphi,\,\psi)=(e_h,\,\varphi),\qquad\forall\varphi\in H_0^1(\Omega), \tag{3.23}$$

has full elliptic regularity, i.e., there is $\gamma_\varepsilon>0$ such that

$$\|\psi\|_2\leq\gamma_\varepsilon\|e_h\|_0\,. \tag{3.24}$$

THEOREM 3.7. *Under the hypotheses of Theorem 3.5 there exists $C(\varepsilon)$ such that*

$$\|\phi_h-\phi\|_0\leq C(\varepsilon)\,h\,\|\phi\|_2. \tag{3.25}$$

Proof. Let $\psi_{\mathcal{I}}$ be the interpolant of the solution of the adjoint problem and ϕ_h the stabilized solution. Setting $\varphi=e_h$ in (3.23) and using that $a(\phi,\psi_{\mathcal{I}})=a^h(\phi_h,\psi_{\mathcal{I}})$ gives

$$\begin{aligned}\|e_h\|_0^2 &= a(e_h,\psi)\pm a(e_h,\psi_{\mathcal{I}})\\ &= a(\phi-\phi_h,\psi-\psi_{\mathcal{I}})+a(\phi-\phi_h,\psi_{\mathcal{I}})\\ &= a(\phi-\phi_h,\psi-\psi_{\mathcal{I}})+\big(a^h(\phi_h,\psi_{\mathcal{I}})-a(\phi_h,\psi_{\mathcal{I}})\big)\\ &= a(\phi-\phi_h,\psi-\psi_{\mathcal{I}})-Q(\phi_h,\psi_{\mathcal{I}})\,.\end{aligned}$$

Using the Cauchy-Schwartz inequality for the first term gives the bound

$$a(\phi-\phi_h,\psi-\psi_{\mathcal{I}}) \le \varepsilon\|\nabla e_h\|_0\,\|\nabla\psi-\nabla\psi_{\mathcal{I}}\|_0 + |\mathbf{u}|\,\|e_h\|_0\,\|\nabla\psi-\nabla\psi_{\mathcal{I}}\|_0$$

Using interpolation theory (1.9) and elliptic regularity (3.24)

$$\|\nabla\psi-\nabla\psi_{\mathcal{I}}\|_0 \le C\,h\|\psi\|_2 \le C\,h\|e_h\|_0\,,$$

while from Corollary 3.6 $\sqrt{\varepsilon}\|\nabla e_h\|_0 \le C\sqrt{h}\|\phi\|_2$. Therefore,

$$a(\phi-\phi_h,\psi-\psi_{\mathcal{I}}) \le C\left(h^{3/2}\sqrt{\varepsilon}\|\phi\|_2\|e_h\|_0 + h\,|\mathbf{u}|\,\|e_h\|_0^2\right).$$

Using the Cauchy-Schwartz inequality, (3.19) in Lemma 3.4, (1.10) and elliptic regularity (3.24),

$$\begin{aligned} Q(\phi_h,\psi_{\mathcal{I}}) &\le Q(\phi_h,\phi_h)^{1/2}\,Q(\psi_{\mathcal{I}},\psi_{\mathcal{I}})^{1/2} \\ &\le C\,h\,|\mathbf{u}|\,\|\nabla\phi_h\|_0\,\|\nabla\psi_{\mathcal{I}}\|_0 \le C\,h\,|\mathbf{u}|\,\|\nabla\phi_h\|_0\,\|\psi\|_2 \le C\,h\,|\mathbf{u}|\,\|\nabla\phi_h\|_0\,\|e_h\|_0\,. \end{aligned}$$

This gives the following intermediate upper bound for $\|e_h\|_0$:

$$\|e_h\|_0^2 \le C\left(h^{3/2}\sqrt{\varepsilon}\|\phi\|_2\|e_h\|_0 + h\,|\mathbf{u}|\,\|e_h\|_0^2 + h\,|\mathbf{u}|\,\|\nabla\phi_h\|_0\,\|e_h\|_0\right)$$

For h small enough, $Ch\,|\mathbf{u}| < 1/2$ and $Ch\,|\mathbf{u}|\,\|e_h\|_0^2$ can be absorbed into the left hand side. After dividing the result by $\|e_h\|_0$ the intermediate bound assumes the form

$$\|e_h\|_0 \le C\,h\,\left(\sqrt{\varepsilon\,h}\,\|\phi\|_2 + |\mathbf{u}|\,\|\nabla\phi_h\|_0\right).$$

Using Corollary 3.6, the regularity assumption $\phi\in H^2(\Omega)$ and the triangle inequality,

$$\|\nabla\phi_h\|_0 \le \|\nabla\phi_h-\nabla\phi\|_0 + \|\nabla\phi\|_0 \le \frac{1}{\sqrt{\varepsilon}}\left(\sqrt{\varepsilon}\|\nabla\phi_h-\nabla\phi\|_0\right) + \|\phi\|_2 \le \left(C\sqrt{\frac{h}{\varepsilon}}+1\right)\|\phi\|_2\,.$$

Combining all results together shows that for h small enough

$$\|e_h\|_0 \le C\,h\,\left(\sqrt{\varepsilon\,h} + |\mathbf{u}|\left(C\sqrt{\frac{h}{\varepsilon}}+1\right)\right)\|\phi\|_2\,,$$

which completes the proof. □

4. Computational study. This section provides a brief numerical illustration of the "symmetrized" stabilized method (2.14). We refer to [3] and [4] for thorough numerical studies of the finite element method (2.11) and the related control volume finite element method, respectively, both of which utilize the whole edge element flux F_E.

We solve (1.1) on the unit square $\Omega=[0,1]^2$ using conforming partitions $K_h(\Omega)$ of Ω into quadrilateral and triangular elements. Thus, $N^h(\Omega)$ is either the isoparametric bilinear finite element space on quads, or the affine piecewise linear finite element on triangles. The domain boundary $\Gamma=\Gamma_B\cup\Gamma_T\cup\Gamma_L\cup\Gamma_R$, where Γ_B, Γ_T, Γ_L and Γ_R are the bottom, top, left and right sides of Ω, respectively.

4.1. Convergence rates. The objective is to illustrate the theoretical error estimates (3.20), (3.22) and (3.25) in Section 3.2. To this end, we use the manufactured solution $\phi = x^3 - y^2$ and the velocity field $\mathbf{u} = (-\sin\pi/6, \cos\pi/6)$ from [8, Example 3.1.3, p.118]. Substitution of ϕ and $\mathbf{u}$ into (1.1) defines the boundary data and the forcing term.

Convergence rates of the discrete error $|||\phi_{\mathcal{I}} - \phi_h|||$, the H^1-seminorm error $\|\nabla\phi - \nabla\phi_h\|_0$ and the L^2-norm error $\|\phi - \phi_h\|$ are estimated by solving (2.14) on a sequence of uniform quadrilateral and triangular grids for two different values of ε. Triangular grids are obtained by splitting each element in the quadrilateral grids into two triangles. Tables 4.1–4.2 present solution errors in various norms and the corresponding convergence rate estimates.

TABLE 4.1
Errors and convergence rates for (2.14) on uniform quads and triangles with $\varepsilon = 0.001$.

	Quadrilaterals			Triangles		
Grid	$\|\phi - \phi_h\|$	$\|\nabla\phi - \nabla\phi_h\|_0$	$\|\|\|\phi_{\mathcal{I}} - \phi_h\|\|\|$	$\|\phi - \phi_h$	$\|\nabla\phi - \nabla\phi_h\|_0$	$\|\|\|\phi_{\mathcal{I}} - \phi_h\|\|\|$
32	0.4260E-02	0.7533E-01	0.6333E-02	0.7707E-02	0.8804E-01	0.9588E-02
64	0.2073E-02	0.4794E-01	0.3052E-02	0.3854E-02	0.5712E-01	0.4633E-02
128	0.1061E-02	0.2764E-01	0.1425E-02	0.2029E-02	0.3804E-01	0.2338E-02
Rate	1.002	0.723	1.076	0.963	0.605	1.018

TABLE 4.2
Errors and convergence rates for (2.14) on uniform quads and triangles with $\varepsilon = 0.00001$.

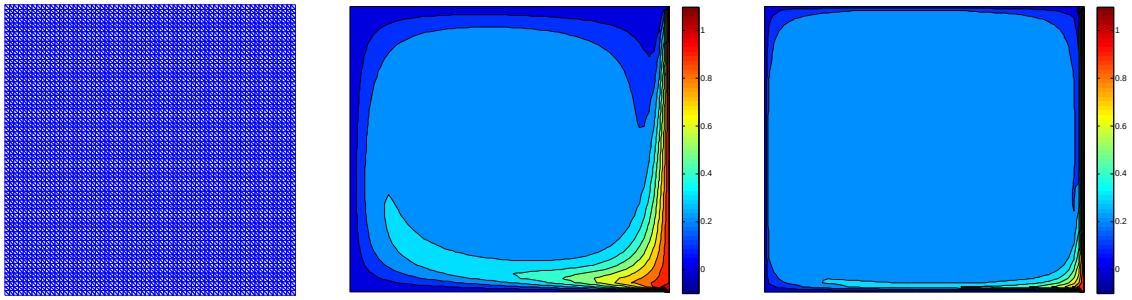
	Quadrilaterals			Triangles		
Grid	$\|\phi - \phi_h\|$	$\|\nabla\phi - \nabla\phi_h\|_0$	$\|\|\|\phi_{\mathcal{I}} - \phi_h\|\|\|$	$\|\phi - \phi_h\|$	$\|\nabla\phi - \nabla\phi_h\|_0$	$\|\|\|\phi_{\mathcal{I}} - \phi_h\|\|\|$
32	0.4739E-02	0.7949E-01	0.6835E-02	0.8594E-02	0.9514E-01	0.1054E-01
64	0.2518E-02	0.5497E-01	0.3552E-02	0.4616E-02	0.6664E-01	0.5485E-02
128	0.1299E-02	0.3842E-01	0.1809E-02	0.2402E-02	0.4684E-01	0.2796E-02
Rate	0.934	0.524	0.959	0.916	0.511	0.957

The results in the tables confirm that the solution of (2.14) is first-order accurate in L^2. These results also suggest that the dependence on ε in the theoretical bound (3.25) appears to be fairly benign. Indeed, decreasing the diffusion coefficient by a factor of 100 does not reduce significantly the L^2 convergence rates, which drop only by 7% and 5%, on triangles and quads, respectively. The H^1-seminorm errors are also in line with the theoretical estimate (3.22). In particular, for $\varepsilon = 0.00001$ the rate of convergence is in an almost perfect agreement with the theoretical prediction.

We note that numerically, the discrete error outperforms the theoretical rate of convergence in (3.20). This may indicate a possible superconvergent behavior of the discrete error on uniform grids that cannot be captured by the proof of Theorem 3.5, which addresses general unstructured grids. Since this superconvergent behavior does not seem to extend to the H^1-seminorm error, its further investigation is beyond the scope of this paper.

Finally, it is worth pointing out that convergence rates on quadrilaterals are slightly, but consistently, higher than on triangles. Theorem 3.3 provides a possible explanation of this observation. Specifically, stability bounds (3.3) and (3.4) suggest that the stabilized method (2.14) is *less diffusive* on quadrilateral girds. We will further examine this conjecture in Section 4.2.

4.2. Qualitative properties. Although asymptotic convergence rates are an important quantitative measure of the accuracy of numerical methods they don't always tell the whole story about their computational properties. In this section we use the challenging "Double Glazing" advection



(a) (b) (c)

FIG. 4.1. *Solution (b) of the Double Glazing test problem (4.1) with* $\varepsilon = 0.00001$ *on a* 64×64 *uniform triangular grid (a) and a* 64×64 *uniform quadrilateral grid (c) by (2.14).*

test [8, Example 3.1.4, p.119] to provide a complementary, qualitative perspective on (2.14). On the unit square the Double Glazing advection test is specified by

$$\mathbf{u} = \begin{pmatrix} 2(2y-1)(1-(2x-1)^2) \\ -2(2x-1)(1-(2y-1)^2) \end{pmatrix}; \quad f = 0; \quad \text{and} \quad g = \begin{cases} 1 & \text{on } \Gamma_R \\ 0 & \text{on } \Gamma_B \cup \Gamma_T \cup \Gamma_L \end{cases}. \tag{4.1}$$

Problem (4.1) models temperature distribution in a cavity with a "hot" external wall (Γ_R). The discontinuities at the two corners of the hot wall create boundary layers near its corners.

Triangles vs. quadrilaterals. According to Theorem 3.3 on tensor product elements the stabilized form $a^h(\cdot,\cdot)$ is bounded from below by the streamline derivative of the finite element solution along the effective elemental advective velocity $\mathbf{u}_k$. On the other hand, for simplicial elements this form is bounded from below by the gradient of the finite element solution. The difference in the stability bounds (3.3) and (3.4) prompts a conjecture that (2.14) will be less diffusive on tensor product elements. To test this conjecture we solve (4.1) with $\varepsilon = 0.00001$ on uniform 64×64 triangular and quadrilateral grids. Figure 4.1 presents the solution plots, which clearly show the more diffusive behavior of the finite element solution on the triangular grid.

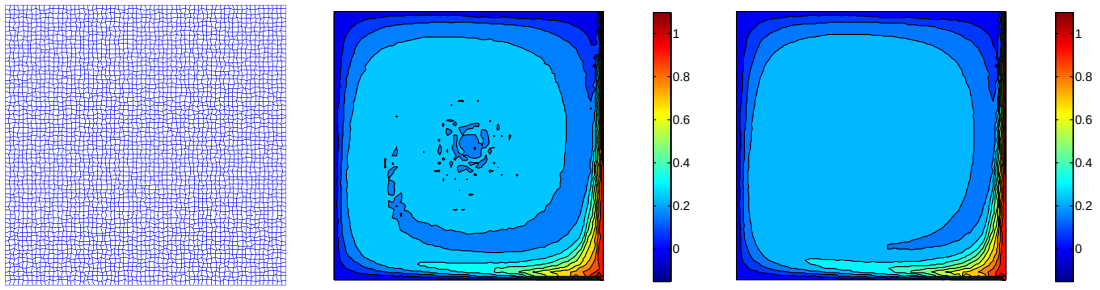
Edge flux stabilized vs. "symmetrized" stabilized formulations. Direct stabilization by the edge element flux $F_E(\phi_h)$ in (2.11) introduces the non-symmetric artificial diffusion kernel exposed in (2.12). The lack of symmetry in this kernel may be detrimental for the accuracy on unstructured grids. Figure 4.2 compares solution of the Double Glazing test problem with $\varepsilon = 0.0001$ by (2.11) and the new "symmetrized" formulation (2.14) on a 64×64 randomly perturbed quadrilateral grid. From the plots it is evident that the solution of (2.11) suffers from numerical noise, whereas the "symmetrized" formulation is not affected by the mesh structure. These observations can be further quantified by measuring the violation of global bounds by the two solutions. For the test problem $\phi^{\min} = 0$ and $\phi^{\max} = 1$ and so, the quantity $\Delta = (\phi_h^{\max} - \phi_h^{\min}) - 1$ provides a measure of the total violation of global solution bounds by the finite element approximations. For the solution of (2.11)

$$\phi_h^{\max} = 1.0936, \quad \phi_h^{\min} = -0.0986 \quad \text{and} \quad \Delta = 0.19,$$

whereas for the solution of (2.14)

$$\phi_h^{\max} = 1.0792, \quad \phi_h^{\min} = -0.0052 \quad \text{and} \quad \Delta = 0.08$$

In other words, violation of global bounds in (2.11) is twice the violation in (2.14).



(a) (b) (c)

FIG. 4.2. *Solution of the Double Glazing test problem (4.1) with* $\varepsilon = 0.0001$ *on a* 64×64 *randomly perturbed quadrilateral mesh (a) by the stabilized formulation (2.11) (b) and the "symmetrized" stabilized method (2.14) (c).*

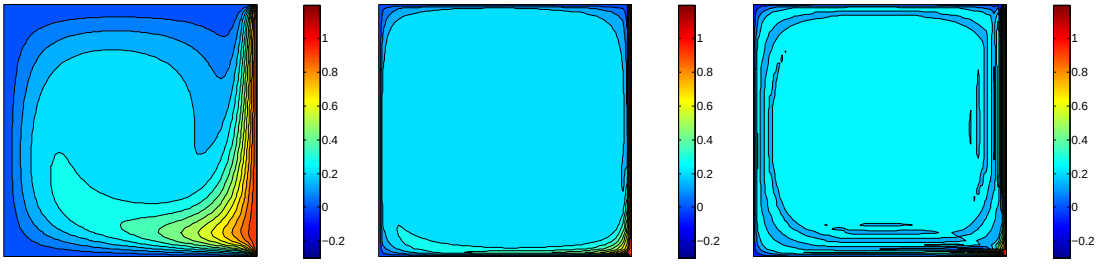
"Symmetrized" stabilized formulation vs. Artificial Diffusion and SUPG. The final qualitative study compares and contrasts (2.14) with the classical artificial diffusion method [14, p.181] and the Streamline Upwind Petrov-Galerkin (SUPG) method [12]. For this test we solve the Double Glazing problem with $\varepsilon = 0.00001$ on a 64×64 uniform quadrilateral grid. Figure 4.3 presents the results. From the solution plots in this figure it is immediately obvious that our formulation is significantly less diffusive than the artificial diffusion method. It also appears to be more robust than the SUPG solution, which develops spurious oscillations near the boundary layers, leading to significant violations of the global solution bounds. In particular, for the SUPG solution

$$\phi_h^{\max} = 1.19, \quad \phi_h^{\min} = -0.28 \quad \text{and} \quad \Delta = 0.47,$$

resulting in a 47% violation of the global solution bounds. In contrast, for the solution of (2.14)

$$\phi_h^{\max} = 1.00, \quad \phi_h^{\min} = -0.0097 \quad \text{and} \quad \Delta = 0.0097,$$

i.e., the total violation of the solution bounds is less than 1%.



(a) (b) (c)

FIG. 4.3. *Solution of the Double Glazing test problem (4.1) with* $\varepsilon = 0.00001$ *on a* 64×64 *uniform quadrilateral mesh by the Artificial Diffusion method (a) the "symmetrized" stabilized method (2.14) (b), and SUPG (c).*

Acknowledgements. The authors acknowledge the support of the Advanced Scientific Computing Research program of the DoE Office of Science and the ASC program of the NNSA in carrying out this research. Discussions with our colleagues from the Charon team X. Gao, G. Hennigan, L. Musso, and T. Smith motivated many aspects of this work.

Appendix A. Extensions to non-constant diffusion and velocity. Throughout this section we assume that ε and $\mathbf{u}$ are of class $C^1(\Omega)$ and $\nabla\cdot\mathbf{u} = 0$. Whenever necessary, dependence on diffusivity and velocity will be indicated explicitly by including these fields in the list of arguments, e.g., $a^h(\phi_h, \varphi_h; \varepsilon, \mathbf{u})$, $\theta_\alpha(\varepsilon, \bar{u}_\alpha)$, $p_\alpha(\varepsilon, \bar{u}_\alpha)$, and so on. To extend the results of Theorem 3.3 to non-constant ε and $\mathbf{u}$ it suffices to show that on every $\boldsymbol{k} \in K(\Omega)$ the element form $a^h_{\boldsymbol{k}}(\phi_h, \phi_h; \varepsilon, \mathbf{u})$ can be bounded from below by a constant times $a^h_{\boldsymbol{k}}(\phi_h, \phi_h; \varepsilon_{\boldsymbol{k}}, \mathbf{u}_{\boldsymbol{k}})$, where $\varepsilon_{\boldsymbol{k}}$ and $\mathbf{u}_{\boldsymbol{k}}$ are constant approximations of ε and $\mathbf{u}$ on element $\boldsymbol{k}$. Here, we set these approximations to

$$\varepsilon_{\boldsymbol{k}} := \max_{\boldsymbol{x}\in\boldsymbol{k}} \varepsilon(\boldsymbol{x}) \quad \text{and} \quad \mathbf{u}_{\boldsymbol{k}} = \frac{1}{|\boldsymbol{k}|}\int_{\boldsymbol{k}} \mathbf{u} dV\,,$$

respectively, i.,e., the largest value of ε on $\boldsymbol{k}$ and the average element velocity, respectively.

LEMMA A.1. *For sufficiently small mesh size h there exists a positive constant C such that*

$$a^h_{\boldsymbol{k}}(\phi_h, \phi_h; \varepsilon, \mathbf{u}) \geq C a^h_{\boldsymbol{k}}(\phi_h, \phi_h; \varepsilon_{\boldsymbol{k}}, \mathbf{u}_{\boldsymbol{k}}) \quad \forall \boldsymbol{k} \in K(\Omega)\,. \tag{A.1}$$

Proof. For variable ε and $\mathbf{u}$

$$a^h_{\boldsymbol{k}}(\phi_h, \phi_h; \varepsilon, \mathbf{u}) = \|\sqrt{\varepsilon}\nabla\phi_h\|^2_{0,\boldsymbol{k}} + Q_{\boldsymbol{k}}(\phi_h,\, \phi_h;\, \varepsilon, \mathbf{u})\,.$$

Regularity assumptions imply the existence of a positive constant γ such that $\varepsilon_{\boldsymbol{k}} \leq (1+\gamma)\,\varepsilon$ on every element $\boldsymbol{k} \in K(\Omega)$. This yields the following bound for the first term in $a^h_{\boldsymbol{k}}(\phi_h, \phi_h; \varepsilon, \mathbf{u})$:

$$\|\sqrt{\varepsilon}\nabla\phi_h\|^2_{0,\boldsymbol{k}} \geq \frac{\varepsilon_{\boldsymbol{k}}}{1+\gamma}\|\nabla\phi_h\|^2_{0,\boldsymbol{k}}\,. \tag{A.2}$$

We bound the second term from below in two steps, starting with an estimate of $Q_{\boldsymbol{k}}(\phi_h,\, \phi_h;\, \varepsilon, \mathbf{u})$ by $Q_{\boldsymbol{k}}(\phi_h,\, \phi_h;\, \varepsilon_{\boldsymbol{k}}, \mathbf{u})$. Proceeding as in the proof of Theorem 3.3, and using that $\theta_\alpha(\varepsilon, \bar{u}_\alpha)$ is monotone decreasing with respect to its first argument, we find that

$$Q_{\boldsymbol{k}}(\phi_h,\, \phi_h;\, \varepsilon, \mathbf{u}) \geq \underline{\lambda}(M_{\boldsymbol{k}}) \sum_{\boldsymbol{e}_\xi\in E(\boldsymbol{k})} \theta_\xi(\varepsilon, \bar{u}_\xi)\, z_\xi^2 \geq \underline{\lambda}(M_{\boldsymbol{k}}) \sum_{\boldsymbol{e}_\xi\in E(\boldsymbol{k})} \theta_\xi(\varepsilon_{\boldsymbol{k}}, \bar{u}_\xi)\, z_\xi^2\,,$$

where, as before, z_ξ is the element of $\mathbf{z}_{\boldsymbol{k}} = \boldsymbol{\phi}_{\boldsymbol{k}} G_{\boldsymbol{k}}$. On the other hand,

$$Q_{\boldsymbol{k}}(\phi_h,\, \phi_h;\, \varepsilon_{\boldsymbol{k}}, \mathbf{u}) \leq \overline{\lambda}(M_{\boldsymbol{k}}) \sum_{\boldsymbol{e}_\xi\in E(\boldsymbol{k})} \theta_\xi(\varepsilon_{\boldsymbol{k}}, \bar{u}_\xi) z_\xi^2\,.$$

Combining these two results yields the intermediate estimate

$$Q_{\boldsymbol{k}}(\phi_h,\, \phi_h;\, \varepsilon, \mathbf{u}) \geq \kappa^{-1}(M_{\boldsymbol{k}})\, Q_{\boldsymbol{k}}(\phi_h,\, \phi_h;\, \varepsilon_{\boldsymbol{k}}, \mathbf{u})\,. \tag{A.3}$$

We proceed to bound $Q_{\boldsymbol{k}}(\phi_h,\, \phi_h;\, \varepsilon_{\boldsymbol{k}}, \mathbf{u})$ from below in terms of $Q_{\boldsymbol{k}}(\phi_h,\, \phi_h;\, \varepsilon_{\boldsymbol{k}}, \mathbf{u}_{\boldsymbol{k}})$. Let

$$D_{\boldsymbol{k}}(\mathbf{u}) := D_{\boldsymbol{k}}(\varepsilon_{\boldsymbol{k}}, \mathbf{u}) = \operatorname{diag}\left(\widetilde{\theta}_\xi(\varepsilon_{\boldsymbol{k}}, \bar{u}_\xi)\right); \quad D_{\boldsymbol{k}} := D_{\boldsymbol{k}}(\varepsilon_{\boldsymbol{k}}, \mathbf{u}_{\boldsymbol{k}}) = \operatorname{diag}\left(\widetilde{\theta}_\xi(\varepsilon_{\boldsymbol{k}}, u_{\boldsymbol{k},\xi})\right)$$
$$\text{and} \quad D^\Delta_{\boldsymbol{k}} = D_{\boldsymbol{k}}(\mathbf{u}) - D_{\boldsymbol{k}} = \operatorname{diag}\left(\widetilde{\theta}_\xi(\varepsilon_{\boldsymbol{k}}, \bar{u}_\xi) - \widetilde{\theta}_\xi(\varepsilon_{\boldsymbol{k}}, u_{\boldsymbol{k},\xi})\right)\,,$$

where $\widetilde{\theta}_\xi = \sqrt{\theta_\xi}$ and $\bar{u}_\xi$ and $\bar{u}_{\boldsymbol{k},\xi}$ are the average tangential components of $\mathbf{u}$ and $\mathbf{u}_{\boldsymbol{k}}$, respectively, along edge $\boldsymbol{e}_\xi$. In terms of these matrices

$$\begin{aligned} Q_{\boldsymbol{k}}(\phi_h,\, \phi_h;\, \varepsilon_{\boldsymbol{k}}, \mathbf{u}) &= \mathbf{z}_{\boldsymbol{k}}^T \left(D_{\boldsymbol{k}} + D_{\boldsymbol{k}}^\Delta\right) M_{\boldsymbol{k}} \left(D_{\boldsymbol{k}} + D_{\boldsymbol{k}}^\Delta\right) \mathbf{z}_{\boldsymbol{k}} \\ &= \mathbf{z}_{\boldsymbol{k}}^T \left(D_{\boldsymbol{k}}\, M_{\boldsymbol{k}}\, D_{\boldsymbol{k}}\right) \mathbf{z} + \mathbf{z}_{\boldsymbol{k}}^T \left(D_{\boldsymbol{k}}^\Delta\, M_{\boldsymbol{k}}\, D_{\boldsymbol{k}}^\Delta\right) \mathbf{z}_{\boldsymbol{k}} + 2\mathbf{z}_{\boldsymbol{k}}^T \left(D_{\boldsymbol{k}}\, M_{\boldsymbol{k}} D_{\boldsymbol{k}}^\Delta\right) \mathbf{z}_{\boldsymbol{k}}\,. \end{aligned}$$

Applying the Schwartz's and Young's inequalities to the last, mixed term gives

$$\begin{aligned} 2\mathbf{z}_{\boldsymbol{k}}^T \left(D_{\boldsymbol{k}}\, M_{\boldsymbol{k}} D_{\boldsymbol{k}}^\Delta\right) \mathbf{z}_{\boldsymbol{k}} &\leq 2 \left(\mathbf{z}_{\boldsymbol{k}}^T \left(D_{\boldsymbol{k}}^\Delta\, M_{\boldsymbol{k}}\, D_{\boldsymbol{k}}^\Delta\right) \mathbf{z}_{\boldsymbol{k}}\right)^{1/2} \left(\mathbf{z}_{\boldsymbol{k}}^T \left(D_{\boldsymbol{k}}\, M_{\boldsymbol{k}}\, D_{\boldsymbol{k}}\right) \mathbf{z}_{\boldsymbol{k}}\right)^{1/2} \\ &\leq \delta\, \mathbf{z}_{\boldsymbol{k}}^T \left(D_{\boldsymbol{k}}^\Delta\, M_{\boldsymbol{k}}\, D_{\boldsymbol{k}}^\Delta\right) \mathbf{z}_{\boldsymbol{k}} + \frac{1}{\delta}\, \mathbf{z}_{\boldsymbol{k}}^T \left(D_{\boldsymbol{k}}\, M_{\boldsymbol{k}}\, D_{\boldsymbol{k}}\right) \mathbf{z}_{\boldsymbol{k}} \end{aligned}$$

with a constant $\delta > 1$ that will be determined later. Therefore,

$$Q_{\boldsymbol{k}}(\phi_h,\, \phi_h;\, \varepsilon_{\boldsymbol{k}}, \mathbf{u}) \geq \left(1 - \frac{1}{\delta}\right) Q_{\boldsymbol{k}}(\phi_h,\, \phi_h;\, \varepsilon_{\boldsymbol{k}}, \mathbf{u}_{\boldsymbol{k}}) - (\delta - 1)\mathbf{z}_{\boldsymbol{k}}^T \left(D_{\boldsymbol{k}}^\Delta\, M_{\boldsymbol{k}}\, D_{\boldsymbol{k}}^\Delta\right) \mathbf{z}_{\boldsymbol{k}}\,. \tag{A.4}$$

We estimate the last term in (A.4) as follows. Note that

$$\mathbf{z}_{\boldsymbol{k}}^T \left(D_{\boldsymbol{k}}^\Delta\, M_{\boldsymbol{k}}\, D_{\boldsymbol{k}}^\Delta\right) \mathbf{z}_{\boldsymbol{k}} \leq \overline{\lambda}(M_{\boldsymbol{k}}) \sum_{\boldsymbol{e}_\xi \in E(\boldsymbol{k})} \left(\widetilde{\theta}_\xi(\varepsilon_{\boldsymbol{k}}, \bar{u}_\xi) - \widetilde{\theta}_\xi(\varepsilon_{\boldsymbol{k}}, \bar{u}_{\boldsymbol{k},\xi})\right)^2 z_\xi^2$$

Recalling that $\widetilde{\theta}_\xi = \sqrt{\theta_\xi}$ allows us to conclude that

$$\left(\widetilde{\theta}_\xi(\varepsilon_{\boldsymbol{k}}, \bar{u}_\xi) - \widetilde{\theta}_\xi(\varepsilon_{\boldsymbol{k}}, \bar{u}_{\boldsymbol{k},\xi})\right)^2 \leq \left|\theta_\xi(\varepsilon_{\boldsymbol{k}}, \bar{u}_\xi) - \theta_\xi(\varepsilon_{\boldsymbol{k}}, \bar{u}_{\boldsymbol{k},\xi})\right|\,.$$

Expanding $\theta_\xi(\varepsilon_{\boldsymbol{k}}, \bar{u}_\xi)$ in a Taylor series about $\bar{u}_{\boldsymbol{k},\xi}$ shows that

$$\left|\theta_\xi(\varepsilon_{\boldsymbol{k}}, \bar{u}_\xi) - \theta_\xi(\varepsilon_{\boldsymbol{k}}, \bar{u}_{\boldsymbol{k},\xi})\right| \leq \left|\frac{\partial \theta_\xi(\varepsilon_{\boldsymbol{k}}, \bar{v}_\xi)}{\partial \bar{u}_\xi}\right| \left|\bar{u}_\xi - \bar{u}_{\boldsymbol{k},\xi}\right|\,,$$

where $\bar{v}_\xi = \lambda \bar{u}_\xi + (1-\lambda)\bar{u}_{\boldsymbol{k},\xi}$ for some $0 < \lambda < 1$. Direct calculation reveals that

$$\left|\frac{\partial \theta_\xi(\varepsilon_{\boldsymbol{k}}, \bar{v}_\xi)}{\partial \bar{u}_\xi}\right| = \frac{h_\xi}{2} \left[\coth p_\xi(\varepsilon_{\boldsymbol{k}}, \bar{v}_\xi) + p_\xi(\varepsilon_{\boldsymbol{k}}, \bar{v}_\xi) \left(1 - \coth^2 p_\xi(\varepsilon_{\boldsymbol{k}}, \bar{v}_\xi)\right)\right] \leq \frac{h_\xi}{2}$$

with the last inequality following from the fact that $0 \leq \coth x + x(1 - \coth^2 x) \leq 1$. Finally, consider the Taylor expansion $\mathbf{u}(\boldsymbol{x}) = \mathbf{u}_{\boldsymbol{k}} + \nabla \mathbf{u}(\boldsymbol{\omega}) \cdot (\boldsymbol{x} - \bar{\boldsymbol{x}})$, where $\mathbf{u}(\bar{\boldsymbol{x}}) = \mathbf{u}_{\boldsymbol{k}}$ and $\boldsymbol{\omega}$ is some point in $\boldsymbol{k}$. Using this expansion we find that

$$\left|\bar{u}_\xi - \bar{u}_{\boldsymbol{k},\xi}\right| = \frac{1}{h_\xi} \left|\int_{\boldsymbol{e}_\xi} \mathbf{t}_\xi \cdot\, \nabla \mathbf{u}(\boldsymbol{\omega}) \cdot (\boldsymbol{x} - \bar{\boldsymbol{x}}) d\ell\right| \leq \frac{1}{h_\xi} \|\nabla \mathbf{u}\|_{0,\infty,\boldsymbol{k}} \int_{\boldsymbol{e}_\xi} |\boldsymbol{x} - \bar{\boldsymbol{x}}| d\ell \leq h_\xi \|\nabla \mathbf{u}\|_{0,\infty,\boldsymbol{k}}\,.$$

After combining all of the above bounds we find that

$$\mathbf{z}_{\boldsymbol{k}}^T \left(D_{\boldsymbol{k}}^\Delta\, M_{\boldsymbol{k}}\, D_{\boldsymbol{k}}^\Delta\right) \mathbf{z}_{\boldsymbol{k}} \leq \overline{\lambda}(M_{\boldsymbol{k}}) \frac{h_{\boldsymbol{k}}^2}{2} \|\nabla \mathbf{u}\|_{0,\infty,\boldsymbol{k}} |\mathbf{z}_{\boldsymbol{k}}|^2\,.$$

Since $\|\nabla\phi_h\|^2_{0,\boldsymbol{k}} \geq \underline{\lambda}(M_{\boldsymbol{k}})|\mathbf{z}_{\boldsymbol{k}}|^2$ it follows that

$$\mathbf{z}_{\boldsymbol{k}}^T \left(D_{\boldsymbol{k}}^{\Delta} M_{\boldsymbol{k}} D_{\boldsymbol{k}}^{\Delta}\right) \mathbf{z}_{\boldsymbol{k}} \leq \kappa(M_{\boldsymbol{k}}) \frac{h_{\boldsymbol{k}}^2}{2} \|\nabla \mathbf{u}\|_{0,\infty,\boldsymbol{k}} \|\nabla\phi_h\|^2_{0,\boldsymbol{k}} .$$

Owing to the regularity assumptions, for a sufficiently small mesh size

$$\kappa(M_{\boldsymbol{k}}) \frac{h_{\boldsymbol{k}}^2}{2} \|\nabla \mathbf{u}\|_{0,\infty,\boldsymbol{k}} \leq \gamma\varepsilon_{\boldsymbol{k}} .$$

Using this inequality together with (A.4) allows us to conclude that

$$Q_{\boldsymbol{k}}(\phi_h,\, \phi_h;\, \varepsilon_{\boldsymbol{k}}, \mathbf{u}) \geq \left(1 - \frac{1}{\delta}\right) Q_{\boldsymbol{k}}(\phi_h,\, \phi_h;\, \varepsilon_{\boldsymbol{k}}, \mathbf{u}_{\boldsymbol{k}}) - (\delta - 1)\gamma\varepsilon_{\boldsymbol{k}} \|\nabla\phi_h\|^2_{0,\boldsymbol{k}} .$$

To complete the proof we combine the above inequality with (A.2) and (A.3):

$$\begin{aligned} a^h_{\boldsymbol{k}}(\phi_h, \phi_h; \varepsilon, \mathbf{u}) &\geq \frac{\varepsilon_{\boldsymbol{k}}}{1+\gamma} \|\nabla\phi_h\|^2_{0,\boldsymbol{k}} + \kappa^{-1}(M_{\boldsymbol{k}}) \left[\left(1 - \frac{1}{\delta}\right) Q_{\boldsymbol{k}}(\phi_h,\, \phi_h;\, \varepsilon_{\boldsymbol{k}}, \mathbf{u}_{\boldsymbol{k}}) - (\delta - 1)\gamma\varepsilon_{\boldsymbol{k}} \|\nabla\phi_h\|^2_{0,\boldsymbol{k}}\right] \\ &= \left(\frac{1}{1+\gamma} - \kappa^{-1}(M_{\boldsymbol{k}})(\delta - 1)\gamma\right) \varepsilon_{\boldsymbol{k}} \|\nabla\phi_h\|^2_{0,\boldsymbol{k}} + \kappa^{-1}(M_{\boldsymbol{k}}) \left(1 - \frac{1}{\delta}\right) Q_{\boldsymbol{k}}(\phi_h,\, \phi_h;\, \varepsilon_{\boldsymbol{k}}, \mathbf{u}_{\boldsymbol{k}}) . \end{aligned}$$

Setting $\delta = 1 + \kappa(M_{\boldsymbol{k}})/(\gamma+1)^2$ gives

$$a^h_{\boldsymbol{k}}(\phi_h, \phi_h; \varepsilon, \mathbf{u}) \geq \frac{\varepsilon_{\boldsymbol{k}} \|\nabla\phi_h\|^2_{0,\boldsymbol{k}}}{(1+\gamma)^2} + \frac{Q_{\boldsymbol{k}}(\phi_h,\, \phi_h;\, \varepsilon_{\boldsymbol{k}}, \mathbf{u}_{\boldsymbol{k}})}{\kappa(M_{\boldsymbol{k}}) + (1+\gamma)^2} \geq C(\gamma)\, a^h_{\boldsymbol{k}}(\phi_h, \phi_h; \varepsilon_{\boldsymbol{k}}, \mathbf{u}_{\boldsymbol{k}}) ,$$

which proves the theorem with $C(\gamma) = 1/(\kappa(M_{\boldsymbol{k}}) + (1+\gamma)^2)$. □

Appendix B. Tensor product finite element spaces. We label the edges of quadrilateral and hexahedral elements according to the versors of the reference coordinate system.

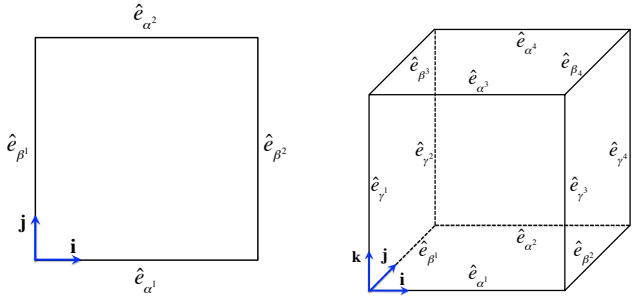


FIG. B.1. *Edge numbering for reference quadrilateral and hexahedral elements.*

A reference quadrilateral $\widehat{\boldsymbol{k}}$ has two pairs of edges $\{\widehat{\boldsymbol{e}}_{\alpha^1}, \widehat{\boldsymbol{e}}_{\alpha^2}\}$ and $\{\widehat{\boldsymbol{e}}_{\beta^1}, \widehat{\boldsymbol{e}}_{\beta^2}\}$, parallel to $\mathbf{i}$ and $\mathbf{j}$, respectively; see Fig. B.1. Let $F_{\boldsymbol{k}}$ be the bilinear map between $\widehat{\boldsymbol{k}}$ and a quadrilateral $\boldsymbol{k} \in K(\Omega)$. The edges of $\boldsymbol{k}$ are images of the reference edges under $F_{\boldsymbol{k}}$ and form the pairs $\{\boldsymbol{e}_{\alpha^1}, \boldsymbol{e}_{\alpha^2}\}$ and $\{\boldsymbol{e}_{\beta^1}, \boldsymbol{e}_{\beta^2}\}$, respectively. The unit tangents to the edges of $\boldsymbol{k}$ form another two pairs labeled by $\{\mathbf{t}_{\alpha^1}, \mathbf{t}_{\alpha^2}\}$ and $\{\mathbf{t}_{\beta^1}, \mathbf{t}_{\beta^2}\}$, respectively. Let $J_{\boldsymbol{k}}(\boldsymbol{x})$ be the Jacobian of $F_{\boldsymbol{k}}$ and define the vector fields

$$\mathbf{t}_{\mathbf{i}}(\boldsymbol{x}) = \frac{\mathbf{J}_{\boldsymbol{k}}(\boldsymbol{x})\,\mathbf{i}}{|\mathbf{J}_{\boldsymbol{k}}(\boldsymbol{x})\,\mathbf{i}|} \qquad \text{and} \qquad \mathbf{t}_{\mathbf{j}}(\boldsymbol{x}) = \frac{\mathbf{J}_{\boldsymbol{k}}(\boldsymbol{x})\,\mathbf{j}}{|\mathbf{J}_{\boldsymbol{k}}(\boldsymbol{x})\,\mathbf{j}|} .$$

It is easy to see that $\{\mathbf{t}_{\alpha^1}, \mathbf{t}_{\alpha^2}\}$ are the restrictions of $\mathbf{t_i}$ to $\{\boldsymbol{e}_{\alpha^1}, \boldsymbol{e}_{\alpha^2}\}$: $\mathbf{t}_{\alpha^1} = \mathbf{t_i}(\boldsymbol{x})|_{\widehat{\boldsymbol{e}}_{\alpha^1}}$ and $\mathbf{t}_{\alpha^2} = \mathbf{t_i}(\boldsymbol{x})|_{\widehat{\boldsymbol{e}}_{\alpha^2}}$. Likewise, $\{\mathbf{t}_{\beta^1}, \mathbf{t}_{\beta^2}\}$ are the restrictions of $\mathbf{t_j}(\boldsymbol{x})$ to $\{\boldsymbol{e}_{\beta^1}, \boldsymbol{e}_{\beta^2}\}$.

A reference hexahedral $\widehat{\boldsymbol{k}}$ has three four-tuples of edges parallel to $\mathbf{i}$, $\mathbf{j}$ and $\mathbf{k}$, respectively, labeled as shown in Figure B.1. The edges of an element $\boldsymbol{k} \in K(\Omega)$ are images of the reference edges under a trilinear map $F_{\boldsymbol{k}}$. They form three four-tuples, which we denote by $\{\boldsymbol{e}_{\alpha^1}, \boldsymbol{e}_{\alpha^2}, \boldsymbol{e}_{\alpha^3}, \boldsymbol{e}_{\alpha^4}\}$, $\{\boldsymbol{e}_{\beta^1}, \boldsymbol{e}_{\beta^2}, \boldsymbol{e}_{\beta^3}, \boldsymbol{e}_{\beta^4}\}$, and $\{\boldsymbol{e}_{\gamma^1}, \boldsymbol{e}_{\gamma^2}, \boldsymbol{e}_{\gamma^3}, \boldsymbol{e}_{\gamma^4}\}$, respectively. Likewise, restrictions of the vector fields

$$\mathbf{t_i}(\boldsymbol{x}) = \frac{\mathbf{J}_{\boldsymbol{k}}(\boldsymbol{x})\,\mathbf{i}}{|\mathbf{J}_{\boldsymbol{k}}(\boldsymbol{x})\,\mathbf{i}|}; \qquad \mathbf{t_j}(\boldsymbol{x}) = \frac{\mathbf{J}_{\boldsymbol{k}}(\boldsymbol{x})\,\mathbf{j}}{|\mathbf{J}_{\boldsymbol{k}}(\boldsymbol{x})\,\mathbf{j}|} \qquad \text{and} \qquad \mathbf{t_k}(\boldsymbol{x}) = \frac{\mathbf{J}_{\boldsymbol{k}}(\boldsymbol{x})\,\mathbf{k}}{|\mathbf{J}_{\boldsymbol{k}}(\boldsymbol{x})\,\mathbf{k}|} \tag{B.1}$$

to four-tuples of reference edges give the unit tangents of the associated four-tuples of edges on $\boldsymbol{k}$.

Stability analysis on tensor product elements requires a notion of an *effective advective velocity* field $\mathbf{u}_{\boldsymbol{k}}$ for element $\boldsymbol{k}$. To define this field consider the normalized projections of the given velocity field $\mathbf{u}$ onto the four-tuples of element edges associated with the versor directions

$$v_{\alpha^i} = \frac{\mathbf{u}\cdot\mathbf{t}_{\alpha^i}}{|\mathbf{u}|}; \quad v_{\beta^i} = \frac{\mathbf{u}\cdot\mathbf{t}_{\beta^i}}{|\mathbf{u}|}; \quad v_{\gamma^i} = \frac{\mathbf{u}\cdot\mathbf{t}_{\gamma^i}}{|\mathbf{u}|}\,, \tag{B.2}$$

and let

$$v_{\mathbf{i}} = \arg\min_i |v_{\alpha^i}|; \quad v_{\mathbf{j}} = \arg\min_i |v_{\beta^i}|; \quad v_{\mathbf{k}} = \arg\min_i |v_{\beta^i}|\,. \tag{B.3}$$

The effective advective velocity on a tensor product element $\boldsymbol{k}$ is then given by

$$\mathbf{u}_{\boldsymbol{k}} = \begin{cases} |\mathbf{u}|\big(v_{\mathbf{i}}\,\mathbf{t_i} + v_{\mathbf{j}}\,\mathbf{t_j}\big) & \text{for quadrilaterals} \\ |\mathbf{u}|\big(v_{\mathbf{i}}\,\mathbf{t_i} + v_{\mathbf{j}}\,\mathbf{t_j} + v_{\mathbf{k}}\,\mathbf{t_k}\big) & \text{for hexahedrons} \end{cases}. \tag{B.4}$$

Finally, note that the edge element basis functions on a quadrilateral element $\boldsymbol{k}$ can be written as

$$\vec{W}_{\alpha^i} = A_i(\boldsymbol{x})\mathbf{J}_{\boldsymbol{k}}^{-T}(\boldsymbol{x})\mathbf{i} \quad \text{and} \quad \vec{W}_{\beta^i} = B_i(\boldsymbol{x})\mathbf{J}_{\boldsymbol{k}}^{-T}(\boldsymbol{x})\mathbf{j}; \quad i = 1, 2; \tag{B.5}$$

where $A_i(\boldsymbol{x})$ and $B_i(\boldsymbol{x})$ are scalar functions, whereas for hexahedral elements

$$\vec{W}_{\alpha^i} = A_i(\boldsymbol{x})\mathbf{J}_{\boldsymbol{k}}^{-T}(\boldsymbol{x})\mathbf{i}; \quad \vec{W}_{\beta^i} = B_i(\boldsymbol{x})\mathbf{J}_{\boldsymbol{k}}^{-T}(\boldsymbol{x})\mathbf{j} \quad \text{and} \quad \vec{W}_{\gamma^i} = C_i(\boldsymbol{x})\mathbf{J}_{\boldsymbol{k}}^{-T}(\boldsymbol{x})\mathbf{k} \quad i = 1, \ldots, 4 \tag{B.6}$$

for some other scalar functions $A_i(\boldsymbol{x})$, $B_i(\boldsymbol{x})$ and $C_i(\boldsymbol{x})$.

REFERENCES

[1] D. N. Arnold, R. S. Falk, and R. Winther. Finite element exterior calculus, homological techniques, and applications. *Acta Numerica*, 15:1–155, 2006.

[2] P. Bochev and M. Hyman. Principles of mimetic discretizations. In D. N. Arnold, P. Bochev, R. Lehoucq, R. Nicolaides, and M. Shashkov, editors, *Compatible Discretizations, Proceedings of IMA Hot Topics Workshop on Compatible Discretizations*, volume IMA 142, pages 89–120. Springer Verlag, 2006.

[3] Pavel Bochev and Kara Peterson. A parameter-free stabilized finite element method for scalar advection-diffusion problems. *Central European Journal of Mathematics*, 11(8):1458–1477, 2013.

[4] Pavel Bochev, Kara Peterson, and Xujiao Gao. A new control volume finite element method for the stable and accurate solution of the drift–diffusion equations on general unstructured grids. *Computer Methods in Applied Mechanics and Engineering*, 254(0):126 – 145, 2013.

[5] Alexander N. Brooks and Thomas 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.

[6] P. Ciarlet. *The Finite Element Method for Elliptic Problems*. SIAM Classics in Applied Mathematics. SIAM, Philadelphia, 2002.

[7] Ramon Codina. Comparison of some finite element methods for solving the diffusion-convection-reaction equation. *Computer Methods in Applied Mechanics and Engineering*, 156(1-4):185 – 210, 1998.

[8] H. C. Elman, D. J. Silvester, and A. J. Wathen. *Finite Elements and Fast Iterative Solvers with Applications in Incompressible Fluid Dynamics*. Numerical Mathematics and Scientific Computation. Oxford University Press, 2005.

[9] A. Ern and J.-L. Guermond. *Theory and Practice of Finite Elements*. Number 159 in Applied Mathematical Sciences. Springer Verlag, New York, 2004.

[10] I. Harari and T. J. R. Hughes. What are C and h?: Inequalities for the analysis and design of finite element methods. *Comput. Meth. Appl. Mech. Eng.*, 97:157–192, 1992.

[11] T. J. R. Hughes. Multiscale phenomena: Green's function, the Dirichlet-to-Neumann map, subgrid scale models, bubbles and the origins of stabilized methods. *Comput. Meth. Appl. Mech. Eng.*, 127:387–401, 1995.

[12] T. J. R. Hughes and A. Brooks. A theoretical framework for Petrov-Galerkin methods with discontinuous weighting functions: Application to the streamline-upwind procedure. In R. H. Gallagher et al, editor, *Finite Elements in Fluids*, volume 4, pages 47–65, New York, 1982. J. Wiley & Sons.

[13] Thomas J.R. Hughes, Michel Mallet, and Mizukami Akira. A new finite element formulation for computational fluid dynamics: Ii. beyond supg. *Computer Methods in Applied Mechanics and Engineering*, 54(3):341 – 355, 1986.

[14] C. Johnson. *Numerical Solution of Partial Differential Equations by the Finite Element Method*. Cambridge University Press, 1992.

[15] P. Knobloch. On the definition of the SUPG parameter. *ETNA*, 32:76–89, 2008.

[16] J. C. Nédélec. Mixed finite elements in R^3. *Numerische Mathematik*, 35:315–341, 1980. 10.1007/BF01396415.

[17] D.L. Scharfetter and H.K. Gummel. Large-signal analysis of a silicon read diode oscillator. *Electron Devices, IEEE Transactions on*, 16(1):64 – 77, jan 1969.

[18] D. Z. Turner, K. B. Nakshatrala, and K. D. Hjelmstad. A stabilized formulation for the advection-diffusion equation using the generalized finite element method. *International Journal for Numerical Methods in Fluids*, 66(1):64–81, 2011.