INTERNATIONAL JOURNAL OF
NUMERICAL ANALYSIS AND MODELING
Volume 16, Number 2, Pages 255–275

AN EMBEDDED SDG METHOD FOR THE CONVECTION-DIFFUSION EQUATION

SIU WUN CHEUNG AND ERIC T. CHUNG

Abstract. In this paper, we present an embedded staggered discontinuous Galerkin method for the convection-diffusion equation. The new method combines the advantages of staggered discontinuous Galerkin (SDG) and embedded discontinuous Galerkin (EDG) method, and results in many good properties, namely local and global conservations, free of carefully designed stabilization terms or flux conditions and high computational efficiency. In applying the new method to convection-dominated problems, the method provides optimal convergence in potential and suboptimal convergence in flux, which is comparable to other existing DG methods, and achieves L^2 stability by making use of a skew-symmetric discretization of the convection term, irrespective of diffusivity. We will present numerical results to show the performance of the method.

Key words. Embedded method, staggered discontinuous Galerkin method, convection-diffusion equation.

1. Introduction

Discontinuous Galerkin (DG) methods were first introduced by Reed and Hill for solving hyperbolic equations [40]. The DG methods have been proven superior to the classical continuous Galerkin (CG) methods for hyperbolic problems. In the past two decades, DG methods have also been applied to second-order elliptic problems. A comprehensive study on DG methods for elliptic problems is given in [1]. The original DG methods for elliptic problems, using polynomial approximations of degree k for both the potential and the flux, converge with optimal order $k + 1$ for the potential but suboptimal order k for the flux. While the same orders of convergence can be obtained by using classical CG finite element methods, the DG methods give rise to a discrete problem with a higher number of degrees of freedom. DG methods have therefore been criticized for its high computation cost and judged to be not being particularly useful for elliptic problems. Later, the hybridizible discontinuous Galerkin (HDG) method was introduced for solving elliptic problems [23]. The HDG method provides optimal orders of convergence for both the potential and the flux in L^2 norm. Moreover, superconvergence can be obtained for the potential through a local postprocessing technique.

In recent years, there are active developments of DG methods for problems in fluid dynamics and wave propagations, see for example [6, 22, 24, 27, 28, 32, 36, 41, 42, 37]. On the other hand, staggered meshes bring the advantages of reducing numerical dissipation in computational fluid dynamics [2, 3, 31], and numerical dispersion in computational wave propagation [9, 10, 11, 12, 13, 14, 17]. Combining the ideas of DG methods and staggered meshes, a new class of staggered discontinuous Galerkin (SDG) methods was proposed for approximations of Stokes system [34], convection-diffusion equation [19], and incompressible Navier-Stokes equations [7]. The new class of SDG methods possesses many good properties, including local and global conservations, stability in energy, and optimal convergence. For a more

Received by the editors July 26, 2017.
2000 *Mathematics Subject Classification.* 65N12, 65N30.

complete discussion on the SDG method, see also [11, 12, 13, 14, 18, 19, 35] and the references therein.

In [15, 16], it was shown that the SDG method can be regarded as a limit of the HDG method. The SDG method can be obtained from the HDG method by setting the stabilization parameter on a set of edges to be zero and letting the parameter on another set of edges to infinity. As a result, the SDG method inherits the advantages of the HDG method, including superconvergence through the use of a local postprocessing technique. Furthermore, in the SDG method for incompressible Navier-Stokes equations [7], using the postprocessing and a spectro-consistent discretizations with a novel splitting of the diffusion and the convection term, stability in L^2 energy is achieved.

The embedded discontinuous Galerkin (EDG) method was first introduced for solving the linear shell problems [30]. Later, an EDG method for solving second order elliptic problems was discussed and analyzed in [26]. The EDG method was obtained from HDG method by enforcing strong continuity for hybrid unknowns [23]. This greatly reduces the number of degrees of freedom in the globally coupled system and makes the EDG method has a higher computational efficiency compared with other DG methods. As a tradeoff for this advantage, the EDG method is not locally conservative and loses the optimal convergence in the flux achieved by the HDG method [38]. The loss in accuracy makes the EDG method a less attractive candidate compared with the HDG method. However, the optimal order of convergence for HDG method is also lost in the case of convection-dominated problems as shown the numerical examples [25]. In this case, the EDG method becomes appealing alternative to all other DG methods including the HDG method, since it has a higher computational efficiency and the same orders of convergence. On the other hand, compared with the CG finite element method, the EDG method provides the same sparsity structure of the stiffness matrix after static condensation, whilist the EDG method is more robust, accurate and stable than the CG finite element method in convection-dominated problems. We remark that the multiscale discontinuous Galerkin (MDG) method [5, 33] are related to the EDG method, which is originally proposed for the convection-diffusion problems. The MDG method and the EDG method are both designed for a globally continuous approximation of the solution and can give rise to identical schemes. Recently, the EDG method has been proposed on Euler equations and Navier-Stokes equations [38, 39]. Due to the advantages shared with other DG methods and the high computational efficiency compared with other DG methods, the EDG method has also been applied to challenging problems in computational fluid dynamics, such as implicit large eddy simulation [29].

In this paper, we propose a combination of the SDG method and the EDG method for the convection-diffusion equation. The new method seeks approximations in the SDG locally conforming finite element spaces, which gives rise to a flux formulation without introducing any carefully designed stabilization terms or flux conditions as in other DG methods. The new method further reduces the size of the global discrete problem compared with SDG method by restricting the numerical approximation for the primal unknown to lie in a proper subspace of the SDG finite element space. Moreover, the new method inherits the stability in L^2 energy thanks to spectro-consistent discretizations in the SDG method. The convergence is optimal with order $k+1$ for the unknown function and suboptimal with order k for its gradient, which are comparable to all other DG methods for convection-dominated problems.

The paper is organized as follows. In Section 2, we will have a derivation on the method. Next, in Section 3, we will provide a stability analysis of the method. Then, in Section 5, we will present extensive numerical examples to see the performance of our method. Finally, a conclusion is given.

2. Method description

2.1. Model problem. Let $\Omega \subset \mathbb{R}^2$ be a polygonal domain. We consider the steady-state convection-diffusion equation with homogeneous Dirichlet boundary condition:

$$\begin{aligned} -\mu\mathbf{\Delta} u + \operatorname{div}(\mathbf{b}u) &= f \quad \text{in } \Omega, \\ u &= 0 \quad \text{on } \partial\Omega. \end{aligned} \tag{1}$$

Here u is the unknown function to be approximated, $\mathbf{b} = (b_1, b_2)$ is a divergence-free convection field, f is a given source term and g is a given boundary condition. Also, μ is the diffusivity, which is assumed to be constants throughout the domain Ω. Before we start the derivation of our method, we shall state the variational formulation of the problem. Suppose $\mathbf{b} \in L^\infty(\Omega)$ and $f \in H^{-1}(\Omega)$. The variational formulation of the convection-diffusion equation is given by: find $u \in H_0^1(\Omega)$ such that for any $v \in H_0^1(\Omega)$, we have

$$\mu(\nabla u, \nabla v)_{0,\Omega} + (\operatorname{div}(\mathbf{b}u), v)_{0,\Omega} = (f, v)_{0,\Omega}. \tag{2}$$

Here $(\cdot,\cdot)_{0,\Omega}$ denotes the standard $L^2(\Omega)$ inner product.

We will derive a mixed method for the problem. Since $\operatorname{div}\mathbf{b} = 0$, it is direct to see that

$$\operatorname{div}(\mathbf{b}u) = \mathbf{b}\cdot\nabla u + (\operatorname{div}\mathbf{b})u = \mathbf{b}\cdot\nabla u. \tag{3}$$

We can therefore rewrite (1) as

$$\begin{aligned} -\mu\mathbf{\Delta} u + \frac{1}{2}\operatorname{div}(\mathbf{b}u) + \frac{1}{2}\mathbf{b}\cdot\nabla u &= f \quad \text{in } \Omega, \\ u &= 0 \quad \text{on } \partial\Omega. \end{aligned} \tag{4}$$

We introduce the auxiliary variables

$$\begin{aligned} \mathbf{w} &= \sqrt{\mu}\,\nabla u - \frac{1}{2\sqrt{\mu}}\mathbf{b}u, \\ \mathbf{p} &= \mathbf{b}u. \end{aligned} \tag{5}$$

Then (4) can be reformulated as a system of first-order linear PDEs:

$$\begin{aligned} -\sqrt{\mu}\operatorname{div}\mathbf{w} + \frac{1}{2\sqrt{\mu}}\mathbf{b}\cdot\mathbf{w} + \frac{1}{4\mu}\mathbf{b}\cdot\mathbf{p} &= f \quad \text{in } \Omega, \\ u &= 0 \quad \text{on } \partial\Omega. \end{aligned} \tag{6}$$

2.2. Staggered meshes. Let $\mathcal{T}_u$ be a triangulation of the two-dimensional domain Ω by a set of triangles without hanging nodes. We introduce the notation $\mathcal{F}_u$ to denote the set of all edges in the triangulation $\mathcal{T}_u$ and $\mathcal{F}_u^0$ to denote the subset of all interior edges in $\mathcal{F}_u$ excluding those on the boundary of Ω. We also denote the set of all vertices in $\mathcal{T}_u$ by $\mathcal{N}_u$. For each triangle in $\mathcal{T}_u$, we take an interior point ν, denote the initial triangle by $\mathcal{S}(\nu)$, and divide $\mathcal{S}(\nu)$ into three triangles by joining the point ν and the three vertices of $\mathcal{S}(\nu)$. We also denote the set of all interior points ν by $\mathcal{N}$, the set of all new edges generated by the subdivision of triangles by $\mathcal{F}_p$, and the triangulation after subdivision by $\mathcal{T}$. Note that the interior point ν of each triangle in $\mathcal{T}_u$ should be chosen such that the new triangulation $\mathcal{T}$ observes the shape regularity criterion. In practice, we can simply choose ν as the centroid of the triangle. Also, $\mathcal{F} = \mathcal{F}_u \cup \mathcal{F}_p$ denotes the set of all edges of triangles in $\mathcal{T}$ and $\mathcal{F}^0 = \mathcal{F}_u^0 \cup \mathcal{F}_p$ denotes the set of all interior edges of triangles in $\mathcal{T}$. For each edge $e \in \mathcal{F}_u$, we let $\mathcal{R}(e)$ be the union of the all triangles in the new triangulation $\mathcal{T}$ sharing the edge e. Figure 1 demonstrates these definitions. The edges $e \in \mathcal{F}_u$ are represented in solid lines and the $e \in \mathcal{F}_p$ are represented in dotted lines.

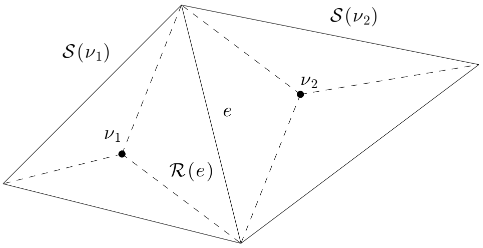

Figure 1. An illustration of the staggered mesh in two dimensions.

For each edge $e \in \mathcal{F}$, we will also define a unit normal vector $\mathbf{n}_e$ in the following way. If $e \in \mathcal{F} \setminus \mathcal{F}^0$ is a boundary edge, then we define $\mathbf{n}_e$ as the outward unit normal vector of e from Ω. If $e \in \mathcal{F}^0$ is an interior edge, then $\mathbf{n}_e$ is fixed as one of the two possible unit normal vectors on e. When it is clear that which edge we are considering, we omit the index e and write the unit normal vector as $\mathbf{n}$.

To end this section, we define the jumps in the following way: for any edge $e \in \mathcal{F}$, denote one of the triangles in the refined triangulation $\mathcal{T}$, which contains e by τ^+, and denote the other triangle, if exists, by τ^-. The outward unit normal vectors on e in τ^+ and τ^- are denoted by $\mathbf{n}^+$ and $\mathbf{n}^-$, respectively. Also, for any quantity ϕ, the notations $\phi^\pm$ are defined on the edge e by the values of $\phi|_{\tau^\pm}$ restricted on e. Then, if ϕ is a scalar quantity, the notation $[\phi]$ over an edge e defined as

$$[\phi]|_e := (\mathbf{n} \cdot \mathbf{n}^+)\phi^+ + (\mathbf{n} \cdot \mathbf{n}^-)\phi^-. \tag{7}$$

If $\mathbf{\Phi}$ is a vector quantity, then the notation $[\mathbf{\Phi} \cdot \mathbf{n}]$ is similarly defined as

$$[\mathbf{\Phi} \cdot \mathbf{n}]|_e := (\mathbf{n} \cdot \mathbf{n}^+)(\mathbf{\Phi}^+ \cdot \mathbf{n}) + (\mathbf{n} \cdot \mathbf{n}^-)(\mathbf{\Phi}^- \cdot \mathbf{n}). \tag{8}$$

2.3. SDG and ESDG finite element spaces. We will define the finite element spaces. Let $k \geq 0$ be a non-negative integer. Let $\tau \in \mathcal{T}$ and $e \in \mathcal{F}$. We define $P^k(\tau)$ and $P^k(e)$ as the space of polynomials whose order is not greater than k on τ and e, respectively. We will also define norms on the spaces. We use the standard notations $\| \cdot \|_{0,\Omega}$ to denote the standard L^2 norm on Ω and $\| \cdot \|_{0,e}$ to denote the L^2 norm on an edge e.

First, we define the following locally $H^1(\Omega)$-conforming finite element space:

$$U^h = \{v \,:\, v|_\tau \in P^k(\tau);\ \tau \in \mathcal{T};\ v \text{ is continuous over } e \in \mathcal{F}_u^0;\ v|_{\partial\Omega} = 0\}. \tag{9}$$

Note that for any $v \in U^h$, we have $v|_{\mathcal{R}(e)} \in H^1(\mathcal{R}(e))$ for each edge $e \in \mathcal{F}_u$. We define the following discrete L^2-norm $\| \cdot \|_X$ and discrete H^1-norm $\| \cdot \|_Z$ on the space U^h:

$$\begin{aligned}
\|v\|_X &= \left(\|v\|_{0,\Omega}^2 + \sum_{e\in\mathcal{F}_u^0} h_e\|v\|_{0,e}^2\right)^{\frac{1}{2}},\\
\|v\|_Z &= \left(\|\nabla_h v\|_{0,\Omega}^2 + \sum_{e\in\mathcal{F}_p} h_e^{-1}\|[v]\|_{0,e}^2\right)^{\frac{1}{2}}.
\end{aligned} \tag{10}$$

Next, we define the following ESDG finite element space, which is a proper subspace of the SDG finite element space:

$$\widetilde{U}^h = \{v \in U^h \ : \ v \text{ is continuous at } \eta;\ \eta \in \mathcal{N}_u\}. \tag{11}$$

Note that the test functions $v \in \widetilde{U}^h$ are continuous at only all the nodes in the initial grid but not the nodes at the refined grid, are therefore they are not globally continuous. We remark that for the EDG method [26], the space of the numerical trace is imposed with a global continuity on the skeleton of the mesh.

Finally, we define the following locally $H(\text{div};\Omega)$-conforming finite element space:

$$W^h = \{\mathbf{\Psi} \ : \ \mathbf{\Psi}|_\tau \in P^k(\tau)^2;\ \tau \in \mathcal{T};\ \mathbf{\Psi}\cdot\mathbf{n} \text{ is continuous over } e \in \mathcal{F}_p\}. \tag{12}$$

Note that for any $\mathbf{\Psi} \in W^h$, we have $\mathbf{\Psi}|_{\mathcal{S}(\nu)} \in H(\text{div};\mathcal{S}(\nu))$ for each $\nu \in \mathcal{N}$. We define the following discrete L^2-norm $\|\cdot\|_{X'}$ and discrete $H(\text{div};\Omega)$-norm $\|\cdot\|_{Z'}$ on the space W^h:

$$\begin{aligned}
\|\mathbf{\Psi}\|_{X'} &= \left(\|\mathbf{\Psi}\|_{0,\Omega}^2 + \sum_{e\in\mathcal{F}_p} h_e\|\mathbf{\Psi}\cdot\mathbf{n}\|_{0,e}^2\right)^{\frac{1}{2}},\\
\|\mathbf{\Psi}\|_{Z'} &= \left(\|\text{div}_h\mathbf{\Psi}\|_{0,\Omega}^2 + \sum_{e\in\mathcal{F}_u^0} h_e^{-1}\|[\mathbf{\Psi}\cdot\mathbf{n}]\|_{0,e}^2\right)^{\frac{1}{2}}.
\end{aligned} \tag{13}$$

By [11, 12], there exist interpolation operators $\mathcal{I}$ onto $\widetilde{U}^h$ and $\mathcal{J}$ onto W^h such that

$$\begin{aligned}
B_h^*(u-\mathcal{I}u,\mathbf{\Psi}) &= 0 \text{ for all } \mathbf{\Psi}\in W^h,\\
B_h(\mathbf{w}-\mathcal{J}\mathbf{w}, v) &= 0 \text{ for all } v \in U^h,
\end{aligned} \tag{14}$$

and

$$\begin{aligned}
\|u-\mathcal{I}u\|_{0,\Omega} &\le Ch^{k+1}|u|_{H^{k+1}(\Omega)},\\
\|\mathbf{w}-\mathcal{J}\mathbf{w}\|_{0,\Omega} &\le Ch^{k+1}|\mathbf{w}|_{[H^{k+1}(\Omega)]^2}.
\end{aligned} \tag{15}$$

2.4. Derivation of the method. We will derive the discrete problem in our SDG formulation starting from the system of first order equations in (5) and (6).

Multiplying the first equation of (5) by $\mathbf{\Psi}_1 \in W^h$ and integrating over $\mathcal{S}(\nu)$ for $\nu \in \mathcal{N}$, we obtain

$$\begin{aligned}
\int_{\mathcal{S}(\nu)} \mathbf{w}\cdot\mathbf{\Psi}_1\,dx =& -\sqrt{\mu}\int_{\mathcal{S}(\nu)} u(\text{div }\mathbf{\Psi}_1)\,dx\\
&+\sqrt{\mu}\int_{\partial\mathcal{S}(\nu)} u(\mathbf{\Psi}_1\cdot\mathbf{n})\,d\sigma - \frac{1}{2\sqrt{\mu}}\int_{\mathcal{S}(\nu)}\mathbf{p}\cdot\mathbf{\Psi}_1\,dx.
\end{aligned} \tag{16}$$

Similarly, multiplying the second equation of (5) by $\mathbf{\Psi}_2 \in W^h$ and integrating over $\mathcal{S}(\nu)$ for $\nu \in \mathcal{N}$, we have

$$\int_{\mathcal{S}(\nu)} \mathbf{p}\cdot\mathbf{\Psi}_2 \, dx = \int_{\mathcal{S}(\nu)} u(\mathbf{b}\cdot\mathbf{\Psi}_2)\, dx. \tag{17}$$

Finally, multiplying the first equation of (6) by $v \in U^h$ and integrating over $\mathcal{R}(e)$ for $e \in \mathcal{F}_u^0$, we have

$$\begin{aligned}\sqrt{\mu}\int_{\mathcal{R}(e)} \mathbf{w}\cdot\nabla v\, dx - \sqrt{\mu}\int_{\partial\mathcal{R}(e)} (\mathbf{w}\cdot\mathbf{n}) v\, d\sigma + \frac{1}{2\sqrt{\mu}}\int_{\mathcal{R}(e)} (\mathbf{b}\cdot\mathbf{w})\, v\, dx \\ + \frac{1}{4\mu}\int_{\mathcal{R}(e)} (\mathbf{b}\cdot\mathbf{p})\, v\, dx = \int_{\mathcal{R}(e)} fv\, dx.\end{aligned} \tag{18}$$

Summing those equations in (16)–(18) over all $\mathcal{R}(e)$ and $\mathcal{S}(\nu)$, we obtain the staggered discontinuous Galerkin method for (1) proposed in [19]: find $(u_h, \mathbf{w}_h, \mathbf{p}_h) \in U^h \times W^h \times W^h$ such that for any $v \in U^h$, $\mathbf{\Psi}_1, \mathbf{\Psi}_2 \in W^h$, we have

$$\begin{aligned}\sqrt{\mu}B_h(\mathbf{w}_h, v) + \frac{1}{2\sqrt{\mu}} R_h\left(\mathbf{w}_h + \frac{1}{2\sqrt{\mu}}\mathbf{p}_h, v\right) &= (f, v)_{0,\Omega}, \\ \sqrt{\mu}B_h^*(u_h, \mathbf{\Psi}_1) - \frac{1}{2\sqrt{\mu}}(\mathbf{p}_h, \mathbf{\Psi}_1)_{0,\Omega} &= (\mathbf{w}_h, \mathbf{\Psi}_1)_{0,\Omega}, \\ R_h^*(u_h, \mathbf{\Psi}_2) &= (\mathbf{p}_h, \mathbf{\Psi}_2)_{0,\Omega},\end{aligned} \tag{19}$$

where bilinear forms $B_h(\mathbf{\Psi}, v)$ and $B_h^*(v, \mathbf{\Psi})$ are defined as

$$\begin{aligned}B_h(\mathbf{\Psi}, v) &= \int_\Omega \mathbf{\Psi}\cdot\nabla_h v\, dx - \sum_{e\in\mathcal{F}_p}\int_e (\mathbf{\Psi}\cdot\mathbf{n})\,[v]\, d\sigma, \\ B_h^*(v, \mathbf{\Psi}) &= -\int_\Omega v\, \mathrm{div}_h\, \mathbf{\Psi}\, dx + \sum_{e\in\mathcal{F}_u^0}\int_e v\,[\mathbf{\Psi}\cdot\mathbf{n}]\, d\sigma,\end{aligned} \tag{20}$$

and the bilinear forms $R_h(\mathbf{\Psi}, v)$ and $R_h^*(v, \mathbf{\Psi})$ are defined as

$$\begin{aligned}R_h(\mathbf{\Psi}, v) &= \int_\Omega (\mathbf{b}\cdot\mathbf{\Psi})\, v\, dx, \\ R_h^*(v, \mathbf{\Psi}) &= \int_\Omega v\,(\mathbf{b}\cdot\mathbf{\Psi})\, dx.\end{aligned} \tag{21}$$

Now, if we consider only the test functions $v \in \widetilde{U}^h$ in (18), and we seek an approximation $\widetilde{u}_h \in \widetilde{U}^h$ for the unknown function u, we obtain the embedded staggered discontinuous Galerkin method for (1): find $(\widetilde{u}_h, \widetilde{\mathbf{w}}_h, \widetilde{\mathbf{p}}_h) \in \widetilde{U}^h \times W^h \times W^h$ such that for any $v \in \widetilde{U}^h$, $\mathbf{\Psi}_1, \mathbf{\Psi}_2 \in W^h$, we have

$$\begin{aligned}\sqrt{\mu}B_h(\widetilde{\mathbf{w}}_h, v) + \frac{1}{2\sqrt{\mu}} R_h\left(\widetilde{\mathbf{w}}_h + \frac{1}{2\sqrt{\mu}}\widetilde{\mathbf{p}}_h, v\right) &= (f, v)_{0,\Omega}, \\ \sqrt{\mu}B_h^*(\widetilde{u}_h, \mathbf{\Psi}_1) - \frac{1}{2\sqrt{\mu}}(\widetilde{\mathbf{p}}_h, \mathbf{\Psi}_1)_{0,\Omega} &= (\widetilde{\mathbf{w}}_h, \mathbf{\Psi}_1)_{0,\Omega}, \\ R_h^*(\widetilde{u}_h, \mathbf{\Psi}_2) &= (\widetilde{\mathbf{p}}_h, \mathbf{\Psi}_2)_{0,\Omega}.\end{aligned} \tag{22}$$

By [12], the two bilinear forms in (20) satisfy the adjoint relation

$$B_h(\mathbf{\Psi}, v) = B_h^*(v, \mathbf{\Psi}) \tag{23}$$

for all $v \in U_h$ and $\mathbf{\Psi} \in W^h$. The bilinear forms B_h and B_h^* are also continuous with respect to suitable discrete norms

$$\begin{aligned} |B_h(\mathbf{\Psi}, v)| &\leq \|\mathbf{\Psi}\|_{X'} \|v\|_Z, \\ |B_h^*(v, \mathbf{\Psi})| &\leq \|v\|_X \|\mathbf{\Psi}\|_{Z'}, \end{aligned} \tag{24}$$

for all $v \in U_h$ and $\mathbf{\Psi} \in W^h$. Moreover, the bilinear forms B_h and B_h^* satisfy a pair of inf-sup conditions: there exists constants β_1 and β_2, independent of h, such that

$$\begin{aligned} \inf_{v \in U^h \setminus \{0\}} \sup_{\mathbf{\Psi} \in W^h \setminus \{\mathbf{0}\}} \frac{B_h(\mathbf{\Psi}, v)}{\|\mathbf{\Psi}\|_{X'} \|v\|_Z} &\geq \beta_1, \\ \inf_{\mathbf{\Psi} \in W^h \setminus \{\mathbf{0}\}} \sup_{v \in U^h \setminus \{0\}} \frac{B_h^*(v, \mathbf{\Psi})}{\|v\|_X \|\mathbf{\Psi}\|_{Z'}} &\geq \beta_2. \end{aligned} \tag{25}$$

Also, it is obvious that the two bilinear forms in (21) satisfy

$$R_h^*(v, \mathbf{\Psi}) = R_h(\mathbf{\Psi}, v) \tag{26}$$

for all $v \in U_h$ and $\mathbf{\Psi} \in W^h$.

2.5. Linear system. In this section, we derive the linear systems resulting from (19) and (22). We denote the corresponding matrix representation of the bilinear forms B_h and R_h by B and R, respectively. Then by the adjoint properties, the matrix representation of the bilinear forms B_h^* and R_h^* are given by B^T and R^T, respectively. Also, the notations for the finite element solutions would be abused to denote their corresponding vector representations.

Using these notations, we can write the SDG method (19) as a linear system of algebraic equations. The second equation of (19) can be written as

$$\sqrt{\mu} B^T u_h - \frac{1}{2\sqrt{\mu}} M \mathbf{p}_h = M \mathbf{w}_h, \tag{27}$$

where M is the mass matrix for the space W^h. Similarly, the last equation of (19) can be written as

$$R^T u_{h,1} = M \mathbf{p}_h. \tag{28}$$

Lastly, the first equations of (19) can be written as

$$\sqrt{\mu} B \mathbf{w}_h + \frac{1}{2\sqrt{\mu}} R \left(\mathbf{w}_h + \frac{1}{2\sqrt{\mu}} \mathbf{p}_h \right) = f_h. \tag{29}$$

We can now obtain a linear system with the unknowns $\mathbf{w}_h$ and $\mathbf{p}_h$ eliminated. Combining (27) and (28), we have

$$\begin{aligned} \mathbf{w}_h &= M^{-1} \left(\sqrt{\mu} B^T u_h - \frac{1}{2\sqrt{\mu}} R^T u_h \right), \\ \mathbf{p}_h &= M^{-1} R^T u_h. \end{aligned} \tag{30}$$

We note that the elimination can be done by solving small problems in each $\mathcal{S}(\nu)$ since M is a block diagonal matrix with each block corresponding to the mass matrix of $W^h|_{\mathcal{S}(\nu)}$.

We further introduce the notations

$$\begin{aligned} \mathbf{\Delta}_h &= -B M^{-1} B^T, \\ \mathbf{b} \cdot \nabla_h &= -\frac{1}{2} B M^{-1} R^T + \frac{1}{2} R M^{-1} B^T, \\ A &= -\mu \mathbf{\Delta}_h + \mathbf{b} \cdot \nabla_h. \end{aligned} \tag{31}$$

We note that the discrete diffusion operator $-\Delta_h$ is symmetric and positive-definite, and the discrete convection operator $\mathbf{b}\cdot\nabla_h$ is skew-symmetric. Combining (29) and (30), the algebraic system of the discrete problem (19) can then be reduced to

$$Au_h = f_h. \tag{32}$$

Now, if we denote the matrix representation of the canonical embedding $\iota : \widetilde{U}^h \to U^h$ by P, then the matrix representations $\widetilde{B}$ and $\widetilde{R}$ of the bilinear forms $B_h|_{W^h\times\widetilde{U}^h}$ and $R_h|_{W^h\times\widetilde{U}^h}$ are related to B and R by

$$\begin{aligned}\widetilde{B} &= PB,\\ \widetilde{R} &= PR.\end{aligned} \tag{33}$$

The corresponding matrix represenations for the discrete diffusion operator and discrete convection operator are

$$\begin{aligned}\widetilde{\mathbf{\Delta}}_h &= -\widetilde{B}M^{-1}\widetilde{B}^T = P\mathbf{\Delta}_hP^T,\\ \mathbf{b}\cdot\widetilde{\nabla}_h &= -\frac{1}{2}\widetilde{B}M^{-1}\widetilde{R}^T + \frac{1}{2}\widetilde{R}M^{-1}\widetilde{B}^T = P(\mathbf{b}\cdot\nabla_h)P^T.\end{aligned} \tag{34}$$

Therefore the algebraic system of the discrete problem (22) is given by

$$\widetilde{A}\widetilde{u}_h = \widetilde{f}_h, \tag{35}$$

where

$$\begin{aligned}\widetilde{A} &= PAP^T,\\ \widetilde{f}_h &= Pf_h.\end{aligned} \tag{36}$$

We remark that in the embedded SDG method, the discretization of the diffusion operator $-\widetilde{\mathbf{\Delta}}_h$ is still symmetric and positive definite. Similarly, the discretization of the convection operator $\mathbf{b}\cdot\widetilde{\nabla}_h$ is still skew-symmetric. Therefore the spectro-consistent discretization is preserved in the new method.

3. Stability analysis

We start this section by stating a stability result in L^2 energy for the variational formulation (2).

Lemma 3.1. *Let $u \in H_0^1(\Omega)$ be the weak solution of the variational formulation (2) of the convection-diffusion equation. Then we have*

$$\mu\|\nabla u\|_{0,\Omega}^2 = (f,u)_{0,\Omega}. \tag{37}$$

Proof. In (2), we take a test function $v = u$. Then we have

$$\mu\|\nabla u\|_{0,\Omega}^2 + (\operatorname{div}(\mathbf{b}u), u)_{0,\Omega} = (f,u)_{0,\Omega}. \tag{38}$$

On the other hand, using an integration by parts and the relation (3), we have

$$\begin{aligned}(\operatorname{div}(\mathbf{b}u), u)_{0,\Omega} &= -(\mathbf{b}u, \nabla u)_{0,\Omega}\\ &= -(\mathbf{b}\cdot\nabla u, u)_{0,\Omega}\\ &= -(\operatorname{div}(\mathbf{b}u), u)_{0,\Omega},\end{aligned} \tag{39}$$

which implies $(\operatorname{div}(\mathbf{b}u), u)_{0,\Omega} = 0$. This completes the proof. □

We will next see that the ESDG method provides a similar stability result. In the SDG method, the stability in L^2 energy is due to a spectro-consistent discretizations with the splitting of the diffusion and the convection term proposed in [19]. The stability in L^2 energy in a numerical method for the convection-diffusion problems is a kind of measure of how well the numerical solution approximates the analytical solution, and has significant effects on the quality of the numerical solution (see, for example, [7], [8]; also see Section 5.3). The ESDG method inherits the stability in L^2 energy from the SDG method due to the same spectro-consistent discretization structure.

The unknowns in the space W^h in both the SDG method and the ESDG method give rise to an approximation of the flux $\mathbf{z} = \nabla u$ in the space W^h. For the ESDG method, suppose $(\widetilde{u}_h, \widetilde{\mathbf{w}}_h, \widetilde{\mathbf{p}}_h) \in \widetilde{U}^h \times W^h \times W^h$ is the solution of (22). An approximation $\widetilde{\mathbf{z}}_h \in W^h$ for the flux $\mathbf{z}$ is then given by

$$\widetilde{\mathbf{z}}_h = \frac{1}{\sqrt{\mu}}\widetilde{\mathbf{w}}_h + \frac{1}{2\mu}\widetilde{\mathbf{p}}_h = M^{-1}B^TP^T\widetilde{u}_h. \tag{40}$$

Likewise for the SDG method, suppose $(u_h, \mathbf{w}_h, \mathbf{p}_h) \in U^h \times W^h \times W^h$ is the solution of (19). An approximation $\mathbf{z}_h \in W^h$ for the flux $\mathbf{z}$ is given by

$$\mathbf{z}_h = \frac{1}{\sqrt{\mu}}\mathbf{w}_h + \frac{1}{2\mu}\mathbf{p}_h = M^{-1}B^Tu_h. \tag{41}$$

We are now ready to state the stability result for the ESDG method:

Lemma 3.2. *Let* $(\widetilde{u}_h, \widetilde{\mathbf{w}}_h, \widetilde{\mathbf{p}}_h) \in \widetilde{U}^h \times W^h \times W^h$ *be the numerical solution of the ESDG method* (22). *Then we have*

$$\mu\|\widetilde{\mathbf{z}}_h\|_{0,\Omega}^2 = (f, \widetilde{u}_h)_{0,\Omega}, \tag{42}$$

where $\widetilde{\mathbf{z}}_h \in W^h$ *is defined in* (40).

Proof. In (22), we take test functions as follows:

$$\begin{aligned} v &= \widetilde{u}_h, \\ \boldsymbol{\Psi}_1 &= -\widetilde{\mathbf{w}}_h \\ \boldsymbol{\Psi}_2 &= -\frac{1}{2}\widetilde{\mathbf{z}}_h. \end{aligned} \tag{43}$$

Then we have

$$\begin{aligned} \sqrt{\mu}B_h(\widetilde{\mathbf{w}}_h, \widetilde{u}_h) + \frac{1}{2\sqrt{\mu}}R_h\left(\widetilde{\mathbf{w}}_h + \frac{1}{2\sqrt{\mu}}\widetilde{\mathbf{p}}_h, \widetilde{u}_h\right) &= (f, \widetilde{u}_h)_{0,\Omega}, \\ -\sqrt{\mu}B_h^*(\widetilde{u}_h, \widetilde{\mathbf{w}}_h) + \frac{1}{2\sqrt{\mu}}(\widetilde{\mathbf{p}}_h, \widetilde{\mathbf{w}}_h)_{0,\Omega} &= -(\widetilde{\mathbf{w}}_h, \widetilde{\mathbf{w}}_h)_{0,\Omega}, \\ -\frac{1}{2}R_h^*(\widetilde{u}_h, \widetilde{\mathbf{z}}_h) &= -\frac{1}{2}(\widetilde{\mathbf{p}}_h, \widetilde{\mathbf{z}}_h)_{0,\Omega}. \end{aligned} \tag{44}$$

Recalling the definition of $\widetilde{\mathbf{z}}_h$ in (40), the above equations can be rewritten as

$$\begin{aligned} \sqrt{\mu}B_h(\widetilde{\mathbf{w}}_h, \widetilde{u}_h) + \frac{1}{2}R_h\left(\widetilde{\mathbf{z}}_h, \widetilde{u}_h\right) &= (f, \widetilde{u}_h)_{0,\Omega}, \\ -\sqrt{\mu}B_h^*(\widetilde{u}_h, \widetilde{\mathbf{w}}_h) + \sqrt{\mu}(\widetilde{\mathbf{z}}_h, \widetilde{\mathbf{w}}_h)_{0,\Omega} &= 0, \\ -\frac{1}{2}R_h^*(\widetilde{u}_h, \widetilde{\mathbf{z}}_h) + \frac{1}{2}(\widetilde{\mathbf{z}}_h, \widetilde{\mathbf{p}}_h)_{0,\Omega} &= 0. \end{aligned} \tag{45}$$

Summing up the equations in (45), using the adjoint relations (23) and (26), and the definition of $\widetilde{\mathbf{z}}_h$ in (40) again, we have

$$\begin{aligned}\sqrt{\mu}(\widetilde{\mathbf{z}}_h,\widetilde{\mathbf{w}}_h)_{0,\Omega}+\frac{1}{2}(\widetilde{\mathbf{z}}_h,\widetilde{\mathbf{p}}_h)_{0,\Omega}&=(f,\widetilde{u}_h)_{0,\Omega}\\ \mu\left(\widetilde{\mathbf{z}}_h,\frac{1}{\sqrt{\mu}}\widetilde{\mathbf{w}}_h+\frac{1}{2\mu}\widetilde{\mathbf{p}}_h\right)_{0,\Omega}&=(f,\widetilde{u}_h)_{0,\Omega}\\ \mu\|\widetilde{\mathbf{z}}_h\|^2_{0,\Omega}&=(f,\widetilde{u}_h)_{0,\Omega}.\end{aligned}\tag{46}$$

□

An important message from Lemma 3.2 is that the convection field $\mathbf{b}$ vanishes in the above L^2 stability estimate for $\widetilde{\mathbf{z}}_h$. This makes the ESDG approximation mimics the weak solution $\mathbf{z}$ better as the convection field $\mathbf{b}$ also vanishes in the above L^2 stability estimate for $\mathbf{z}$ in Lemma 3.1. This is an advantage brought by the novel splitting of the convection term and the diffusion term.

To end this section, we establish the main stability result.

Theorem 3.3. *Let $(\widetilde{u}_h,\widetilde{\mathbf{w}}_h,\widetilde{\mathbf{p}}_h)\in\widetilde{U}^h\times W^h\times W^h$ be the numerical solution of the ESDG method* (22). *Then we have*

$$\mu\|\widetilde{u}_h\|_Z\le C\|f\|_{0,\Omega},\tag{47}$$

where C is a constant independent of mesh size and diffusivity.

Proof. By Lemma 3.2, we have

$$\mu\|\widetilde{z}_h\|^2_{0,\Omega}=(f,\widetilde{u}_h)_{0,\Omega}.\tag{48}$$

By Cauchy-Schwarz inequality and the equivalence of the standard L^2 norm $\|\cdot\|_{0,\Omega}$ and the discrete H^1 norm $\|\cdot\|_Z$ on the finite dimensional space $\widetilde{U}^h$, we obtain the following estimate for the right hand side:

$$(f,\widetilde{u}_h)_{0,\Omega}\le\|f\|_{0,\Omega}\|\widetilde{u}_h\|_{0,\Omega}\le K\|f\|_{0,\Omega}\|\widetilde{u}_h\|_Z,\tag{49}$$

where K is the constant from the equivalence of norms. On the other hand, by the adjoint relation (23) and the first inf-sup condition in (25), we have

$$\begin{aligned}\|\widetilde{u}_h\|_Z&\le\frac{1}{\beta_1}\sup_{\boldsymbol{\Psi}\in W^h\backslash\{\mathbf{0}\}}\frac{B_h^*(\widetilde{u}_h,\boldsymbol{\Psi})}{\|\boldsymbol{\Psi}\|_{X'}}\\ &\le\frac{1}{\beta_1}\sup_{\boldsymbol{\Psi}\in W^h\backslash\{\mathbf{0}\}}\frac{B_h^*(\widetilde{u}_h,\boldsymbol{\Psi})}{\|\boldsymbol{\Psi}\|_{0,\Omega}}\\ &=\frac{1}{\beta_1}\sup_{\boldsymbol{\Psi}\in W^h\backslash\{\mathbf{0}\}}\frac{(\widetilde{\mathbf{z}}_h,\boldsymbol{\Psi})_{0,\Omega}}{\|\boldsymbol{\Psi}\|_{0,\Omega}}\\ &=\frac{1}{\beta_1}\|\widetilde{\mathbf{z}}_h\|_{0,\Omega}.\end{aligned}\tag{50}$$

Therefore we have

$$\mu\|\widetilde{u}_h\|^2_Z\le\beta_1^2(f,\widetilde{u}_h)_{0,\Omega}\le\beta_1^2K\|f\|_{0,\Omega}\|\widetilde{u}_h\|_Z.\tag{51}$$

Dividing $\|\widetilde{u}_h\|_Z$ on both sides, we obtain the desired result. □

4. Convergence analysis

In this section, we present an error estimate between the weak solution u in (6) and the ESDG solution $\widetilde{u}_h$ in (22).

Theorem 4.1. *Let $(u, \mathbf{w}, \mathbf{p})$ be the solution of (5)–(6). Let $(\widetilde{u}_h, \widetilde{\mathbf{w}}_h, \widetilde{\mathbf{p}}_h) \in \widetilde{U}^h \times W^h \times W^h$ be the numerical solution of the ESDG method (22). Then we have the following optimal error bound:*

$$\|u - \widetilde{u}_h\|_{L^2(\Omega)} \leq C(1 + \mu^{-1})h^{k+1}, \tag{52}$$

where C is a constant independent of mesh size and diffusivity.

Proof. First, we note that the solution $(u, \mathbf{w}, \mathbf{p})$ satisfies the following system:

$$\begin{aligned} \sqrt{\mu}B_h(\mathbf{w}, v) + \frac{1}{2\sqrt{\mu}}R_h\left(\mathbf{w} + \frac{1}{2\sqrt{\mu}}\mathbf{p}, v\right) &= (f, v)_{0,\Omega}, \\ \sqrt{\mu}B_h^*(u, \mathbf{\Psi}_1) - \frac{1}{2\sqrt{\mu}}(\mathbf{p}, \mathbf{\Psi}_1)_{0,\Omega} &= (\mathbf{w}_h, \mathbf{\Psi}_1)_{0,\Omega}, \\ R_h^*(u, \mathbf{\Psi}_2) &= (\mathbf{p}, \mathbf{\Psi}_2)_{0,\Omega}. \end{aligned} \tag{53}$$

for any $v \in \widetilde{U}^h, \mathbf{\Psi}_1, \mathbf{\Psi}_2 \in W^h$. Subtracting (22) from (53), we have

(54)

$$\begin{aligned} \sqrt{\mu}B_h(\mathbf{w} - \widetilde{\mathbf{w}}_h, v) + \frac{1}{2\sqrt{\mu}}R_h\left((\mathbf{w} - \widetilde{\mathbf{w}}_h) + \frac{1}{2\sqrt{\mu}}(\mathbf{p} - \widetilde{\mathbf{p}}_h), v\right) &= 0, \\ \sqrt{\mu}B_h^*(u - \widetilde{u}_h, \mathbf{\Psi}_1) - \frac{1}{2\sqrt{\mu}}(\mathbf{p} - \widetilde{\mathbf{p}}_h, \mathbf{\Psi}_1)_{0,\Omega} &= (\mathbf{w} - \widetilde{\mathbf{w}}_h, \mathbf{\Psi}_1)_{0,\Omega} \\ R_h^*(u - u_h, \mathbf{\Psi}_2) &= (\mathbf{p} - \widetilde{\mathbf{p}}_h, \mathbf{\Psi}_2)_{0,\Omega}. \end{aligned}$$

Introduce the notations

$$\begin{aligned} \delta_u &= \mathcal{I}u - \widetilde{u}_h \in \widetilde{U}^h, & \varepsilon_u &= u - \mathcal{I}u, \\ \delta_w &= \mathcal{J}\mathbf{w} - \widetilde{\mathbf{w}}_h \in W^h, & \varepsilon_w &= \mathbf{w} - \mathcal{J}\mathbf{w}, \\ \delta_p &= \mathcal{J}\mathbf{p} - \widetilde{\mathbf{p}}_h \in W^h, & \varepsilon_p &= \mathbf{p} - \mathcal{J}\mathbf{p}. \end{aligned} \tag{55}$$

Using the properties in (14), we can rewrite (54) as

$$\begin{aligned} \sqrt{\mu}B_h(\delta_w, v) + \frac{1}{2\sqrt{\mu}}R_h\left(\delta_w + \frac{1}{2\sqrt{\mu}}\delta_p, v\right) &= -\frac{1}{2\sqrt{\mu}}R_h\left(\varepsilon_w + \frac{1}{2\sqrt{\mu}}\varepsilon_p, v\right), \\ \sqrt{\mu}B_h^*(\delta_u, \mathbf{\Psi}_1) - \left(\delta_w + \frac{1}{2\sqrt{\mu}}\delta_p, \mathbf{\Psi}_1\right)_{0,\Omega} &= \left(\varepsilon_w + \frac{1}{2\sqrt{\mu}}\varepsilon_p, \mathbf{\Psi}_1\right)_{0,\Omega}, \\ R_h^*(\delta_u, \mathbf{\Psi}_2) - (\delta_p, \mathbf{\Psi}_2)_{0,\Omega} &= -R_h^*(\varepsilon_u, \mathbf{\Psi}_2) + (\varepsilon_p, \mathbf{\Psi}_2)_{0,\Omega}. \end{aligned} \tag{56}$$

Using the argument as in (50), by the second equation in (56), we have

$$\|\delta_u\|_Z \leq \frac{1}{\beta_1\sqrt{\mu}}\left(\left\|\delta_w + \frac{1}{2\sqrt{\mu}}\delta_p\right\|_{0,\Omega} + \left\|\varepsilon_w + \frac{1}{2\sqrt{\mu}}\varepsilon_p\right\|\right). \tag{57}$$

Moreover, using a discrete Poincaré inequality, we have

$$\|\delta_u\|_{0,\Omega} \leq \frac{K}{\beta_1\sqrt{\mu}}\left(\left\|\delta_w + \frac{1}{2\sqrt{\mu}}\delta_p\right\|_{0,\Omega} + \left\|\varepsilon_w + \frac{1}{2\sqrt{\mu}}\varepsilon_p\right\|_{0,\Omega}\right). \tag{58}$$

On the other hand, in (56), we take we take test functions as follows:

$$\begin{aligned} v &= \delta_u, \\ \boldsymbol{\Psi}_1 &= -\delta_w, \\ \boldsymbol{\Psi}_2 &= -\frac{1}{2\sqrt{\mu}}\delta_w - \frac{1}{4\mu}\delta_p. \end{aligned} \tag{59}$$

Then we have

(60)

$$\begin{aligned} \sqrt{\mu}B_h(\delta_w, \delta_u) + \frac{1}{2\sqrt{\mu}}R_h\left(\delta_w + \frac{1}{2\sqrt{\mu}}\delta_p, \delta_u\right) &= -\frac{1}{2\sqrt{\mu}}R_h\left(\varepsilon_w + \frac{1}{2\sqrt{\mu}}\varepsilon_p, \delta_u\right), \\ -\sqrt{\mu}B_h^*(\delta_u, \delta_w) + \left(\delta_w + \frac{1}{2\sqrt{\mu}}\delta_p, \delta_w\right)_{0,\Omega} &= -\left(\varepsilon_w + \frac{1}{2\sqrt{\mu}}\varepsilon_p, \delta_w\right)_{0,\Omega}, \\ -\frac{1}{2\sqrt{\mu}}R_h^*\left(\delta_u, \delta_w + \frac{1}{2\sqrt{\mu}}\delta_p\right) + \frac{1}{2\sqrt{\mu}}\left(\delta_p, \delta_w + \frac{1}{2\sqrt{\mu}}\delta_p\right)_{0,\Omega} & \\ = \frac{1}{2\sqrt{\mu}}R_h^*\left(\varepsilon_u, \delta_w + \frac{1}{2\sqrt{\mu}}\delta_p\right) & \\ - \frac{1}{2\sqrt{\mu}}\left(\varepsilon_p, \delta_w + \frac{1}{2\sqrt{\mu}}\delta_p\right)_{0,\Omega}&. \end{aligned}$$

Summing up the equations in (60) and using the adjoint relations (23) and (26), we have

$$\left\|\delta_w + \frac{1}{2\sqrt{\mu}}\delta_p\right\|_{0,\Omega}^2 = T_1 + T_2 + T_3 + T_4, \tag{61}$$

where

$$\begin{aligned} T_1 &= \frac{1}{2\sqrt{\mu}}R_h^*\left(\varepsilon_u, \delta_w + \frac{1}{2\sqrt{\mu}}\delta_p\right), \\ T_2 &= -\frac{1}{2\sqrt{\mu}}\left(\varepsilon_p, \delta_w + \frac{1}{2\sqrt{\mu}}\delta_p\right)_{0,\Omega}, \\ T_3 &= -\frac{1}{2\sqrt{\mu}}R_h\left(\varepsilon_w + \frac{1}{2\sqrt{\mu}}\varepsilon_p, \delta_u\right), \\ T_4 &= -\left(\varepsilon_w + \frac{1}{2\sqrt{\mu}}\varepsilon_p, \delta_w\right)_{0,\Omega}. \end{aligned} \tag{62}$$

Next, we will estimate each of these terms. Using Young's inequality, we have

$$\begin{aligned} |T_1| &\leq \frac{1}{2\sqrt{\mu}}\|\mathbf{b}\|_{L^\infty(\Omega)}\|\varepsilon_u\|_{0,\Omega}\left\|\delta_w + \frac{1}{2\sqrt{\mu}}\delta_p\right\|_{0,\Omega} \\ &\leq \frac{1}{4\mu}\|\mathbf{b}\|_{L^\infty(\Omega)}^2\|\varepsilon_u\|_{0,\Omega}^2 + \frac{1}{4}\left\|\delta_w + \frac{1}{2\sqrt{\mu}}\delta_p\right\|_{0,\Omega}^2. \end{aligned} \tag{63}$$

Similarly, for T_2, we imply

$$\begin{aligned} |T_2| &\leq \frac{1}{2\sqrt{\mu}}\|\varepsilon_p\|_{0,\Omega}\left\|\delta_w + \frac{1}{2\sqrt{\mu}}\delta_p\right\|_{0,\Omega} \\ &\leq \frac{1}{4\mu}\|\varepsilon_p\|_{0,\Omega}^2 + \frac{1}{4}\left\|\delta_w + \frac{1}{2\sqrt{\mu}}\delta_p\right\|_{0,\Omega}^2. \end{aligned} \tag{64}$$

For T_3, we have

$$\begin{aligned}|T_3| &\le \frac{1}{2\sqrt{\mu}}\|\mathbf{b}\|_{L^\infty(\Omega)}\left\|\varepsilon_w+\frac{1}{2\sqrt{\mu}}\varepsilon_p\right\|_{0,\Omega}\|\delta_u\|_{0,\Omega},\\ &\le \frac{2K^2}{\beta_1^2\mu^2}\|\mathbf{b}\|^2_{L^\infty(\Omega)}\left\|\varepsilon_w+\frac{1}{2\sqrt{\mu}}\varepsilon_p\right\|^2_{0,\Omega}+\frac{\beta_1^2\mu}{32K^2}\|\delta_u\|^2_{0,\Omega}.\end{aligned}\tag{65}$$

For T_4, we first observe that

$$\begin{aligned}|T_4| &\le \left\|\varepsilon_w+\frac{1}{2\sqrt{\mu}}\varepsilon_p\right\|_{0,\Omega}\|\delta_w\|_{0,\Omega}\\ &\le \left\|\varepsilon_w+\frac{1}{2\sqrt{\mu}}\varepsilon_p\right\|_{0,\Omega}\left(\left\|\delta_w+\frac{1}{2\sqrt{\mu}}\delta_p\right\|_{0,\Omega}+\frac{1}{2\sqrt{\mu}}\|\delta_p\|_{0,\Omega}\right).\end{aligned}\tag{66}$$

Taking $\mathbf{\Psi}_2=-\delta_p$ in the last equation of (56), we have

$$\begin{aligned}\|\delta_p\|^2_{0,\Omega} &= R_h^*(\delta_u,\delta_p)+R_h^*(\varepsilon_u,\delta_p)-(\varepsilon_p,\delta_p)_{0,\Omega}\\ &\le \|\mathbf{b}\|_{L^\infty(\Omega)}(\|\delta_u\|_{0,\Omega}+\|\varepsilon_u\|_{0,\Omega})\|\delta_p\|_{0,\Omega}+\|\varepsilon_p\|_{0,\Omega}\|\delta_p\|_{0,\Omega}\end{aligned}\tag{67}$$

Hence we imply

$$\begin{aligned}|T_4| &\le \left\|\varepsilon_w+\frac{1}{2\sqrt{\mu}}\varepsilon_p\right\|_{0,\Omega}\|\delta_w\|_{0,\Omega}\\ &\le \left\|\varepsilon_w+\frac{1}{2\sqrt{\mu}}\varepsilon_p\right\|_{0,\Omega}\left(\left\|\delta_w+\frac{1}{2\sqrt{\mu}}\delta_p\right\|_{0,\Omega}\right.\\ &\quad\left.+\frac{1}{2\sqrt{\mu}}\left(\|\mathbf{b}\|_{L^\infty(\Omega)}(\|\delta_u\|_{0,\Omega}+\|\varepsilon_u\|_{0,\Omega})+\|\varepsilon_p\|_{0,\Omega}\right)\right)\\ &\le \left\|\varepsilon_w+\frac{1}{2\sqrt{\mu}}\varepsilon_p\right\|^2_{0,\Omega}+\frac{1}{4}\left\|\delta_w+\frac{1}{2\sqrt{\mu}}\delta_p\right\|^2_{0,\Omega}\\ &\quad+\frac{2K^2}{\beta_1^2\mu^2}\|\mathbf{b}\|^2_{L^\infty(\Omega)}\left\|\varepsilon_w+\frac{1}{2\sqrt{\mu}}\varepsilon_p\right\|^2_{0,\Omega}+\frac{\beta_1^2\mu}{32K^2}\|\delta_u\|^2_{0,\Omega}\\ &\quad+\left\|\varepsilon_w+\frac{1}{2\sqrt{\mu}}\varepsilon_p\right\|^2_{0,\Omega}+\frac{1}{16\mu}\left(\|\mathbf{b}\|^2_{L^\infty(\Omega)}\|\varepsilon_u\|^2_{0,\Omega}+\|\varepsilon_p\|^2_{0,\Omega}\right)\end{aligned}\tag{68}$$

Combining all these estimates with (61), we have

$$\begin{aligned}\frac{1}{4}\left\|\delta_w+\frac{1}{2\sqrt{\mu}}\delta_p\right\|^2_{0,\Omega} \le& C\left(\mu^{-1}\|\varepsilon_u\|^2_{0,\Omega}+\mu^{-1}\|\varepsilon_p\|^2_{0,\Omega}+(1+\mu^{-2})\left\|\varepsilon_w+\frac{1}{2\sqrt{\mu}}\varepsilon_p\right\|^2_{0,\Omega}\right)\\ &+\frac{\beta_1^2\mu}{16K^2}\|\delta_u\|^2_{0,\Omega}.\end{aligned}\tag{69}$$

Combining (58) and (69), we have

(70)

$$\begin{aligned}\|\delta_u\|_{0,\Omega}^2 \leq & \frac{K^2}{\beta_1^2\mu}\left(\left\|\delta_w + \frac{1}{2\sqrt{\mu}}\delta_p\right\|_{0,\Omega} + \left\|\varepsilon_w + \frac{1}{2\sqrt{\mu}}\varepsilon_p\right\|_{0,\Omega}\right)^2 \\ \leq & \frac{2K^2}{\beta_1^2\mu}\left(\left\|\delta_w + \frac{1}{2\sqrt{\mu}}\delta_p\right\|_{0,\Omega}^2 + \left\|\varepsilon_w + \frac{1}{2\sqrt{\mu}}\varepsilon_p\right\|_{0,\Omega}^2\right) \\ \leq & \frac{2K^2}{\beta_1^2\mu}\left(C\left(\mu^{-1}\|\varepsilon_u\|_{0,\Omega}^2 + \mu^{-1}\|\varepsilon_p\|_{0,\Omega}^2 + (1+\mu^{-2})\left\|\varepsilon_w + \frac{1}{2\sqrt{\mu}}\varepsilon_p\right\|_{0,\Omega}^2\right)\right. \\ & \left. + \frac{\beta_1^2\mu}{4K^2}\|\delta_u\|_{0,\Omega}^2\right).\end{aligned}$$

Therefore, we have

$$\|\delta_u\|_{0,\Omega}^2 \leq \frac{C}{\mu}\left(\mu^{-1}\|\varepsilon_u\|_{0,\Omega}^2 + \mu^{-1}\|\varepsilon_p\|_{0,\Omega}^2 + (1+\mu^{-2})\left\|\varepsilon_w + \frac{1}{2\sqrt{\mu}}\varepsilon_p\right\|_{0,\Omega}^2\right). \quad (71)$$

Finally, using the approximation properties (15), we imply

$$\|\delta_u\|_{0,\Omega}^2 \leq C(1+\mu^{-2})h^{2(k+1)}. \quad (72)$$

Using triangle inequality on $u - \widetilde{u}_h = \varepsilon_u + \delta_u$, we obtain our desired result. □

We remark that the error $u - \widetilde{u}_h$ consists of two parts. The difference δ_u between the numerical solution and the interpolation image depends on the viscosity coefficient μ, while the interpolation error ε_u does not. As we will see in our numerical results, the error does not vary significantly with the viscosity coefficient μ.

5. Numerical results

In this section, we illustrate some numerical examples. We carry out numerical experiments to see and compare the rates of convergence of the SDG method and the ESDG method. Polynomials with degree $k = 1$ is used for SDG approximations. We are interested in the L^2 error of the unknown function u and also that of the flux $\mathbf{z} = \nabla u$. We recall the definitions in (40) and (41) for numerical approximations of the flux.

Throughout this section, we will take $\Omega = [0,1]^2 \subset \mathbb{R}^2$ and use a family of staggered meshes in all the experiments. We denote the number of uniform divisions in $[0,1]$ in the mesh by N. The domain Ω is partitioned into N^2 sub-squares with length $h = N^{-1}$. Each sub-square is then divided into two identical right-angled triangles by its diagonal. This constructs the initial triangulation $\mathcal{T}_u$. To construct the staggered mesh $\mathcal{T}$, we simply take the interior point ν as the centroid in each initial triangle $\mathcal{S}(\nu) \in \mathcal{T}_u$. Figure 2 illustrates a member of this family with $N = 4$. Table 1 compares the numbers of degrees of freedom in the discrete problems (32) for the SDG method and (35) for the ESDG method for this family of mesh. It can be seen that the number of degrees of freedom for the ESDG method is almost half of that of the SDG method for each level of mesh. Indeed, it can be worked out that for this family of mesh, the ratio of the numbers of the degrees of freedom is 7/12 asymptotically:

$$\begin{aligned}\dim(U^h) &= 12N^2 + 4N, \\ \dim(\widetilde{U}^h) &= 7N^2 + 2N + 1.\end{aligned} \quad (73)$$

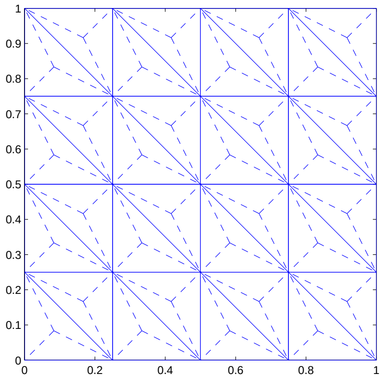

FIGURE 2. An illustration of a staggered mesh on $\Omega = [0,1]^2$ with $N = 4$.

TABLE 1. Comparison of numbers of degrees of freedom: $\dim(U^h)$ for SDG and $\dim(\widetilde{U}^h)$ for ESDG.

N	$\dim(U^h)$	$\dim(\widetilde{U}^h)$
2	56	33
4	208	121
8	800	465
16	3136	1825
32	12416	7233
64	49408	28801

5.1. Experiment 1: comparison to the EDG method. The purpose of this experiment is to compare our method with [38] in the same setting. In this experiment, the convection field $\mathbf{b} = (b_1, b_2)$ is a constant vector. The analytic solution of this experiment is given by

$$u(x,y) = xy\frac{(1-e^{b_1(x-1)})(1-e^{b_2(y-1)})}{(1-e^{b_1})(1-e^{b_2})}. \tag{74}$$

For large values of b_1 and b_2, there is a boundary layer around the segments $x = 1$ and $y = 1$. The diffusivity μ are set to be 1. The constant convection field is chosen to be $\mathbf{b} = (20, 20)$ and the problem is weakly convection-dominated. A homogeneous Dirichlet boundary condition $u|_{\partial\Omega} = 0$ is prescribed. The source function f is computed accordingly. The SDG method (19) and the ESDG method (22) are used to solve the problem numerically. We examine the performance of the ESDG method by comparing the L^2 errors and the orders of L^2 convergence to the EDG method and the SDG method.

Figure 3 shows a plot of the numerical solution $\widetilde{u}_h$ of the ESDG method. Tables 2 compare the convergence results of the SDG method and the ESDG method with various scales of diffusivity μ. The second to the firth columns record the L^2 error and the orders of convergence of the potential and the flux for the SDG method. The sixth to the ninth columns record the L^2 error and the orders of convergence of the potential and the flux for the ESDG method. It can be seen that the approximated potential converge with an optimal order 2 in L^2 error for both the SDG method and the ESDG method, which is the same for the EDG method in [38]. In particular, for $N = 32$ and $N = 64$, the ESDG method gives a L^2 error smaller that the SDG method and also the EDG method. However, similar to the EDG method for second-order elliptic problems [26], our numerical results show that the ESDG method only provides a suboptimal order 1 of convergence of L^2 error of the approximated flux, while the SDG method provides an optimal order 2.

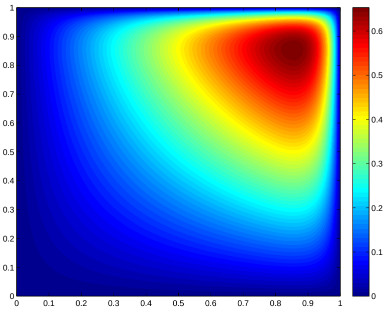

FIGURE 3. A plot for the numerical solution $\widetilde{u}_h$ in Experiment 1.

TABLE 2. History of convergence in Experiment 1.

Mesh	$\|u - u_h\|_{0,\Omega}$		$\|\mathbf{z} - \mathbf{z}_h\|_{0,\Omega}$		$\|u - \widetilde{u}_h\|_{0,\Omega}$		$\|\mathbf{z} - \widetilde{\mathbf{z}}_h\|_{0,\Omega}$	
N	Error	Order	Error	Order	Error	Order	Error	Order
2	1.50e-01	–	2.48e+00	–	1.03e-01	–	2.31e+00	–
4	8.42e-02	0.84	1.63e+00	0.60	5.80e-02	0.82	2.00e+00	0.21
8	3.37e-02	1.32	6.96e-01	1.23	2.23e-02	1.38	1.38e+00	0.54
16	1.03e-02	1.72	2.15e-01	1.70	6.15e-03	1.86	7.92e-01	0.80
32	2.72e-03	1.91	5.71e-02	1.91	1.55e-03	1.99	4.13e-01	0.94
64	6.91e-04	1.98	1.45e-02	1.98	3.86e-04	2.00	2.09e-01	0.98

5.2. Experiment 2: sensitivity of orders of convergence to diffusivity.

The purpose of this experiment is to examine the performance of the ESDG method in terms of L^2 convergence and compare the ESDG method with the SDG method in various scales of diffusivity. In this experiment, the convection field $\mathbf{b} = (b_1, b_2)$ is set to be

$$\begin{aligned} b_1 &= (1 - \cos(2\pi x))\sin(2\pi y), \\ b_2 &= -\sin(2\pi x)(1 - \cos(2\pi y)). \end{aligned} \tag{75}$$

The analytic solution of this experiment is given by

$$u = \sin(2\pi x)\cos(2\pi y). \tag{76}$$

Figure 4 shows a plot of the convection field $\mathbf{b}$ in this experiment. We perform the experiment with different scales of diffusivity μ. In particular, we are interested in observing the behaviour of the solutions when μ is small, i.e. the problem is convection-dominated. An inhomogeneous Dirichlet boundary condition is prescribed. The source function f is computed accordingly. The SDG method (19) and the ESDG method (22) are used to solve the problem numerically. We examine the performance of the ESDG method by comparing the L^2 errors and the orders of L^2 convergence to the SDG method.

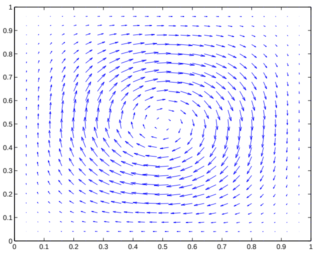

FIGURE 4. The convection field $\mathbf{b}$ in Experiment 2.

Tables 3–9 compare the convergence results of the SDG method and the ESDG method with various scales of diffusivity μ. The second to the firth columns record the L^2 error and the orders of convergence of the potential and the flux for the SDG method. The sixth to the ninth columns record the L^2 error and the orders of convergence of the potential and the flux for the ESDG method. It can be seen that when the diffusivity μ is close to unity, the SDG method clearly outperforms the ESDG method. The convergence of the potential is optimal for both methods, while the convergence of the flux is optimal for the SDG method and suboptimal for the ESDG method. However, when the diffusivity μ reduces in scale, the optimal convergence of the flux for the SDG method is lost. By comparing the second column with the fourth column, and comparing the third column with the column, it can be seen that for convection-dominated situations, say $\mu \leq 10^{-3}$ in Tables 6–9, the ESDG method has a comparable performance to the SDG method. With the considerable reduction in the size of the discrete problem, these results suggest that the ESDG method is favourable in convection-dominated situations. These observations in moderate problems and convection-dominated problems are in good agreement with the descriptions in [38]. Furthermore, we observe that, with a fixed mesh size, the L^2 error of the potential does not vary significantly with the viscosity coefficient μ.

TABLE 3. History of convergence for $\mu = 10^0$ in Experiment 2.

Mesh	$\|u - u_h\|_{0,\Omega}$		$\|\mathbf{z} - \mathbf{z}_h\|_{0,\Omega}$		$\|u - \tilde{u}_h\|_{0,\Omega}$		$\|\mathbf{z} - \tilde{\mathbf{z}}_h\|_{0,\Omega}$	
N	Error	Order	Error	Order	Error	Order	Error	Order
2	3.28e-01	–	2.02e+00	–	6.37e-02	–	3.11e+00	–
4	1.23e-01	1.42	1.19e+00	0.76	1.05e-01	-0.72	2.41e+00	0.37
8	3.58e-02	1.78	3.44e-01	1.79	3.81e-02	1.46	1.47e+00	0.71
16	9.31e-03	1.94	8.95e-02	1.94	1.06e-02	1.85	7.80e-01	0.92
32	2.35e-03	1.99	2.26e-02	1.99	2.72e-03	1.96	3.96e-01	0.98
64	5.89e-04	2.00	5.67e-03	2.00	6.85e-04	1.99	1.99e-01	0.99

TABLE 4. History of convergence for $\mu = 10^{-2}$ in Experiment 2.

Mesh	$\|u - u_h\|_{0,\Omega}$		$\|\mathbf{z} - \mathbf{z}_h\|_{0,\Omega}$		$\|u - \tilde{u}_h\|_{0,\Omega}$		$\|\mathbf{z} - \tilde{\mathbf{z}}_h\|_{0,\Omega}$	
N	Error	Order	Error	Order	Error	Order	Error	Order
2	1.07e+00	–	1.19e+01	–	6.40e-01	–	9.39e+00	–
4	2.16e-01	2.31	3.89e+00	1.61	2.62e-01	1.29	6.11e+00	0.62
8	5.58e-02	1.95	1.49e+00	1.38	5.55e-02	2.24	2.60e+00	1.23
16	1.39e-02	2.01	4.96e-01	1.59	1.24e-02	2.16	1.02e+00	1.35
32	3.38e-03	2.03	1.43e-01	1.79	2.88e-03	2.11	4.35e-01	1.23
64	8.33e-04	2.02	3.79e-02	1.92	7.01e-04	2.04	2.04e-01	1.09

Table 5. History of convergence for $\mu = 2 \times 10^{-3}$ in Experiment 2.

Mesh	$\|u - u_h\|_{0,\Omega}$		$\|\mathbf{z} - \mathbf{z}_h\|_{0,\Omega}$		$\|u - \widetilde{u}_h\|_{0,\Omega}$		$\|\mathbf{z} - \widetilde{\mathbf{z}}_h\|_{0,\Omega}$	
N	Error	Order	Error	Order	Error	Order	Error	Order
2	1.91e+00	–	3.03e+01	–	1.35e+00	–	2.16e+01	–
4	3.11e-01	2.62	8.52e+00	1.83	4.37e-01	1.63	1.36e+01	0.66
8	8.17e-02	1.93	3.93e+00	1.12	8.27e-02	2.40	4.91e+00	1.47
16	2.07e-02	1.98	1.57e+00	1.32	1.82e-02	2.18	1.74e+00	1.50
32	4.78e-03	2.11	5.35e-01	1.55	4.15e-03	2.13	6.71e-01	1.37
64	1.09e-03	2.13	1.58e-01	1.76	9.71e-04	2.10	2.67e-01	1.33

Table 6. History of convergence for $\mu = 10^{-3}$ in Experiment 2.

Mesh	$\|u - u_h\|_{0,\Omega}$		$\|\mathbf{z} - \mathbf{z}_h\|_{0,\Omega}$		$\|u - \widetilde{u}_h\|_{0,\Omega}$		$\|\mathbf{z} - \widetilde{\mathbf{z}}_h\|_{0,\Omega}$	
N	Error	Order	Error	Order	Error	Order	Error	Order
2	2.33e+00	–	4.23e+01	–	2.42e+00	–	3.98e+01	–
4	3.61e-01	2.69	1.10e+01	1.95	6.06e-01	2.00	2.09e+01	0.93
8	9.92e-02	1.86	5.66e+00	0.95	1.09e-01	2.47	7.20e+00	1.54
16	2.58e-02	1.94	2.37e+00	1.26	2.30e-02	2.25	2.41e+00	1.58
32	5.99e-03	2.11	8.87e-01	1.42	4.92e-03	2.23	8.58e-01	1.49
64	1.31e-03	2.20	2.77e-01	1.68	1.14e-03	2.11	3.34e-01	1.36

Table 7. History of convergence for $\mu = 5 \times 10^{-4}$ in Experiment 2.

Mesh	$\|u - u_h\|_{0,\Omega}$		$\|\mathbf{z} - \mathbf{z}_h\|_{0,\Omega}$		$\|u - \widetilde{u}_h\|_{0,\Omega}$		$\|\mathbf{z} - \widetilde{\mathbf{z}}_h\|_{0,\Omega}$	
N	Error	Order	Error	Order	Error	Order	Error	Order
2	2.78e+00	–	5.53e+01	–	4.65e+00	–	7.80e+01	–
4	4.34e-01	2.68	1.39e+01	2.00	9.42e-01	2.31	3.34e+01	1.22
8	1.22e-01	1.83	7.92e+00	0.81	1.67e-01	2.50	1.15e+01	1.53
16	3.20e-02	1.93	3.37e+00	1.23	3.13e-02	2.41	3.77e+00	1.61
32	7.69e-03	2.06	1.39e+00	1.28	6.15e-03	2.35	1.20e+00	1.65
64	1.66e-03	2.21	4.72e-01	1.56	1.33e-03	2.20	4.27e-01	1.49

Table 8. History of convergence for $\mu = 2 \times 10^{-4}$ in Experiment 2.

Mesh	$\|u - u_h\|_{0,\Omega}$		$\|\mathbf{z} - \mathbf{z}_h\|_{0,\Omega}$		$\|u - \widetilde{u}_h\|_{0,\Omega}$		$\|\mathbf{z} - \widetilde{\mathbf{z}}_h\|_{0,\Omega}$	
N	Error	Order	Error	Order	Error	Order	Error	Order
2	3.78e+00	–	8.05e+01	–	1.14e+01	–	1.94e+02	–
4	6.29e-01	2.59	2.08e+01	1.96	2.05e+00	2.48	7.21e+01	1.43
8	1.66e-01	1.93	1.20e+01	0.79	3.41e-01	2.59	2.40e+01	1.59
16	4.28e-02	1.95	5.03e+00	1.26	5.65e-02	2.59	7.94e+00	1.60
32	1.07e-02	2.00	2.26e+00	1.15	9.58e-03	2.56	2.28e+00	1.80
64	2.39e-03	2.16	8.86e-01	1.35	1.79e-03	2.42	6.88e-01	1.73

Table 9. History of convergence for $\mu = 10^{-4}$ in Experiment 2.

Mesh	$\|u - u_h\|_{0,\Omega}$		$\|\mathbf{z} - \mathbf{z}_h\|_{0,\Omega}$		$\|u - \widetilde{u}_h\|_{0,\Omega}$		$\|\mathbf{z} - \widetilde{\mathbf{z}}_h\|_{0,\Omega}$	
N	Error	Order	Error	Order	Error	Order	Error	Order
2	5.33e+00	–	1.20e+02	–	2.28e+01	–	3.87e+02	–
4	8.87e-01	2.59	3.06e+01	1.98	3.91e+00	2.54	1.38e+02	1.49
8	2.21e-01	2.01	1.64e+01	0.90	6.20e-01	2.66	4.37e+01	1.66
16	5.43e-02	2.02	6.75e+00	1.28	1.03e-01	2.60	1.50e+01	1.54
32	1.39e-02	1.97	3.10e+00	1.12	1.52e-02	2.76	4.10e+00	1.87
64	3.17e-03	2.13	1.32e+00	1.23	2.52e-03	2.59	1.15e+00	1.84

5.3. Experiment 3: uniform stability in L^2 energy with respect to diffusivity. The purpose of this experiment is to examine the stability in L^2 energy of the ESDG method in various scales of diffusivity. We first observe that it is actually possible to derive different discretizations for the convection term and the diffusion term, and we will compare our skew-symmetric discretization with two types of non-skew-symmetric discretizations. Given $\theta \in [0, 1]$. we modify the definitions of the auxiliary variables in (5) by

$$\begin{aligned} \mathbf{w} &= \sqrt{\mu}\nabla u - \frac{\theta}{\sqrt{\mu}}\mathbf{b}u, \\ \mathbf{z} &= \mathbf{b}u. \end{aligned} \tag{77}$$

Then we can use the same idea as (22) to obtain a new method. The discrete convection term is then modified accordingly as

$$\mathbf{b} \cdot \nabla_h = -\theta \widetilde{B} M^{-1} \widetilde{R}^T + (1-\theta) \widetilde{R} M^{-1} \widetilde{B}^T. \tag{78}$$

In particular, when $\theta = 1/2$, it is reduced to ESDG method (22) with a skew-symmetric discretization of the convection term proposed in Section 2. We will compare the discretizations with $\theta = 0$, $\theta = 1/2$ and $\theta = 1$, and observe the advantages brought by the spectro-consistent discretization with the novel splitting of the convection term and the diffusion term. We remark that a similar experiment is performed on the SDG method for incompressible Navier-Stokes equations in [7].

In this experiment, the convection field $\mathbf{b} = (b_1, b_2)$ is identical to Experiment 2.

$$\begin{aligned} b_1 &= (1 - \cos(2\pi x)) \sin(2\pi y), \\ b_2 &= -\sin(2\pi x)(1 - \cos(2\pi y)). \end{aligned} \tag{79}$$

The analytic solution of this experiment is given by

$$u = \sin(2\pi x) \sin(2\pi y). \tag{80}$$

We perform the experiment with different scales of diffusivity μ. In particular, we are interested in observing the behaviour of the solutions when μ is small, i.e. the problem is convection-dominated. A homogeneous Dirichlet boundary condition $u|_{\partial\Omega} = 0$ is prescribed. The source function f is computed accordingly. The ESDG method (22) is used to solve the problem numerically. We use a mesh with size $N = 32$. We are interested in the L^2 norm $\|\widetilde{\mathbf{z}}_h\|_{0,\Omega}$ of the approximation $\widetilde{\mathbf{z}}_h$ of the flux $\mathbf{z}$. By a direct computation, it is easy to see that $\|\mathbf{z}\|_{0,\Omega} = \sqrt{2}\pi \approx 4.4429$.

Tables 10 records the L^2 norm $\|\widetilde{\mathbf{z}}_h\|_{0,\Omega}$ of the approximation $\widetilde{\mathbf{z}}_h$ with the three different discretizations. For more moderate problems $\mu > 10^{-3}$, it can be seen that all the three discretizations provide a approximation $\widetilde{\mathbf{z}}_h$ with the L^2 norm close to the value $\sqrt{2}\pi$. However, for convection dominated problems, the skew symmetric discretization $\theta = 1/2$ clearly outperforms the other two discretizations. In spite of the machine error due to an ill-conditioned linear system as the diffusivity tends to zero, the L^2 norm $\|\widetilde{\mathbf{z}}_h\|_{0,\Omega}$ of the approximation $\widetilde{\mathbf{z}}_h$ is around a constant when $\theta = 1/2$. Meanwhile, for the other two discretizations, the L^2 norm $\|\widetilde{\mathbf{z}}_h\|_{0,\Omega}$ of the approximation $\widetilde{\mathbf{z}}_h$ blows up as the diffusivity tends to zero.

TABLE 10. Record of $\|\mathbf{z}_h\|_{0,\Omega}$ in Experiment 3.

Diffusivity	$\|z_h\|_{0,\Omega}$		
μ	$\theta = 0$	$\theta = 1/2$	$\theta = 1$
10^0	4.43e+00	4.43e+00	4.43e+00
10^{-2}	4.47e+00	4.47e+00	4.47e+00
2×10^{-3}	4.48e+00	4.49e+00	4.59e+00
10^{-3}	6.33e+00	4.52e+00	9.31e+00
5×10^{-4}	1.12e+03	4.59e+00	2.01e+03
2×10^{-4}	1.81e+03	4.88e+00	8.73e+02
10^{-4}	1.55e+05	5.52e+00	8.51e+04

6. Conclusion

In this paper, we develop an embedded staggered discontinuous Galerkin method for the convection-diffusion equation. Thanks to the design of the SDG finite element spaces, the new method provides local and global conservations, and does not require the introduction of carefully designed stabilization terms or flux conditions. Furthermore, L^2 stability is achieved by a skew-symmetric discretization of the convection term. Numerical results are presented to show the robustness of the method with respect to diffusivity. On the other hand, the method seeks reduced approximations in a subspace of the SDG finite element space. In convection-dominated problems, like other DG methods, the convergence are optimal in potential and suboptimal in flux, as our numerical results have shown.

References

[1] D. N. Arnold, F. Brezzi, B. Cockburn, and L. D. Marini, Unified analysis of discontinuous Galerkin methods for elliptic problems, SIAM J. Numer. Anal., 39 (2002), pp. 1749–1779.

[2] P. Blanc, R. Eymard, R. Herbin, A staggered finite volume scheme on general meshes for the generalized Stokes problem in two space dimensions, International Journal on Finite Volumes, 2 (2005), pp. 1–31.

[3] B. J. Boersma, A staggered compact finite difference formulation for the compressible Navier-Stokes equations, J. Comput. Phys., 208 (2005), pp. 675–690.

[4] D. Braess, Finite elements. Theory, fast solvers, and applications in elasticity theory, Cambridge University Press, Cambridge, 2007.

[5] A. Buffa, T. J. R. Hughes, G. Sangalli, Analysis of a multiscale discontinuous Galerkin method for convection-diffusion problems, SIAM J. Numer. Anal., 44 (2006), pp. 1420–1440.

[6] J. Carrero, B. Cockburn, D. Schötzau, Hybridized globally divergence-free LDG methods. Part I: The Stokes problem, Math. Comput., 75 (2005), pp. 533–563.

[7] S. W. Cheung, E. Chung, H. H. Kim, Y. Qian, Staggered discontinuous Galerkin methods for incompressible Navier-Stokes equations, J. Comput. Phys., 302 (2015), pp. 251–266.

[8] S. W. Cheung, E. Chung, H. H. Kim, Staggered discontinuous Galerkin approximation for immersed boundary method, arXiv preprint, arXiv:1609.01046, 2016.

[9] E. T. Chung, Q. Du, J. Zou, Convergence analysis on a finite volume method for Maxwell's equations in non-homogeneous media, SIAM J. Numer. Anal., 41 (2003), pp. 37–63.

[10] E. T. Chung, B. Engquist, Convergence analysis of fully discrete finite volume methods for Maxwell's equations in nonhomogeneous media, SIAM J. Numer. Anal., 43 (2005), pp. 303–317.

[11] E. T. Chung, B. Engquist, Optimal discontinuous Galerkin methods for wave propagation, SIAM J. Numer. Anal., 44 (2006), pp. 2131–2158.

[12] E. T. Chung, B. Engquist, Optimal discontinuous Galerkin methods for the acoustic wave equation in higher dimensions, SIAM J. Numer. Anal., 47 (2009), pp. 3820–3848.

[13] E. T. Chung, P. Ciarlet, A staggered discontinuous Galerkin method for wave propagation in media with dielectrics and meta-materials, J. Comput. Appl. Math., 239 (2013), pp. 189–207.

[14] E. T. Chung, P. Ciarlet, T. F. Yu, Convergence and superconvergence of staggered discontinuous Galerkin methods for the three-dimensional Maxwell's equations on Cartesian grids, J. Comput. Phys., 235 (2013), pp. 14–31.

[15] E. Chung, B. Cockburn, G. Fu, The staggered DG method is the limit of a hybridizable DG method, SIAM J. Numer. Anal., 52 (2014), pp. 915–932.

[16] E. Chung, B. Cockburn, G. Fu, The staggered DG method is the limit of a hybridizable DG method. Part II: The Stokes flow., J. Sci. Comput., 66 (2016), pp. 870-887.

[17] E. T. Chung, C. Y. Lam and J. Qian, A staggered discontinuous Galerkin method for the simulation of seismic waves with surface topography, Geophysics, 80 (2015), pp. T119-T135.

[18] E. T. Chung, C. S. Lee, A staggered discontinuous Galerkin method for the curl-curl operator, IMA J. Numer. Anal., 32 (2012), pp. 1241–1265.

[19] E. T. Chung, C. S. Lee, A staggered discontinuous Galerkin method for the convection-diffusion equation, J. Numer. Math., 20 (2012), pp. 1–31.

[20] E. T. Chung, W. Qiu, Analysis of a SDG method for the incompressible Navier-Stokes equations, Submitted.

[21] P. Ciarlet, The Finite Element Method for Elliptic Problems, North-Holland, Amsterdam.

[22] B. Cockburn, J. Gopalakrishnan, N.C. Nguyen, J. Peraire, F.-J. Sayas, Analysis of an HDG method for Stokes flow, Math. Comput., 80 (2011), pp. 723–760.

[23] B. Cockburn, J. Gopalakrishnan, R. Lazarov, Unified hybridization of discontinuous Galerkin, mixed and continuous Galerkin methods for second order elliptic problems, SIAM J. Numer. Anal. 47 (2009) pp. 1319–1365.

[24] B. Cockburn, G. Kanschat, D. Schötzau, A locally conservative LDG method for the incompressible Navier-Stokes equations, Math. Comp., 74 (2005), pp. 1067–1095.

[25] B. Cockburn, B. Dong, J. Guzman, M. Restelli, R. Sacco, A hybridizable discontinuous Galerkin method for steady-state convection-diffusion-reaction problems, SIAM J. Sci. Comput., 31 (2009), 3827–3846.

[26] B. Cockburn, J. Guzman, S.-C. Soon, H.K. Stolarski, An analysis of the embedded discontinuous Galerkin method for second-order elliptic problems, SIAM J. Numer. Anal., 47 (2009) pp. 2686–2707.

[27] B. Cockburn, G. Kanschat, D. Schötzau, C. Schwab, Local discontinuous Galerkin methods for the Stokes system, SIAM J. Numer. Anal., 40 (2002), pp. 319–343.

[28] B. Cockburn, C.-W. Shu, The local discontinuous Galerkin method for time-dependent convection-diffusion systems, SIAM J. Numer. Anal., 35 (1998), 2440–2463.

[29] P. Fernandez, N. C. Nguyen, X. Roca, J. Peraire, Implicit large-eddy simulation of compressible flows using the Interior Embedded Discontinuous Galerkin method, 54th AIAA Aerospace Sciences Meeting, AIAA SciTech, (AIAA 2016-1332)

[30] S. Güzey, B. Cockburn, and H.K. Stolarski, The embedded discontinuous Galerkin methods: Application to linear shells problems, Internat. J. Numer. Methods Engrg., 70 (2007), 757–790.

[31] F. H. Harlow, J. E. Welch, Numerical calculation of time-dependent viscous incompressible flow of fluid with a free surface, Phys. Fluids, 8 (1965), pp. 2182–2189.

[32] P. Houston, D. Schötzau, X. Wei, A mixed DG method for linearized incompressible magnetohydrodynamics, J. Sci. Comp., 40 (2009), pp. 281–314.

[33] J. T. R. Hughes, G. Scovazzi, P. B. Bochev, A. Buffa, A multiscale discontinuous Galerkin method with the computational structure of a continuous Galerkin method, Comput. Methods Appl. Mech. Engrg., 195 (2006), pp. 2761–2787.

[34] H. H. Kim, E. T. Chung, C. S. Lee, A staggered discontinuous Galerkin method for the Stokes system, SIAM J. Numer. Anal., 51 (2013), pp. 3327–3350.

[35] H. H. Kim, E. T. Chung, C. S. Lee, FETI-DP preconditioners for a staggered discontinuous Galerkin formulation of the two-dimensional Stokes problem, Comput. & Math. Appl., 68 (2014), pp. 2233-2250.

[36] J.-G. Liu, C.-W. Shu, A high-order discontinuous Galerkin method for 2D incompressible flows, J. Comput. Phys., 160 (2000), pp. 577–596.

[37] N. C. Nguyen, J. Peraire, B. Cockburn, An implicit high-order hybridizable discontinuous Galerkin method for the incompressible Navier-Stokes equations, J. Comput. Phys., 230 (2011), pp. 1147–1170.

[38] N.C. Nguyen, J. Peraire, B. Cockburn, A class of embedded discontinuous Galerkin methods for computational fluid dynamics, J. Comput. Phys., 302 (2015), pp. 674–692.

[39] J. Peraire, N.C. Nguyen, B. Cockburn, An embedded discontinuous Galerkin method for the compressible Euler and NavierStokes equations , (AIAA Paper 2011-3228), in: Proceedings of the 20th AIAA Computational Fluid Dynamics Conference, Honolulu, HI, 2011.

[40] W. H. Reed, T. R. Hill, Triangular Mesh Methods for the Neutron Transport Equation, Tech. Report LA-UR-73-479, Los Alamos Scientific Laboratory, Los Alamos, NM, 1973.

[41] D. Schötzau, C. Schwab, A. Toselli, Mixed hp-DGFEM for incompressible flows, SIAM J. Numer. Anal., 40 (2003), pp. 2171–2194.

[42] K. Shahbazi, P. F Fischer, C.R. Ethier, A high-order discontinuous Galerkin method for the unsteady incompressible Navier-Stokes equations, J. Comput. Phys. 222 (2007), pp. 391–407.

Department of Mathematics, Texas A&M University, USA.

Department of Mathematics, The Chinese University of Hong Kong, Hong Kong SAR.