ICES REPORT 15-19

September 2015

A Locally Conservative Enriched Galerkin Approximation and User-Friendly Efficient Solver for Elliptic and Parabolic Problems

by

Sanghyun Lee, Young-Ju Lee, Mary F. Wheeler

The Institute for Computational Engineering and Sciences
The University of Texas at Austin
Austin, Texas 78712

Reference: Sanghyun Lee, Young-Ju Lee, Mary F. Wheeler, "A Locally Conservative Enriched Galerkin Approximation and User-Friendly Efficient Solver for Elliptic and Parabolic Problems," ICES REPORT 15-19, The Institute for Computational Engineering and Sciences, The University of Texas at Austin, September 2015.

A LOCALLY CONSERVATIVE ENRICHED GALERKIN APPROXIMATION AND USER-FRIENDLY EFFICIENT SOLVER FOR ELLIPTIC AND PARABOLIC PROBLEMS

SANGHYUN LEE [†,*] , YOUNG-JU LEE [‡], AND MARY F. WHEELER [†]

Abstract. We present and analyze an enriched Galerkin finite element method (EG) to solve elliptic and parabolic equations with jump coefficients. The EG is formulated by enriching the conforming continuous Galerkin finite element method (CG) with the piecewise constant functions, which can be considered as an additional penalty stabilization. The method is shown to be locally and globally conservative, while keeping lower degree of freedoms in comparisons with discontinuous Galerkin finite element methods (DG). Moreover, we present and analyze a fast, effective and user-friendly EG solver simpler than DG and whose cost is roughly that of CG and can handle an arbitrary order of approximations. A number of numerical tests in two and three dimensions are presented to confirm our theoretical results as well as to demonstrate the advantages of the EG when coupled with transport.

Key words. Enriched Galerkin Finite Element Methods, Auxiliary Space Preconditioner, Locally Conservative Methods, Transport Equations

AMS subject classifications.

1. Introduction.

It is well known that classical continuous Galerkin conforming finite element methods (CG) are relatively easy to apply and implement for general equations. However, the CG approximation suffers from the lack of local conservation on the existing mesh. This can yield non-physical results especially when the flow is coupled with transport. It is important to preserve the local conservation in flow to avoid spurious sources.

There are many published works in designing conservative schemes, which include mixed finite elements [34], dual-grid and control volumes [12], finite volume elements, mimetic finite differences, discontinuous Galerkin finite element methods (DG) [25, 26, 30, 35, 40], and post processing CG [13, 14, 23, 24, 33]. The disadvantages of applying the first two classes of methods lie in their complexities and difficulties in the application of dynamic grid adaptations. Perhaps, the most popular and successful methods in terms of the local conservation property of the flux variables is DG. DG can deal robustly with general partial differential equations as well as with equations whose type changes within the computational domain such as from advection dominated to diffusion dominated [29, 32]. They are naturally suited for multi-physics applications, and for problems with highly varying material properties. However, several disadvantages in the use of the DG are that the method is computationally costly due to the number of degrees of freedom, and the lack of efficient user-friendly solvers that are presently limited to first order. [5]

In [28], Sun and Liu defined the enriched Galerkin finite element method (EG) for second order elliptic equations by enriching CG with the piecewise constant functions.

[†]The Center for Subsurface Modeling, Institute for Computational Engineering and Sciences, University of Texas at Austin, TX 78712 (shlee@ices.utexas.edu, mfw@ices.utexas.edu). The authors are supported by funding from the Center for Frontiers of Subsurface Energy Security, an Energy Frontier Research Center funded by the U.S. Department of Energy, Office of Science, and Office of Basic Energy Sciences, DOE Project #DE-SC0001114.

[‡]Department of Mathematics, Texas State University, TX 78666 (y_l39@txstate.edu). This author's research was supported in part by the Center for Subsurface Modeling at University of Texas at Austin and NSF-DMS # 1358953.

Several advantages of EG are local conservation of the flux with fewer degrees of freedom than that of DG; EG also allows for non-matching grids. In addition, one is not required to enrich the entire computational domain but only the subdomains that require local conservation. Similarly Boffi et al. [8] applied the EG to the fluids problem such as Stokes.

The novelty in this work is

- we design and analyze a fast, effective and user-friendly EG solver for elliptic and parabolic problems, simpler than DG and whose cost is roughly that of CG and which can handle an arbitrary order of approximations for all of the interior penalty bilinear forms; Our algorithm is user-friendly in the sense that classical algebraic multigrid methods can be effectively used as a preconditioner [38].
- we extend and analyze EG to parabolic equations with jump coefficients by using the weighted interior penalty method, e.g. [19, 27];
- we couple EG with a transport system that includes dynamic adaptivity with hanging nodes in two and three dimensions.

The paper is organized as follows. We introduce the governing equations in Section 2, then formulate the corresponding EG approximations. Local and global conservation and the error estimations are established in Section 3. In Section 4, the design and analysis of an efficient solver for EG is presented. In Section 5 we provide numerical experiments to verify the theory and assess the performance of our approach for both elliptic and parabolic as well as coupling with transport.

Throughout the paper, we will use the notation $x_1 \lesssim y_1$, and $x_2 \gtrsim y_2$, whenever there exist positive constants C_1, C_2 independent of the mesh size h such that $x_1 \leq C_1 y_1$ and $x_2 \geq C_2 y_2$, respectively. We also use the notation $x \simeq y$ for $C_1 x \leq y \leq C_2 x$. We shall use the standard notation for Sobolev spaces [1] and their norms. For example, let $E \subseteq \Omega$, then $\|\cdot\|_{1,E}$ and $|\cdot|_{1,E}$ denote the $H^1(E)$ norm and seminorm, respectively. If $E = \Omega$, we shall not indicate the domain where the (semi)norms are taken. For any vector space X, X^d will denote the vector space of size d, whose components belong to X and $X^{d\times d}$ will denote the $d \times d$ matrix whose components belong to X.

2. The Mathematical Model. Let $\Omega \subset \mathbb{R}^d$ be a bounded polygon (for $d = 2$) or polyhedron (for $d = 3$) with Lipschitz boundary $\partial\Omega$ which is decomposed into two parts Γ_D and Γ_N so that $\overline{\partial\Omega} = \overline{\Gamma}_D \cup \overline{\Gamma}_N$, and $(0, \mathbb{T}]$ is the computational time interval with $\mathbb{T} > 0$. We consider the following second order parabolic equation

$$\begin{cases} \rho_0 \dfrac{\partial}{\partial t}(\varphi c_F p) + \nabla \cdot (\rho_0 \mathbf{u}) = q^* & \text{in } \Omega \times (0, \mathbb{T}], \\ p = g_D & \text{on } \Gamma_D \times (0, \mathbb{T}], \\ \mathbf{u} \cdot \mathbf{n} = g_N & \text{on } \Gamma_N \times (0, \mathbb{T}], \end{cases} \tag{2.1}$$

where $p : \Omega \times [0, \mathbb{T}] \to \mathbb{R}$ is the fluid pressure, φ is the porosity, $\mathbf{n}$ denotes the outward pointing unit normal on Γ_N, $q^* \in L^2(\Omega)$ is the source (q^+) or sink (q^-) term and $g_D \in L^2(\Gamma_D)$ and $g_N \in L^2(\Gamma_N)$ are the each Dirichlet and Neumann boundary conditions respectively. We assume that the fluid is slightly compressible with $c_F \ll 1$ in the relation $\rho(p) \approx \rho_0(1 + c_F p)$. Here the initial density of the fluid is simply given as $\rho_0 = 1$. The velocity $\mathbf{u} : \Omega \times [0, \mathbb{T}] \to \mathbb{R}^d$ is defined by Darcy's law

$$\mathbf{u} = -\boldsymbol{\kappa} \nabla p, \tag{2.2}$$

where $\boldsymbol{\kappa}$ is the conductivity, the ratio between the permeability K and the viscosity μ. For $\mathbf{x} \in \Omega$, $K = K(\mathbf{x}) = (K^{ij}(\mathbf{x}))_{i,j=1,\cdots,d}$ denotes the permeability coefficient in

$[L^\infty(\Omega)]^{d\times d}$ and the function $\mu = \mu(\mathbf{x})$ is the fluid viscosity in $L^\infty(\Omega)$, both of which may have the jump discontinuities. We assume that $\boldsymbol{\kappa}$ is uniformly symmetric positive definite, with respect to an initial non-overlapping (open) subdomain partition of the domain Ω. Set $\mathcal{T}_S = \{\Omega_m\}_{m=1}^M$, with $\cup_{m=1}^M \overline{\Omega}_m = \overline{\Omega}$ and $\Omega_m \cap \Omega_n = \emptyset$ for $n \neq m$.

The velocity $\mathbf{u}$ from (2.1) can be coupled to the following reactive transport equation in porous media given as follows:

$$\frac{\partial}{\partial t}(\varphi c) + \nabla \cdot (c\mathbf{u} - D_M \nabla c) = c_b q^*, \quad \text{in } \Omega \times (0,T], \tag{2.3}$$

subject to the boundary conditions

$$\begin{aligned} (c\mathbf{u} - D_M\nabla c)\cdot \mathbf{n} &= c_{\text{in}}\mathbf{u}\cdot\mathbf{n} \quad \text{on } \Gamma_{\text{in}} \times (0,T], \\ D_M \nabla c \cdot \mathbf{n} &= 0, \qquad\qquad \text{on } \Gamma_{\text{out}} \times (0,T], \end{aligned}$$

together with the initial condition

$$c(\mathbf{x},0) = c_0(\mathbf{x}), \quad \forall \mathbf{x} \in \Omega,$$

where c is the species concentration advected, D_M is the molecular diffusion-dispersion coefficient. Here c_b is the given source term and we assume $\varphi = 1$ for simplicity.

The boundary of Ω, denoted by $\partial\Omega$ is decomposed into two parts Γ_{in} and Γ_{out}, the inflow and outflow boundary, respectively, defined as

$$\Gamma_{\text{in}} := \{\mathbf{x} \in \partial\Omega : \mathbf{u}\cdot\mathbf{n} \geq 0\} \quad \text{and} \quad \Gamma_{\text{out}} := \{\mathbf{x} \in \partial\Omega : \mathbf{u}\cdot\mathbf{n} < 0\}. \tag{2.4}$$

where $\overline{\partial\Omega} = \overline{\Gamma}_{\text{in}} \cup \overline{\Gamma}_{\text{out}}$. The coefficients may have large discontinuities across the interfaces between different regions with different material properties.

3. Enriched Galerkin Finite Element Approximations. In this section, we introduce EG to handle (2.1). Let $\mathcal{T}_h$ be the shape-regular (in the sense of Ciarlet) triangulation by a family of partitions of Ω into d-simplices T (triangles/squares in $d = 2$ or tetrahedra/cubes in $d = 3$). We denote by h_T the diameter of T and we set $h = \max_{T\in\mathcal{T}_h} h_T$. Also we denote by $\mathcal{E}_h$ the set of all edges and by $\mathcal{E}_h^o$ and $\mathcal{E}_h^\partial$ the collection of all interior and boundary edges, respectively. In the following notation, we assume 2D edges but the results hold analogously for 3D faces. The boundary edges $\mathcal{E}_h^\partial$ can be further decomposed into $\mathcal{E}_h^\partial = \mathcal{E}_h^{D,\partial} \cup \mathcal{E}_h^{N,\partial}$, where $\mathcal{E}_h^{D,\partial}$ is the collection of edges where the Dirichlet boundary condition is imposed, while $\mathcal{E}_h^{N,\partial}$ is the collection of edges where the Neumann boundary condition is imposed. In addition, we let $\mathcal{E}_h^1 := \mathcal{E}_h^o \cup \mathcal{E}_h^{D,\partial}$ and $\mathcal{E}_h^2 := \mathcal{E}_h^o \cup \mathcal{E}_h^{N,\partial}$.

The space $H^s(\mathcal{T}_h)$ ($s \in \mathbb{R}$) is the set of element-wise H^s functions on $\mathcal{T}_h$, and $L^2(\mathcal{E}_h)$ refers to the set of functions whose traces on the elements of $\mathcal{E}_h$ are square integrable. Let $\mathbb{Q}_k(T)$ denote the space of polynomials of partial degree at most k.

We introduce the space of piecewise discontinuous polynomials of degree k as

$$M^k(\mathcal{T}_h) := \left\{p \in L^2(\Omega) |\ p_{|_T} \in \mathbb{Q}_k(T),\ \forall T \in \mathcal{T}_h\right\}, \tag{3.1}$$

and let $M_0^k(\mathcal{T}_h)$ by the subspace of $M^k(\mathcal{T}_h)$ consisting of continuous piecewise polynomials;

$$M_0^k(\mathcal{T}_h) = M^k(\mathcal{T}_h) \cap \mathbb{C}_0(\Omega).$$

The enriched Galerkin finite element space, denoted by $V_{h,k}^{EG}$ is defined as

$$V_{h,k}^{EG} := M_0^k(\mathcal{T}_h) + M^0(\mathcal{T}_h), \tag{3.2}$$

where $k \geq 1$. We would like to remark that the degrees of freedom for $V_{h,k}^{EG}$ when $k = 1$, is approximately one half and one fourth the degrees of freedom of the linear DG space $M^1(\mathcal{T}_h)$, in two and three space dimensions, respectively.

For the EG formulation, we employ the weighted Interior Penalty (IP) methods as presented in e.g., [27, 19]. To begin with, we define the coefficient $\boldsymbol{\kappa}_T$ by

$$\boldsymbol{\kappa}_T := \boldsymbol{\kappa}|_T, \quad \forall T \in \mathcal{T}_h. \tag{3.3}$$

Following [2], for any $e \in \mathcal{E}_h^o$, let T^+ and T^- be two neighboring elements such that $e = \partial T^+ \cap \partial T^-$. We denote by h_e the length of the edges e. Let $\mathbf{n}^+$ and $\mathbf{n}^-$ be the outward normal unit vectors to ∂T^+ and ∂T^-, respectively ($\mathbf{n}^\pm := \mathbf{n}_{|T^\pm}$). For any given function ξ and vector function $\boldsymbol{\xi}$, defined on the triangulation $\mathcal{T}_h$, we denote $\xi^\pm$ and $\boldsymbol{\xi}^\pm$ by the restrictions of ξ and $\boldsymbol{\xi}$ to $T^\pm$, respectively. Given certain weight $\delta_e \in [0,1]$, we define the weighted average $\{\!\{\cdot\}\!\}_{\delta_e}$ as follows: for $\zeta \in L^2(\mathcal{T}_h)$ and $\boldsymbol{\tau} \in L^2(\mathcal{T}_h)^d$,

$$\{\!\{\zeta\}\!\}_{\delta_e} := \delta_e\zeta^+ + (1-\delta_e)\zeta^- \quad \text{and} \quad \{\!\{\boldsymbol{\tau}\}\!\}_{\delta_e} := \delta_e\boldsymbol{\tau}^+ + (1-\delta_e)\boldsymbol{\tau}^- \quad \text{on } e \in \mathcal{E}_h^o. \tag{3.4}$$

The usual average $\{\!\{\cdot\}\!\}_{1/2}$ will be simply denoted by $\{\!\{\cdot\}\!\}$,

$$\{\!\{\zeta\}\!\} := \{\!\{\zeta\}\!\}_{1/2} \quad \text{and} \quad \{\!\{\boldsymbol{\tau}\}\!\} := \{\!\{\boldsymbol{\tau}\}\!\}_{1/2}, \quad \text{on } e \in \mathcal{E}_h^o.$$

On the other hand, for $e \in \mathcal{E}_h^\partial$, we set $\{\!\{\zeta\}\!\}_{\delta_e} := \{\!\{\zeta\}\!\} = \zeta$ and $\{\!\{\boldsymbol{\tau}\}\!\}_{\delta_e} := \{\!\{\boldsymbol{\tau}\}\!\} = \boldsymbol{\tau}$. The jump across the interior edge will be defined as usual:

$$[\![\zeta]\!] = \zeta^+\mathbf{n}^+ + \zeta^-\mathbf{n}^- \quad \text{and} \quad [\![\boldsymbol{\tau}]\!] = \boldsymbol{\tau}^+ \cdot \mathbf{n}^+ + \boldsymbol{\tau}^- \cdot \mathbf{n}^- \quad \text{on } e \in \mathcal{E}_h^o.$$

For $e \in \mathcal{E}_h^\partial$, we set $[\![\zeta]\!] = \zeta\mathbf{n}$. The choice of the weights has been investigated in [9] and references cited therein. In this paper, we consider the following choice of weights in terms of $\boldsymbol{\kappa}$ given as follows:

$$\beta_e := \frac{\boldsymbol{\kappa}^-}{\boldsymbol{\kappa}^+ + \boldsymbol{\kappa}^-}, \tag{3.5}$$

and let $\delta_e := \beta_e$. The coefficient $\boldsymbol{\kappa}_e$ is defined as the harmonic mean of $\boldsymbol{\kappa}^+$ and $\boldsymbol{\kappa}^-$ by

$$\boldsymbol{\kappa}_e := \frac{2\boldsymbol{\kappa}^+\boldsymbol{\kappa}^-}{\boldsymbol{\kappa}^+ + \boldsymbol{\kappa}^-}. \tag{3.6}$$

We note that the weights $\{\beta_e\}_{e\in\mathcal{E}_h^o}$ depend on the coefficient $\boldsymbol{\kappa}$ and they may vary over all interior edges. We also note that for each $e \in \mathcal{E}_h^o$ the weighted average $\{\!\{\boldsymbol{\kappa}\nabla v\}\!\}_{\beta_e}$ for $\forall v \in H^1(\mathcal{T}_h)$, can be rewritten as $\boldsymbol{\kappa}_e\{\!\{\nabla v\}\!\}$ since

$$\begin{aligned}
\{\!\{\boldsymbol{\kappa}\nabla v\}\!\}_{\beta_e} &= \beta_e(\boldsymbol{\kappa}^+(\nabla v)^+) + (1-\beta_e)(\boldsymbol{\kappa}^-(\nabla v)^-) \\
&= \frac{\boldsymbol{\kappa}^-}{\boldsymbol{\kappa}^+ + \boldsymbol{\kappa}^-}\boldsymbol{\kappa}^+(\nabla v)^+ + \frac{\boldsymbol{\kappa}^+}{\boldsymbol{\kappa}^+ + \boldsymbol{\kappa}^-}\boldsymbol{\kappa}^-(\nabla v)^- \\
&= \frac{\boldsymbol{\kappa}^+\boldsymbol{\kappa}^-}{\boldsymbol{\kappa}^+ + \boldsymbol{\kappa}^-}[(\nabla v)^+ + (\nabla v)^-] = \boldsymbol{\kappa}_e\{\!\{\nabla v\}\!\}\,.
\end{aligned}$$

Finally, we will use the notation:

$$(v,w)_{\mathcal{T}_h} := \sum_{T\in\mathcal{T}_h}\int_T v\,w dx, \quad \forall\ v,w\in L^2(\mathcal{T}_h),$$
$$\langle v,w\rangle_{\mathcal{E}_h} := \sum_{e\in\mathcal{E}_h}\int_e v\,w\,d\gamma, \quad \forall\, v,w\in L^2(\mathcal{E}_h).$$

The inner product on the EG space $V_{h,k}^{EG}$ will then be given as

$$(v,w)_{EG} = (\boldsymbol{\kappa}\nabla v,\nabla w)_{\mathcal{T}_h} + \alpha(k)\sum_{e\in\mathcal{E}_h^1} h_e^{-1}\boldsymbol{\kappa}_e\langle[\![v]\!],[\![w]\!]\rangle, \quad \forall v,w\in V_{h,k}^{EG}. \tag{3.7}$$

The induced EG norm is defined by

$$\|v\|_{EG}^2 := |v|_{1,h,\boldsymbol{\kappa}}^2 + \alpha(k)|[\![v]\!]|_{*,\boldsymbol{\kappa}_e}^2, \tag{3.8}$$

where $\alpha(k)$ is a penalty parameter that is dependent on the degree of polynomial k used for the space $V_{h,k}^{EG}$, $|\cdot|_{1,h,\boldsymbol{\kappa}}^2$ and $|[\![\cdot]\!]|_{*,\boldsymbol{\kappa}_e}^2$ are

$$|v|_{1,h,\boldsymbol{\kappa}}^2 := \sum_{T\in\mathcal{T}_h}\boldsymbol{\kappa}_T\,\|\nabla v\|_{0,T}^2, \ \text{ and } \ |[\![v]\!]|_{*,\boldsymbol{\kappa}_e}^2 := \sum_{e\in\mathcal{E}_h^1} h_e^{-1}\boldsymbol{\kappa}_e\,\|[\![v]\!]\|_{0,e}^2, \tag{3.9}$$

respectively. In addition, we set

$$V(h) := V_{h,k}^{EG} + H^2(\Omega)\cap H_{\Gamma_D}^1 \subset H^2(\mathcal{T}_h), \tag{3.10}$$

where $H_{\Gamma_D}^1 := \{v\in H^1(\Omega) : v|_{\Gamma_D} = g_D\}$ and for $v\in V(h)$, we define the norm,

$$|||v|||^2 := \|v\|_{EG}^2 + \sum_{T\in\mathcal{T}_h} h_T^2\boldsymbol{\kappa}_T|\nabla v|_{1,T}^2. \tag{3.11}$$

We introduce the interpolation operator Π_h for the space $V_{h,k}^{EG}$ as

$$\Pi_h v = \Pi_0^k v + Q^0(v-\Pi_0^k v), \tag{3.12}$$

where Π_0^k is the standard continuous interpolation operator onto the the space $M_0^k(\mathcal{T}_h)$, and Q^0 the L^2 projection onto the space $M^0(\mathcal{T}_h)$. Since the Π_h is a projection for the space of polynomials of degree up to k, the following approximation property holds (see [6])

$$\|v-\Pi_h v\|_{q,T} \lesssim h_T^{\min\{k+1,s\}-q}\|v\|_{s,T}, \quad 0\le q<s. \tag{3.13}$$

In addition, we recall the trace theorem

$$\|\phi\|_{0,e}^2 \lesssim \left(h_e^{-1}\|\phi\|_{0,T}^2 + h_e|\phi|_{1,T}^2\right), \quad \forall\phi\in H^1(\Omega), \tag{3.14}$$

where the estimate depends only on the minimum angle of T. Therefore, we have the following estimate:

$$\|v-\Pi_h v\|_{q,\partial T} \lesssim h_T^{\min\{k+1,s\}-q-1/2}\|v\|_{s,T}, \quad 0\le q<s. \tag{3.15}$$

We introduce the backward Euler EG approximation to (2.1). We define a partition of the time interval $0 =: t^0 < t^1 < \cdots < t^N := \mathbb{T}$ and denote the uniform time step size by $\delta t := t^n - t^{n-1}$. For simplicity, we set $\varphi = c_F = 1$. The EG finite element space approximation of the pressure function $p(\mathbf{x}, t)$ is denoted by $P(\mathbf{x}, t) \in V_{h,k}^{EG}$. Let $P^n := P(\mathbf{x}, t^n)$ for $0 \le n \le N$. We set an initial condition for the pressure as

$$P^0 := \Pi_h p(\cdot, 0). \tag{3.16}$$

Then, the time stepping algorithm reads as follows: Given P^{n-1},

$$\text{Find } P^n \in V_{h,k}^{EG} \text{ such that } \mathcal{S}_\theta(P^n, w) = \mathcal{F}_\theta(w), \quad \forall\, w \in V_{h,k}^{EG}, \tag{3.17}$$

where $\mathcal{S}_\theta$ and $\mathcal{F}_\theta$ are the bilinear form and linear functional as defined in (3.18) and (3.20), respectively, below. First of all,

$$\mathcal{S}_\theta(v, w) := \frac{1}{\delta t}(v, w)_{\mathcal{T}_h} + \mathcal{A}_\theta(v, w), \quad \forall v, w \in V_{h,k}^{EG}, \tag{3.18}$$

where

$$\begin{aligned}\mathcal{A}_\theta(v, w) := (\boldsymbol{\kappa}\nabla v, \nabla w)_{\mathcal{T}_h} - \langle \{\!\{\boldsymbol{\kappa}\nabla v\}\!\}_{\beta_e}, [\![w]\!]\rangle_{\mathcal{E}_h^1} \\ +\theta\, \langle [\![v]\!], \{\!\{\boldsymbol{\kappa}\nabla w\}\!\}_{\beta_e}\rangle_{\mathcal{E}_h^1} + \alpha(k)\langle h_e^{-1}\boldsymbol{\kappa}_e [\![v]\!], [\![w]\!]\rangle_{\mathcal{E}_h^1},\end{aligned} \tag{3.19}$$

and

$$\begin{aligned}\mathcal{F}_\theta(w) := \frac{1}{\delta t}(P^{n-1}, w)_{\mathcal{T}_h} + (f(\cdot, t^n), w)_{\mathcal{T}_h} - \langle g_N, [\![w]\!]\rangle_{\mathcal{E}_h^{N,\partial}} \\ +\theta\, \langle g_D, \{\!\{\boldsymbol{\kappa}\nabla w\}\!\}_{\beta_e}\rangle_{\mathcal{E}_h^{D,\partial}} + \alpha(k)\langle h_e^{-1}\boldsymbol{\kappa}_e g_D, [\![w]\!]\rangle_{\mathcal{E}_h^{D,\partial}}.\end{aligned} \tag{3.20}$$

Here $\alpha(k)$ is a penalty parameter that can vary on edges and h_e denotes the maximum length of the edge $e \in \mathcal{E}_h$. The choice of θ lead to different EG algorithm. For example, $\theta = -1$ leads to SIPG(β)$-k$ methods [27] and [18, Section 4], which later has been extended to the advection-diffusion problems in [9, 17], $\theta = 1$ leads to NIPG(β)$-k$ methods [22], and $\theta = 0$ leads to IIPG(β)$-k$ methods [15].

3.1. Error Analysis for the EG for the Fully Discrete Parabolic Equation. In this section, we present the error analysis of the EG solution for the fully discrete parabolic system. Here we apply the standard approach to extend the elliptic case to the parabolic case. We remark that elliptic and parabolic analysis is presented for general interior penalty schemes $\theta = -1, 0,$ and 1 with the harmonic average of $\boldsymbol{\kappa}$. These cases are not treated in [19, 28]. We establish the following three results i) coercivity ii) continuity and iii) consistency.

3.1.1. Error Analysis for the EG for the Elliptic Equation. First, we show the coercivity for the elliptic part (3.19) of the system without the time derivative.

LEMMA 3.1. *For $\alpha(k) \ge 0$ sufficiently large, the bilinear form $\mathcal{A}_\theta : V_{h,k}^{EG} \times V_{h,k}^{EG}$ is coercive, i.e*

$$\mathcal{A}_\theta(v, v) \gtrsim \|v\|_{EG}^2, \quad \forall v \in V_{h,k}^{EG}. \tag{3.21}$$

Proof. Let $w = v$ in (3.19) to get

$$\mathcal{A}_\theta(v, v) = |v|_{1,h,\boldsymbol{\kappa}}^2 - (1-\theta)\, \langle \boldsymbol{\kappa}_e \{\!\{\nabla v\}\!\}, [\![v]\!]\rangle_{\mathcal{E}_h^1} + \alpha(k)|[\![v]\!]|_{*,\boldsymbol{\kappa}_e}^2, \ \forall v \in V_{h,k}^{EG}. \tag{3.22}$$

Then, we observe the following inequalities hold:

$$\begin{aligned}(3.23)\quad \left|\langle \boldsymbol{\kappa}_e\{\!\{\nabla v\}\!\},[\![v]\!]\rangle_{\mathcal{E}_h^1}\right| &= \left|\sum_{e\in\mathcal{E}_h^1}\int_e \boldsymbol{\kappa}_e\{\!\{\nabla v\}\!\}[\![v]\!]\,d\gamma\right| \\ &\le \left(\sum_{e\in\mathcal{E}_h^1}(\alpha(k))^{-1}h_e\boldsymbol{\kappa}_e\|\{\!\{\nabla v\}\!\}\|_{0,e}^2\right)^{1/2}\left(\sum_{e\in\mathcal{E}_h^1}\alpha(k)h_e^{-1}\boldsymbol{\kappa}_e\|[\![v]\!]\|_{0,e}^2\right)^{1/2} \\ &\le \sum_{e\in\mathcal{E}_h^1}(\alpha(k))^{-1}h_e\boldsymbol{\kappa}_e\|\{\!\{\nabla v\}\!\}\|_{0,e}^2+\frac{\alpha(k)}{4}|[\![v]\!]|_{*,\boldsymbol{\kappa}_e}^2 \\ &\le \sum_{e\in\mathcal{E}_h^1}\frac{C_t(1+C_{\rm inv}^2)}{\alpha(k)}(\boldsymbol{\kappa}^+\|\nabla v\|_{0,T^+}^2+\boldsymbol{\kappa}^-\|\nabla v\|_{0,T^-}^2)+\frac{\alpha(k)}{4}|[\![v]\!]|_{*,\boldsymbol{\kappa}_e}^2,\end{aligned}$$

where C_t and $C_{\rm inv}$ are a positive constant from the Trace theorem and inverse inequality, respectively. The last step of the inequality is due to the following observation

$$\begin{aligned}(3.24)\quad h_e\boldsymbol{\kappa}_e\|\{\!\{\nabla v\}\!\}\|_{0,e}^2 &\le C_t\boldsymbol{\kappa}_e\left(\|\nabla v\|_{0,T^\pm}^2+h_T^2|\nabla v|_{1,T^\pm}^2\right) \\ &\le 2C_t(1+C_{\rm inv}^2)(\boldsymbol{\kappa}^+\|\nabla v\|_{0,T^+}^2+\boldsymbol{\kappa}^-\|\nabla v\|_{0,T^-}^2),\end{aligned}$$

since, $\min\{\boldsymbol{\kappa}^+,\boldsymbol{\kappa}^-\}\le\boldsymbol{\kappa}_e\le 2\min\{\boldsymbol{\kappa}^+,\boldsymbol{\kappa}^-\}$. Then by using (3.8) we obtain,

$$(3.25)\qquad \left|\langle \boldsymbol{\kappa}_e\{\!\{\nabla v\}\!\},[\![v]\!]\rangle_{\mathcal{E}_h^1}\right| \lesssim \frac{C_t(1+C_{\rm inv}^2)}{\alpha(k)}|v|_{1,h,\boldsymbol{\kappa}}^2+\frac{\alpha(k)}{4}|[\![v]\!]|_{*,\boldsymbol{\kappa}_e}^2.$$

Finally, we get

$$\mathcal{A}_\theta(v,v)\gtrsim\left(1-(1-\theta)\frac{12C_t(1+C_{\rm inv}^2)}{\alpha(k)}\right)|v|_{1,h,\boldsymbol{\kappa}}^2+\frac{4-(1-\theta)}{4}\alpha(k)|[\![v]\!]|_{*,\boldsymbol{\kappa}_e}^2.$$

This completes the proof. □

We would like to remark that the parameter $\alpha(k)\ge 0$ imposed for the stability is given independent of the jump coefficients. In addition, for NIPG(β) $-$ k methods (i.e., $\theta=1$), we can choose $\alpha(k)=0$, which still leads to the stability. At (3.24), we observe the choice of the harmonic mean $\boldsymbol{\kappa}_e$ helps the stability in comparison with the standard average, in particular when $\boldsymbol{\kappa}^+\ll 1$ or $\boldsymbol{\kappa}^-\ll 1$. Next, we show the continuity for the elliptic equation.

LEMMA 3.2. *The following continuity results hold true:*

$$(3.26a)\qquad \mathcal{A}_\theta(v,w)\lesssim\|v\|_{EG}\|w\|_{EG},\quad \forall v,w\in V_{h,k}^{EG},$$

$$(3.26b)\qquad \mathcal{A}_\theta(v,w)\lesssim|\!|\!|v|\!|\!|\,|\!|\!|w|\!|\!|,\quad \forall v,w\in V(h).$$

Proof. We recall

$$(3.27)\qquad h_e\boldsymbol{\kappa}_e\|\{\!\{\nabla v\}\!\}\|_{0,e}^2\le C_t\boldsymbol{\kappa}_e\left(\|\nabla v\|_{0,T^\pm}^2+h_T^2|\nabla v|_{1,T^\pm}^2\right).$$

Then, for $v\in V(h)$ we get

$$\begin{aligned}\sum_{e\in\mathcal{E}_h^1}h_e\boldsymbol{\kappa}_e\|\{\!\{\nabla v\}\!\}\|_{0,e}^2 \lesssim& \sum_{T\in\mathcal{T}_h}\boldsymbol{\kappa}^+(\|\nabla v\|_{0,T^+}^2+h_T^2|\nabla v|_{1,T^+}^2) \\ &+\sum_{T\in\mathcal{T}_h}\boldsymbol{\kappa}^-(\|\nabla v\|_{0,T^-}^2+h_T^2|\nabla v|_{1,T^-}^2).\end{aligned}$$

Since we have

$$\left|\langle \boldsymbol{\kappa}_e \{\!\{\nabla v\}\!\}, [\![w]\!]\rangle_{\mathcal{E}_h^1}\right| \leq \left(\sum_{e\in\mathcal{E}_h^1} h_e \boldsymbol{\kappa}_e \|\{\!\{\nabla v\}\!\}\|_{0,e}^2\right)^{1/2} \left(\sum_{e\in\mathcal{E}_h^1} h_e^{-1}\boldsymbol{\kappa}_e \|[\![w]\!]\|_{0,e}^2\right)^{1/2},$$

the following inequality holds true;

$$\left|\langle \boldsymbol{\kappa}_e \{\!\{\nabla v\}\!\}, [\![w]\!]\rangle_{\mathcal{E}_h^1}\right| \lesssim |\!|\!| v |\!|\!| |\!|\!| w |\!|\!|.$$

Also, the other terms have the same upper bound and we obtain

$$\mathcal{A}_\theta(v,w) \lesssim |\!|\!| v |\!|\!| |\!|\!| w |\!|\!|. \tag{3.28}$$

On the other hand, if $v, w \in V_{h,k}^{EG}$, then we apply the inverse inequality in (3.27) to get

$$\mathcal{A}_\theta(v,w) \lesssim \|v\|_{EG}\|w\|_{EG}.$$

This completes the proof. □

Now, we show the consistency of the EG solution for the elliptic case. For this purpose, first we consider the steady state (elliptic problem) of (2.1). Let $\overline{f} \in L^2(\Omega), \overline{g}_D \in L^2(\Gamma_D), \overline{g}_N \in L^2(\Gamma_N)$ and $\overline{p} \in H^s(\mathcal{T}_h)$ with $(s \geq 2)$ be the solution to the following elliptic system,

$$\begin{cases} -\nabla\cdot(\boldsymbol{\kappa}\overline{p}) = \overline{f} & \text{in } \Omega, \\ \overline{p} = \overline{g}_{_D} & \text{on } \Gamma_D, \\ -\boldsymbol{\kappa}\nabla\overline{p}\cdot\mathbf{n} = \overline{g}_{_N} & \text{on } \Gamma_N. \end{cases} \tag{3.29}$$

Let $P_h\overline{p} \in V_{h,k}^{EG}$ be the EG solution to the above system i.e,

$$\begin{aligned} \mathcal{A}_\theta(P_h\overline{p}, w) = (\overline{f}, w)_{\mathcal{T}_h} - \langle \overline{g}_N, [\![w]\!]\rangle_{\mathcal{E}_h^{N,\partial}} + \theta\,\langle \overline{g}_D, \{\!\{\boldsymbol{\kappa}\nabla w\}\!\}_{\beta_e}\rangle_{\mathcal{E}_h^{D,\partial}} \\ +\alpha(k)\langle h_e^{-1}\boldsymbol{\kappa}_e\overline{g}_D, [\![w]\!]\rangle_{\mathcal{E}_h^{D,\partial}}, \quad \forall w \in V_{h,k}^{EG}. \end{aligned} \tag{3.30}$$

Now we state and prove the result on the consistency.

Lemma 3.3. *The following consistency holds :*

$$\mathcal{A}_\theta(\overline{p} - P_h\overline{p}, v) = 0, \quad \forall v \in V_{h,k}^{EG}. \tag{3.31}$$

Proof. Due to the regularity assumption for $\overline{p} \in H^2(\Omega)$, the following jumps and normal fluxes of the solution $\overline{p}$ can be shown [16]:

$$[\![\overline{p}]\!] = [\![\boldsymbol{\kappa}\nabla\overline{p}]\!] = 0, \quad \forall e \in \mathcal{E}_h^o. \tag{3.32}$$

First, we have

$$(\boldsymbol{\kappa}\nabla\overline{p}, \nabla v) = -\sum_{T\in\mathcal{T}_h}\int_T \nabla\cdot(\boldsymbol{\kappa}\nabla\overline{p})v + \sum_{T\in\mathcal{T}_h}\int_{\partial T}\boldsymbol{\kappa}\nabla\overline{p}\cdot\mathbf{n}_T v\, d\gamma, \ \forall v \in V_{h,k}^{EG}, \tag{3.33}$$

by the integration by parts, where $\mathbf{n}_T$ is the outward unit normal to ∂T. The last term can be written as follows :

$$\begin{aligned}\sum_{T\in\mathcal{T}_h}\int_{\partial T}\boldsymbol{\kappa}\nabla\overline{p}\cdot\mathbf{n}_T v\,d\gamma &= \sum_{e\in\mathcal{E}_h}\int_e [\![\boldsymbol{\kappa}\nabla\overline{p}\,v]\!]d\gamma\\ &= \sum_{e\in\mathcal{E}_h}\int_e \{\!\{\boldsymbol{\kappa}\nabla\overline{p}\}\!\}_{\beta_e}[\![v]\!]d\gamma + \sum_{e\in\mathcal{E}_h^o}\int_e [\![\boldsymbol{\kappa}\nabla\overline{p}]\!]\{\!\{v\}\!\}_{\overline{\beta}_e}d\gamma\\ &= \sum_{e\in\mathcal{E}_h}\int_e \{\!\{\boldsymbol{\kappa}\nabla\overline{p}\}\!\}_{\beta_e}[\![v]\!]d\gamma, \quad \text{due to (3.32)},\end{aligned}$$

where $\{\!\{v\}\!\}_{\overline{\beta}_e} := (1-\beta_e)v^+ + \beta_e v^-$. Therefore, we obtain

$$(\boldsymbol{\kappa}\nabla\overline{p},\nabla v) = -\sum_{T\in\mathcal{T}_h}\int_T \nabla\cdot(\boldsymbol{\kappa}\nabla\overline{p})v + \sum_{e\in\mathcal{E}_h}\int_e \{\!\{\boldsymbol{\kappa}\nabla\overline{p}\}\!\}_{\beta_e}[\![v]\!]d\gamma. \tag{3.34}$$

Finally,

$$\begin{aligned}\mathcal{A}_\theta(\overline{p},v) &= (-\nabla\cdot(\boldsymbol{\kappa}\nabla\overline{p}),v)_{\mathcal{T}_h} + \langle\{\!\{\boldsymbol{\kappa}\nabla\overline{p}\}\!\}_{\beta_e}[\![v]\!]\rangle_{\mathcal{E}_h^{N,\partial}}\\ &\quad +\theta\,\langle[\![\overline{p}]\!],\{\!\{\boldsymbol{\kappa}\nabla v\}\!\}_{\beta_e}\rangle_{\mathcal{E}_h^{D,\partial}} + \alpha(k)\langle h_e^{-1}\boldsymbol{\kappa}_e[\![\overline{p}]\!],[\![v]\!]\rangle_{\mathcal{E}_h^{D,\partial}}\\ &= \mathcal{A}_\theta(P_h\overline{p},v), \quad \forall v\in V_{h,k}^{EG}.\end{aligned}$$

□

We now state and prove the main result in this section.

THEOREM 3.4. *Let $\overline{p}\in H^s(\Omega)$ with $s\geq 2$. Then we obtain the following error estimate for the elliptic system* (3.30)*:*

$$\|\overline{p}-P_h\overline{p}\|_{EG} \lesssim h^{\min\{k+1,s\}-1}\|\overline{p}\|_{s,\Omega}. \tag{3.35}$$

Proof. By the coercivity, consistency, and the continuity, we have that

$$\begin{aligned}\|\Pi_h\overline{p}-P_h\overline{p}\|_{EG}^2 &\lesssim \mathcal{A}_\theta(\Pi_h\overline{p}-P_h\overline{p},\Pi_h\overline{p}-P_h\overline{p})\\ &= \mathcal{A}_\theta(\Pi_h\overline{p}-\overline{p},\Pi_h\overline{p}-P_h\overline{p})\\ &\lesssim |||\overline{p}-\Pi_h\overline{p}|||\,\|\Pi_h\overline{p}-P_h\overline{p}\|_{EG}.\end{aligned}$$

Therefore, we arrive at

$$\|\Pi_h\overline{p}-P_h\overline{p}\|_{EG} \lesssim |||\overline{p}-\Pi_h\overline{p}|||.$$

The approximation property of the interpolation operator yields the completion of the proof. □

3.1.2. The L^2 estimate for the SIPG(β)$-k$. In this section, we obtain the L^2 error estimate for the elliptic problem in SIPG(β)$-k$. According to [4], the L^2 estimate for NIPG(β)-k is still an open problem as well as for IIPG(β)-k, ($k>0$). Now, we consider the following adjoint problem to find $\psi\in H^1_{\Gamma_D} := \{\psi\in H^1(\Omega) : \psi = 0 \text{ on } \Gamma_D\}$ such that

$$-\nabla\cdot\boldsymbol{\kappa}\nabla\psi = \overline{p}-P_h\overline{p}, \quad \text{in } \cup_{m=1}^M\Omega_m, \tag{3.36}$$

$$[\![\psi]\!] = 0 \text{ and } [\![\boldsymbol{\kappa}\nabla\psi\cdot\mathbf{n}]\!] = 0, \quad \text{on } \cup_{m=1}^M\partial\Omega_m\setminus\partial\Omega. \tag{3.37}$$

Then, we have [36],

$$\sum_{m=1}^{M} \|\psi\|_{H^2(\Omega_m)} \lesssim \|\overline{p} - P_h\overline{p}\|_{\mathcal{T}_h}. \tag{3.38}$$

We also observe that the following relation holds true :

$$\mathcal{A}_{-1}(\overline{p},\psi) = (\overline{p}, \overline{p} - P_h\overline{p})_{\mathcal{T}_h} \quad \text{and} \quad \mathcal{A}_{-1}(P_h\overline{p},\psi) = (P_h\overline{p}, \overline{p} - P_h\overline{p})_{\mathcal{T}_h}. \tag{3.39}$$

Therefore, due to the orthogonality of the elliptic operator P_h for $\mathcal{A}_\theta$ inner product, we arrive at the following identity

$$\|\overline{p} - P_h\overline{p}\|^2_{\mathcal{T}_h} = \mathcal{A}_{-1}(\overline{p} - P_h\overline{p}, \psi) = \mathcal{A}_{-1}(\overline{p} - P_h, \psi - P_h\psi). \tag{3.40}$$

Finally, we obtain

$$\begin{aligned} \|\overline{p} - P_h\overline{p}\|^2_{\mathcal{T}_h} &\lesssim |||\overline{p} - P_h\overline{p}|||\,|||\psi - P_h\psi||| \lesssim h|||\overline{p} - P_h\overline{p}||| \\ &\lesssim h\|\overline{p} - P_h\overline{p}\|_{\mathcal{T}_h}|||\overline{p} - P_h\overline{p}|||. \end{aligned} \tag{3.41}$$

This completes the proof of the L^2 error estimate.

3.1.3. Error Analysis for the Fully Discrete Parabolic Equation. Finally, we establish the approximation property of the EG solution for the fully discrete parabolic system, (3.17). For the following discussion, we shall denote $p(\cdot,t), p_t(\cdot,t)$ and $p_{tt}(\cdot,t)$ by $p(t), p_t(t)$ and $p_{tt}(t)$ for simplicity.

THEOREM 3.5. *Let P^n and $p(t^n)$ be the discrete and analytic solutions at the time step t^n, respectively. Assume that $p(t^0) \in H^s(\Omega)$ and $p_t \in L^1([0,\mathbb{T}], H^s(\Omega))$ with $s \geq 2$, and $p_{tt} \in L^1([0,\mathbb{T}], H^1(\Omega))$, then we obtain the following error estimate:*

$$\|p(t^n) - P^n\|_{EG} \lesssim h^{\min\{k+1,s\}-1}\left(\|p(t^0)\|_{s,\Omega} + \int_0^{t^n} \|p_t\|_{s,\Omega}\, dt\right) + \delta t \int_0^{t^n} \|p_{tt}\|_{EG}\, dt.$$

Proof. We begin by writing the error expression

$$P^n - p(t^n) = \chi^n + \eta^n, \text{ for } n \geq 0,$$

where $\chi^n = P^n - P_h p(t^n)$ and $\eta^n := P_h p(t^n) - p(t^n)$. From the approximation property of the elliptic projection operator in the EG norm (Theorem 3.4), we have

$$\begin{aligned} \|\eta^n\|_{EG} &\lesssim h^{\min\{k+1,s\}-1}\|p(t^n)\|_{s,\Omega} \\ &\lesssim h^{\min\{k+1,s\}-1}\left(\|p(t^0)\|_{s,\Omega} + \int_0^{t^n} \|p_t\|_{s,\Omega} dt\right). \end{aligned} \tag{3.42}$$

It remains to estimate χ^n. We observe that the following identity holds true

$$\begin{aligned} &(\overline{\partial}_t\chi^n, w)_{\mathcal{T}_h} + \mathcal{A}_\theta(\chi^n, w) \\ &\quad = -((P_h - I)\overline{\partial}_t p(t^n) + (\overline{\partial}_t p(t^n) - p_t(t^n)), w), \quad \forall w \in V^{EG}_{h,k}, \end{aligned} \tag{3.43}$$

with $\overline{\partial}_t$ being the first order Euler approximation of the time derivative of p at t^n. By the coercivity of the bilinear form $\mathcal{A}_\theta(\chi,\chi) \geq 0, \forall \chi \in V^{EG}_{h,k}$, (Lemma 3.1), we arrive at the following estimate:

$$\|\chi^n\|_{EG} \leq \|\chi^0\|_{EG} + \delta t \sum_{j=1}^{n} \|(P_h - I)\overline{\partial}_t p(t^j)\|_{EG} + \delta t \sum_{j=1}^{n} \|\overline{\partial}_t p(t^j) - p_t(t^j)\|_{EG}. \tag{3.44}$$

We observe that

$$\begin{aligned}(3.45)\quad \|\chi^0\|_{EG} &= \|P^0 - P_h p(t_0)\|_{EG} \\ &\le \|P^0 - p(t_0)\|_{EG} + \|p(t_0) - P_h p(t_0)\|_{EG} \lesssim h^{s-1}\|p(t_0)\|_{s,\Omega}.\end{aligned}$$

On the other hand,

$$\begin{aligned}(3.46)\quad \|(P_h - I)\overline{\partial}_t p(t^j)\|_{EG} &= \frac{1}{\delta t}\left\| \int_{t^{j-1}}^{t^j} (P_h - I) p_t dt \right\|_{EG} \\ &\le \frac{1}{\delta t}\int_{t^{j-1}}^{t^j} \|(P_h - I)p_t\|_{EG}\, dt \lesssim \frac{h^{s-1}}{\delta t}\int_{t^{j-1}}^{t^j} \|p_t\|_{s,\Omega}\, dt.\end{aligned}$$

From the Taylor's theorem

$$\overline{\partial}_t p(t^j) - p_t(t^j) = \frac{1}{\delta t}\left(p(t^j) - p(t^{j-1})\right) - p_t(t^j) = \frac{1}{\delta t}\int_{t^{j-1}}^{t^j} (t - t^{j-1}) p_{tt}(t) dt,$$

we get

$$\|\overline{\partial}_t p(t^j) - p_t(t^j)\|_{EG} \le \int_{t^{j-1}}^{t^j} \|p_{tt}(t)\|_{EG} dt.$$

Therefore, we obtain

$$\|\chi^n\|_{EG} \le h^{\min\{k+1,s\}-1}\|p\|_{s,\Omega} + h^{\min\{k+1,s\}-1}\int_0^{t^n} \|p_t(t)\|_{s,\Omega}\, dt + \delta t \int_0^{t^n} \|p_{tt}(t)\|_{EG}\, dt.$$

□

This analysis can be extended to obtain $L^2([0,T]; L^2(\Omega))$ results using arguments presented in [31].

3.2. Conservative Properties of EG. In this section, we introduce the flux variable from the pressure obtained by the EG in such a way that the flux variable enjoys the conservative properties. We can transfer the definition for the continuous level conservative flux variable to its discrete counterpart as follows:

Definition 3.6 (Local Conservation and Global Conservation). *At the time step t^n, a computed velocity $\mathbf{U}^n$ is said to be locally and globally conservative, respectively, if the followings hold true:*

$$\int_{\partial T} \mathbf{U}^n \cdot \mathbf{n}\, ds = \int_T f\, dx - \int_T \overline{\partial}_t P^n\, dx, \quad \forall T \in \mathcal{T}_h, \tag{3.47}$$

$$\int_{\partial\Omega} \mathbf{U}^n \cdot \mathbf{n}\, ds = \int_\Omega f\, dx - \int_\Omega \overline{\partial}_t P^n\, dx, \tag{3.48}$$

where $\mathbf{n}$ is the unit outward normal vector on ∂T.

Such conservative flux variables $\mathbf{U}^n$ can be obtained as follows. Let P^n be the solution to the (3.17), we define the flux variables $\mathbf{U}^n$ at time step t^n, by the following procedure:

$$\begin{aligned}
(3.49a)\quad \mathbf{U}^n|_T &= -\boldsymbol{\kappa}\nabla P^n, \quad \forall T \in \mathcal{T}_h \\
(3.49b)\quad \mathbf{U}^n \cdot \mathbf{n}|_e &= -\boldsymbol{\kappa}_e \{\!\{\nabla P^n\}\!\} \cdot \mathbf{n} + \alpha(k) h_e^{-1} \boldsymbol{\kappa}_e [\![P^n]\!], \quad \forall e \in \mathcal{E}_h^o, \\
(3.49c)\quad \mathbf{U}^n \cdot \mathbf{n}|_e &= g_N, \quad \forall e \in \mathcal{E}_h^{N,\partial}, \\
(3.49d)\quad \mathbf{U}^n \cdot \mathbf{n}|_e &= -\boldsymbol{\kappa}\nabla P^n \cdot \mathbf{n} + \alpha(k) h_e^{-1} \boldsymbol{\kappa}\,(P^n - g_D), \quad \forall e \in \mathcal{E}_h^{D,\partial},
\end{aligned}$$

where $\mathbf{n}$ is the unit normal vector of the boundary edge e of T, in which the value of the normal component of $\mathbf{U}^n$ is defined [28]. We now state and prove the local and global conservation property of such defined flux variables as well as their approximation estimate.

THEOREM 3.7. *Let $\mathbf{u}(t^n)$ be the Darcy velocity (2.2) at time step t^n, and $\mathbf{U}^n$ be the numerical flux obtained as in (3.49) from P^n obtained by the EG as given in (3.17). Then the flux variable $\mathbf{U}^n$ is conservative and also satisfies following error estimate:*

$$\|\mathbf{u}(t^n) - \mathbf{U}^n\|_{\mathcal{T}_h} \lesssim h^{\min\{k+1,s\}-1}\|p(t^n)\|_{s,\Omega}, \tag{3.50}$$

$$\sum_{e\in\mathcal{E}_h} \|\mathbf{u}(t^n)\cdot\mathbf{n} - \mathbf{U}^n\cdot\mathbf{n}\|_{0,e}^2 \lesssim h^{2\min\{k+1,s\}-3}\|p(t^n)\|_{s,\Omega}^2. \tag{3.51}$$

Proof. First, we establish the conservative properties of the flux $\mathbf{U}^n$. For any given $T \in \mathcal{T}_h$, we take $w \in M^0(\mathcal{T}_h)$ such that $w|_T = 1$ and $w|_{\Omega\setminus T} = 0$ in (3.17). Then, it is easy to see that the conservation equation (3.47) holds for the given T. Now, we choose $w \in M^0(\mathcal{T}_h)$ such that w is identically one on Ω for the equation (3.17). Then, we obtain the global conservation as given in (3.48).

Secondly, we perform the error analysis. The $L^2(\mathcal{T}_h)$ estimate can be obtained easily from the following inequalities:

$$\|\mathbf{u}(\cdot,t^n) - \mathbf{U}^n\|_{\mathcal{T}_h} = |p(t^n) - P^n|_{1,h,\boldsymbol{\kappa}} \le \|p(t^n) - P^n\|_{EG}. \tag{3.52}$$

To estimate the error on the interior edge, we define the following gradient of the interpolation:

$$\overline{\mathbf{u}}|_e = -\boldsymbol{\kappa}_e\{\!\{\nabla\Pi_h p(t^n)\}\!\} + \alpha(k)h_e^{-1}\boldsymbol{\kappa}_e[\![p(t^n)]\!] = -\boldsymbol{\kappa}_e\{\!\{\nabla\Pi_h p(t^n)\}\!\}. \tag{3.53}$$

We note that the velocity $\mathbf{u}\cdot\mathbf{n}$ is continuous across the interface [16]. Therefore, we have that

$$\mathbf{u}\cdot\mathbf{n} = -\{\!\{\boldsymbol{\kappa}\nabla p\}\!\}_{\beta_e}\cdot\mathbf{n} = -\boldsymbol{\kappa}_e\{\!\{\nabla p\}\!\}\cdot\mathbf{n}. \tag{3.54}$$

Therefore, by the trace theorem, the boundedness of $\boldsymbol{\kappa}$, and the inverse estimate, we arrive at the following estimate:

$$\begin{aligned}\|\mathbf{u}(t^n)\cdot\mathbf{n} - \mathbf{U}^n\cdot\mathbf{n}\|_{0,e}^2 &\lesssim \|\mathbf{u}(t^n)\cdot\mathbf{n} - \overline{\mathbf{u}}\cdot\mathbf{n}\|_{0,e}^2 + \|\mathbf{U}^n\cdot\mathbf{n} - \overline{\mathbf{u}}\cdot\mathbf{n}\|_{0,e}^2\\ &\lesssim h_T^{2\min\{k+1,s\}-3}\|p(t^n)\|_s^2 + h_T^{2\min\{k+1,s\}-2}\|p(t^n)\|_{s,\Omega}^2.\end{aligned}$$

Note that there is no error on the Neumann boundary edge, while the error estimate on the Dirichlet boundary edge can be done similarly to the case for the interior edges. This completes the proof. ∎

4. Efficient and User-Friendly Solver for EG discretizations. In this section, we design an efficient and user-friendly solver for EG discretizations for elliptic and parabolic equations (2.1). We let

$$M_0^k(\mathcal{T}_h) = \text{span}\{\phi_j\}_{j=1}^{N_0} \quad \text{and} \quad M^0(\mathcal{T}_h) = \text{span}\{\psi_j\}_{j=1}^{N_e},$$

where ϕ_j is the standard $k-$degree nodal basis for the conforming finite element method while ψ_j is the element-wise constant function defined by $\psi_j|_{T_\ell} = \delta_{j\ell}$ for $j,\ell = 1,\cdots,N_e$. Here N_0 and N_e are the number of nodes and elements for the

triangulation $\mathcal{T}_h$, respectively. Under this setting, independent of the choice of $\theta = -1, 0$ and 1, we have the following norm equivalence:

$$\|v\|_{EG}^2 \lesssim \mathcal{A}_\theta(v,v) \lesssim \|v\|_{EG}^2, \quad \forall v \in V_{h,k}^{EG}. \tag{4.1}$$

However,

$$|[\![v]\!]|_{*,\kappa_e}^2 \neq |[\![v_c]\!]|_{*,\kappa_e}^2 + |[\![v_d]\!]|_{*,\kappa_e}^2, \tag{4.2}$$

where $v = v_c + v_d$ with $v_c \in M_0^k(\mathcal{T}_h)$ and $v_d \in M^0(\mathcal{T}_h)$, due to weakly imposed Dirichlet boundary conditions. On the other hand, the space decomposition given in (4.1) is nearly orthogonal with respect to the EG inner product. Therefore one could expect that a block preconditioner similar to DG [3] would be efficient. In following sections, we show theoretically that a robust preconditioner can be achieved if one adds pre- and post- smoothings to the block preconditioner. We note that in a recent work [3], a block preconditioner was devised using a decomposition of the linear DG space into a Crouzeix-Raviart space and a complementary space. However, this requires solution to the Crouzeix-Raviart (C-R) discrete system. Even if it can be solved by a two-grid method using a CG space as a coarse grid, this leads to the introduction of an additional transfer operator between the C-R space and the CG space. Furthermore, for a higher order DG system, the development of a relevant solver has not been reported in the literature to the best of our knowledge. We emphasize that our proposed user-friendly solver for EG systems can be proven to perform well for a arbitrary order of approximations.

For simplicity, we assume a homogenous Dirichlet boundary condition. Furthermore, due to (4.1), it is enough to consider to solve the following system: find $v \in V_{h,k}^{EG}$ such that

$$a(v,w) = \langle f, w\rangle, \quad \forall w \in V_{h,k}^{EG}, \tag{4.3}$$

where

$$a(v,w) := \frac{1}{\delta t}(v,w)_{\mathcal{T}_h} + (v,w)_{EG}, \quad \forall v, w \in V_{h,k}^{EG}. \tag{4.4}$$

The discrete system for the above equation (4.3) with the basis arrangement of $V_{h,k}^{EG}$ leads to the following linear system of equations :

$$\mathbb{S}\underline{\mathrm{p}} = \underline{\mathrm{f}}, \tag{4.5}$$

where $\underline{\mathrm{p}}$ and $\underline{\mathrm{f}}$ are vector representation of P and the linear functional $\langle f, \cdot\rangle$, respectively. Note that the system matrix $\mathbb{S} \in \mathbb{R}^{N\times N}$ with $N = N_0 + N_e$ can be written as follows:

$$\mathbb{S} = \begin{pmatrix} \mathbb{S}_{11} & \mathbb{S}_{12} \\ \mathbb{S}_{21} & \mathbb{S}_{22} \end{pmatrix}, \tag{4.6}$$

where (i,j) component of each block matrix $\mathbb{S}_{ab}$ with $a, b = 1, 2$, is given by the

following:

$$\begin{aligned}
\mathbb{S}_{11}(i,j) &= \frac{1}{\delta t}(\phi_j,\phi_i)_{\mathcal{T}_h} + (\boldsymbol{\kappa}\nabla\phi_j,\nabla\phi_i)_{\mathcal{T}_h} - \langle \{\!\{\boldsymbol{\kappa}\nabla\phi_j\}\!\}_{\beta_e}, [\![\phi_i]\!]\rangle_{\mathcal{E}_h^{D,\partial}} \\
&\quad - \langle [\![\phi_j]\!], \{\!\{\boldsymbol{\kappa}\nabla\phi_i\}\!\}_{\beta_e}\rangle_{\mathcal{E}_h^{D,\partial}} + \alpha(k)\langle h_e^{-1}\boldsymbol{\kappa}_e[\![\phi_j]\!], [\![\phi_i]\!]\rangle_{\mathcal{E}_h^{D,\partial}}, \\
\mathbb{S}_{12}(i,j) &= \frac{1}{\delta t}(\psi_j,\phi_i)_{\mathcal{T}_h} - \langle [\![\psi_j]\!], \{\!\{\boldsymbol{\kappa}\nabla\phi_i\}\!\}_{\beta_e}\rangle_{\mathcal{E}_h^1} + \alpha(k)\langle h_e^{-1}\boldsymbol{\kappa}_e[\![\psi_j]\!], [\![\phi_i]\!]\rangle_{\mathcal{E}_h^{D,\partial}}, \\
\mathbb{S}_{21}(i,j) &= \frac{1}{\delta t}(\phi_j,\psi_i)_{\mathcal{T}_h} - \langle \{\!\{\boldsymbol{\kappa}\nabla\phi_j\}\!\}_{\beta_e}, [\![\psi_i]\!]\rangle_{\mathcal{E}_h^1} + \alpha(k)\langle h_e^{-1}\boldsymbol{\kappa}_e[\![\phi_j]\!], [\![\psi_i]\!]\rangle_{\mathcal{E}_h^{D,\partial}}, \\
\mathbb{S}_{22}(i,j) &= \frac{1}{\delta t}(\psi_j,\psi_i)_{\mathcal{T}_h} + \alpha(k)\langle h_e^{-1}\boldsymbol{\kappa}_e[\![\psi_j]\!], [\![\psi_i]\!]\rangle_{\mathcal{E}_h^1}.
\end{aligned}$$

We observe that the first block $\mathbb{S}_{11}$ is the system matrix that corresponds to the CG bilinear form of the elliptic equation with weakly imposed boundary condition.

4.1. Auxiliary Space Preconditioner. Our algorithm is based on the preconditioned conjugate gradient (PCG) methods with an auxiliary space preconditioner [37]. Here, we propose a preconditioner consisting of three steps: step i) pre-smoothing (4.3), step ii) solving each diagonal block system, $\mathbb{S}_{11}$ and $\mathbb{S}_{22}$, which are the restrictions of (4.3) to $M_0^k(\mathcal{T}_h)$ and $M^0(\mathcal{T}_h)$, respectively, and step iii) post-smoothing (4.3). In addition, we note that one could consider the CG space $M_0^k(\mathcal{T}_h)$ as a coarse grid and apply the two grid method. However, this requires a transfer operator between $M_0^k(\mathcal{T}_h)$ and $M^0(\mathcal{T}_h)$, which is not as user-friendly as our proposed method.

To define an auxiliary space preconditioner, we introduce a fictitious space given by

$$\overline{V} = V_0 \times V_1 \times V_2, \tag{4.7}$$

where $V_0 = V_{h,k}^{EG}$, $V_1 = M_0^k(\mathcal{T}_h)$ and $V_2 = M^0(\mathcal{T}_h)$ and three inner products for each spaces V_i with $i = 0,1,2$. For V_0, we consider $b(\cdot,\cdot) : V_0 \times V_0 \mapsto \mathbb{R}$ defined by $\forall v = v_1 + v_2 \in V_0, w = w_1 + w_2 \in V_0$, $v_i, w_i \in V_i$ for $i = 1,2$, $b(v_1,w_2) = 0, \forall v_1 \in V_1, w_2 \in V_2$, and

$$\begin{aligned}
b(v_1,w_1) &:= \frac{1}{\delta t}(v_1,w_1)_{\mathcal{T}_h} + h^{-2}\sum_{T\in\mathcal{T}_h}\boldsymbol{\kappa}_T\int_T v_1\, w_1\, dx, \\
b(v_2,w_2) &:= a(v_2,w_2).
\end{aligned}$$

For V_1 and V_2, we consider the restriction of $a(\cdot,\cdot)$ on V_1 and V_2, respectively. We then equip the space $\overline{V}$ with the inner product defined as follows:

$$\overline{a}(\overline{v},\overline{v}) := b(v_0,v_0) + a(v_1,v_1) + a(v_2,v_2), \quad \forall \overline{v} = (v_0,v_1,v_2) \in \overline{V}. \tag{4.8}$$

For $i = 1,2$, we introduce an operator $\Pi_i^* : V_0 \mapsto V_i$ defined as simple injections, i.e., $\Pi_i^* v = v_i$, $\forall v = v_1 + v_2 \in V_0$, with $v_i \in V_i$ for $i = 1,2$, and inclusion maps, $\Pi_i : V_i \mapsto V_0$. The operator $\Pi : \overline{V} \mapsto V_0$ can be defined by

$$\Pi\overline{v} = v_0 + \Pi_1 v_1 + \Pi_2 v_2, \quad \forall \overline{v} = (v_0,v_1,v_2) \in \overline{V}. \tag{4.9}$$

We remark that Π_1 and Π_2 are simply the identity which are simpler than those for DG. The preconditioner can be stated as follows:

$$\mathsf{B} = \mathsf{S}^{-1} + \Pi_1 \circ \mathsf{A}_1^{-1} \circ \Pi_1^* + \Pi_2 \circ \mathsf{A}_2^{-1} \circ \Pi_2^*, \tag{4.10}$$

where S, A_1 and A_2 are matrix representations of operators from bilinear forms $b(\cdot,\cdot)$ on V_0, $a(\cdot,\cdot)$ on V_1 and $a(\cdot,\cdot)$ on V_2, respectively. The quality of the preconditioner can be measured from the following estimates:

$$\|v\|_a \le c_b \|v\|_b, \quad \forall v \in V_0, \tag{4.11}$$

where $\|v\|_a^2 := a(v,v)$, $\|v\|_b^2 := b(v,v)$. This inequality (4.11) is due to the Cauchy-Schwarz and inverse inequality. In addition, we establish the lemma:

LEMMA 4.1. *For all $v \in V_0$, there exist $v_0 \in V_0$, $v_1 \in V_1$ and $v_2 \in V_2$ such that $v = v_0 + v_1 + v_2$ and*

$$b(v_0, v_0) + \|v_1\|_a^2 + \|v_2\|_a^2 \le c_0^2 \|v\|_a^2. \tag{4.12}$$

Proof. For the simplicity, we only show the elliptic case, since the time dependent case is similar. Given $v \in V_0$, we define v_i for $i = 1, 2$, by the solution to the following equation:

$$(v_i, w_i)_a = (v, w_i)_a, \quad \forall w_i \in V_i. \tag{4.13}$$

Then, it is immediate that $\|v_i\|_a \lesssim \|v\|_a$ for $i = 1, 2$. We define v_0 by

$$v_0 = v - v_1 - v_2 \in V_0, \tag{4.14}$$

then, it is enough to show that

$$\|v_0\|_b \lesssim \|v\|_a.$$

On the other hand, we note that

$$\|v_0\|_b = \|v - v_1 - v_2\|_b \lesssim \|v - v_1\|_b + \|v_2\|_b \lesssim \|v - v_1\|_b + \|v\|_b.$$

Therefore, in fact, it is enough to show that $\|v - v_1\|_b \lesssim \|v\|_a$. Let χ_E be a characteristic function, which is one on any subset $E \subset \Omega$ and zero elsewhere. We then consider the adjoint problem to find ψ such that

$$-\nabla \cdot \boldsymbol{\kappa} \nabla \psi = \sum_{T \in \mathcal{T}_h} \boldsymbol{\kappa}_T (v - v_1)\chi_T, \quad \text{in } \Omega, \tag{4.15}$$

subject to the interface conditions and homogeneous boundary condition on $\partial\Omega$, given as in the equation (3.36). Then we have

$$\begin{aligned}\sum_{T \in \mathcal{T}_h} \boldsymbol{\kappa}_T \int_T (v - v_1)^2 \, dx = (v - v_1, \psi)_a = (v - v_1, \psi - \chi)_a &\lesssim \|v - v_1\|_a \|\psi - \chi\|_a \\ &\lesssim h \left\| \sum_{T \in \mathcal{T}_h} \boldsymbol{\kappa}_T (v - v_1)\chi_T \right\|_{\mathcal{T}_h} \|v - v_1\|_a, \, \forall \chi \in M_0^{k,0}(\mathcal{T}_h).\end{aligned}$$

Combining the boundedness of $\max_{T \in \mathcal{T}_h} \boldsymbol{\kappa}_T$, we have

$$h^{-2} \sum_{T \in \mathcal{T}_h} \boldsymbol{\kappa}_T \int_T (v - v_1)^2 \, dx \lesssim \|v - v_1\|_a^2 \lesssim \|v\|_a^2.$$

This completes the proof. ∎

Namely, under the aforementioned two estimates (4.11) and Lemma 4.1, we can show that [37]

$$\kappa(\mathsf{BA}) \lesssim (c_0 c_b)^2, \tag{4.16}$$

where A is the matrix that corresponds to the bilinear form $a(\cdot,\cdot) : V_0 \times V_0 \mapsto \mathbb{R}$.

In the remaining section, we show that each block matrices $\mathbb{S}_{ii}$ for $i = 1, 2$ can be easily solved by a classical algebraic multigrid method.

LEMMA 4.2. *The block matrix $\mathbb{S}_{11}$ can be preconditioned by CG discretization part of* (4.4) *with strongly imposed boundary condition.*

Proof. Let $v \in M_0^k(\mathcal{T}_h)$ be the restricted solution to (4.4) with weakly imposed boundary condition while $v_h \in M_0^{k,0}(\mathcal{T}_h)$ is with strongly imposed boundary condition. The framework of the auxiliary space preconditioner requires to show

$$\|v - v_h\|_b + \|v_h\|_a \lesssim \|v\|_a. \tag{4.17}$$

We define $v_h \in M_0^{k,0}(\mathcal{T}_h)$ by

$$a(v_h, w) = a(v, w), \quad \forall w \in H_0^1(\Omega). \tag{4.18}$$

Then, it is immediate to see that $\|v_h\|_a \lesssim \|v\|_a$. On the other hand, using the auxiliary problem considered in Lemma 4.1, we can obtain that $\|v - v_h\|_b \lesssim \|v\|_a$. ∎

The following Lemma indicates that the block matrix $\mathbb{S}_{22}$ can be easily handled by a classical algebraic multigrid method.

LEMMA 4.3. *The matrix $\mathbb{S}_{22}$ is M-matrix and weakly diagonally dominant.*

Proof. It is immediate to see that the time derivative term restricted on V_2 is positive diagonal and

$$\mathbb{S}_{22}(i,i) = \frac{1}{\delta t}(\psi_i, \psi_i)_{\mathcal{T}_h} + \alpha(k) \sum_{e \in \mathcal{E}_h} \frac{\boldsymbol{\kappa}_e}{h_e} |e| [\![\psi_i]\!][\![\psi_i]\!] > 0. \tag{4.19}$$

On the other hand, for $i \neq j$, $\mathbb{S}_{22}(i,j) = \alpha(k) \sum_{e \in \mathcal{E}_h} \frac{\boldsymbol{\kappa}_e}{h_e} |e| [\![\psi_j]\!][\![\psi_i]\!] < 0$. Furthermore, we have $\sum_{j=1, j\neq i}^{N_e} |\mathbb{S}(i,j)| \leq \mathbb{S}(i,i), \forall i \geq 1$ and the strict inequality can be achieved when the triangle $T_i \in \mathcal{T}_h$ has at least one edge that is not shared by a neighbor triangle. This shows that the matrix $\mathbb{S}$ is weakly diagonally dominant. ∎

5. Numerical Results. This last section illustrates and evaluates the performance of the EG coupled with the new user-friendly solver and transport system. All examples are computed with the open-source finite element package deal.II [7] and FASP package [39]. In particular, the three dimensional implementation is based on the MPI-library[20], and the subdivision and mesh distribution is implemented using the p4est library [10].

We start in Section 5.1 by validating the theoretical convergence results for the EG method. The numerical tests of the proposed AMG based solver is then presented in Section 5.2. In Section 5.3, we investigate the model problem (2.1)-(2.4) and compare concentration values and local conservative residuals between CG and EG approximations.

	Dirichlet		Neumann	
δt	Error	Rate	Error	Rate
0.01	0.080252	0.0000	0.080256	0.000
0.005	0.040158	1.0458	0.040158	1.0459
0.0025	0.020083	1.0227	0.020083	1.0228

Table 1: Convergence rate regarding the spatial discretization for each boundary condition.

	SIPG(β)-1		NIPG(β)-1		IIPG(β)-1	
δt	Error	Rate	Error	Rate	Error	Rate
0.01	0.080257	0.0000	0.080257	0.0000	0.080256	0.0000
0.005	0.040162	1.0457	0.040160	1.0458	0.040160	1.0458
0.0025	0.020086	1.0227	0.020085	1.0227	0.020084	1.0227

Table 2: Convergence rate regarding the spatial discretization for each different EG methods.

5.1. Convergence rate estimates. Let $\Omega = (0,1)^2$ and we consider the system (2.1) with the known analytic solution given as

$$p = \cos(t + x - y), \quad \text{in } \Omega \times [0, \mathbb{T}]. \tag{5.1}$$

For the simplicity, we assume $\mu = c_F = \rho_0 = \varphi = 1$, and $K = diag(k)$ with its entry $k = 1$. We consider both Dirichlet and Neumann boundary condition cases. The error defined in the Theorem 3.5 is computed at the final time $\mathbb{T} = 0.2$. Table 1 illustrates three computations done on three uniform meshes with constant time steps. The mesh size and the time step are divided by two each time. The meshes are composed of 145, 545, and 2113 $\mathbb{Q}_1$-EG degree of freedoms and here $\theta = 0$ is used, i.e. the IIPG(β)-1 method. The time discretization is chosen fine enough not to influence the time error. The penalty constant is given large enough as $\alpha = 100$ with $r = 1$. Note that the Table 1 is divided into two columns, the first of which shows the results obtained by using the solely Dirichlet boundary condition while the second column shows the results obtained by using the solely Neumann boundary condition. Apparently, we have achieved the expected convergence rate for both cases.

Next, in Table 2 we test the SIPG(β)-1, NIPG(β)-1 and IIPG(β)-1 cases. The Dirichlet boundary condition is given on the left and top boundaries and Neumann boundary condition is given on the right and bottom boundaries. We also observe the proven convergence rates shown at Theorem 3.5.

5.2. Effectiveness of the solver. In this section, we present a number of numerical tests that can confirm the efficacy of our numerical algorithm to solve the EG system provided in Section 4. We solve (2.1) in 2D and 3D cases by varying the K values with and without the time derivative term to see the performance of each different solvers. Here $q^* = 0$, and the homogeneous Dirichlet boundary condition $g_D = 0$ is weakly imposed.

First, we test PCG-AMG (the preconditioned conjugate gradient method with AMG preconditioner), PCG-BMG (the diagonal block preconditioner with added sin-

DOFs	PCG-AMG				PCG-BMG0				PCG-BMG			
	Ell.		Par.		Ell.		Par.		Ell.		Par.	
	(i)	(ii)	(i)	(ii)	(i)	(ii)	(i)	(ii)	(i)	(ii)	(i)	(ii)
545	16	14	16	13	9	9	9	9	6	6	6	6
2113	69	48	58	38	9	10	9	11	6	6	7	6
8321	>100	>100	>100	>100	9	11	9	11	7	7	7	7
33025	>100	>100	>100	>100	9	11	9	10	7	7	7	7
131585	>100	>100	>100	>100	9	10	9	10	7	7	7	7

Table 3: The iteration numbers for each solver with different cases in 2D.

DOFs	PCG-AMG				PCG-BMG0				PCG-BMG			
	Ell.		Par.		Ell.		Par.		Ell.		Par.	
	(i)	(ii)	(i)	(ii)	(i)	(ii)	(i)	(ii)	(i)	(ii)	(i)	(ii)
189	1	1	1	1	9	9	9	9	5	5	5	5
1241	>100	>100	>100	>100	13	13	13	13	6	6	6	6
9009	>100	>100	>100	>100	11	12	11	12	7	6	7	7
68705	>100	>100	>100	>100	11	12	11	12	7	7	7	7

Table 4: The iteration numbers for each solver with different cases in 3D.

gle pre- and post- smoothing step), and PCG-BMG0 (the diagonal block preconditioner without smoothing steps) with applying SIPG(β)-1 method. The test results are provided in Table 3 for 2D and Table 4 for 3D problem. Each result shows iteration numbers taken for PCG-AMG, PCG-BMG0 and PCG-BMG to achieve the stopping criterion given in terms of the preconditioned relative residual less than 10^{-7}. In all cases, a single V-cycle of the classical AMG is used as a preconditioner. The time independent elliptic problem ($\partial_t p = 0$) is denoted by Ell. and the time dependent parabolic problem is denoted by Par. with $\delta t = 0.5$. The first column (i) is for uniform permeability ($k = 1$) and the second column (ii) is for random permeability $k \in [10^{-3}, 1]$.

In Table 3 and Table 4, we clearly observe that the direct application of the AMG (PCG-AMG) deteriorates in its performance. We remark that PCG-BMG0 also work robustly which seems to be due to the near orthogonality discussed in Section 4.

5.2.1. Preconditioner for $\mathbb{S}_{11}$. In this section, we solve the elliptic EG system (4.3) on $M_0^1(\Omega)$ with weakly imposed Dirichlet boundary condition using the preconditioner discussed in Section 4.2. We apply three different preconditions denoted by PCG-AMG, PCG-type I (the system (4.3) on $M_0^1(\Omega)$ with strongly imposed boundary condition as a preconditioner with pre and post smoothings), and PCG-type II (the system (4.3) on $M_0^1(\Omega)$ with strongly imposed boundary condition as a preconditioner without pre and post smoothings). As predicted in Section 4.2, Table 5 shows that PCG-Type I provides robust results. However, the other two preconditioners still show uniform convergence whose proof is beyond the scope of the paper.

Next, we solve the elliptic EG discrete system (3.17) with $\theta = 0$ and 1 for both 2D and 3D, using PCG-Type I for inverting the block matrix $\mathbb{S}_{11}$ and PCG-AMG for the block matrix $\mathbb{S}_{22}$. The outer iteration is done with GMRES method. This algorithm is denoted by PGMRES-BMG. The results with uniform and random K are presented in Table 6 for 2D and Table 7 for 3D, respectively.

DOFs	PCG-AMG	PCG-Type I	PCG-Type II
4225	7	5	25
16641	7	6	28
66049	7	6	27
263169	7	6	27
1050625	7	7	29

Table 5: Iteration number for PCG-AMG, PCG-Type I and PCG-Type-II applied for SIPG(β)-1.

DOFs	PGMRES-BMG for Uniform K		PGMRES-BMG for Random K	
	IIPG(β)-1	NIPG(β)-1	IIPG(β)-1	NIPG(β)-1
545	6	6	6	6
2113	7	8	6	8
8321	7	9	7	8
33025	7	9	7	9

Table 6: Iteration numbers for PGMRES-BMG for IIPG(β)-1 and NIPG(β)-1 for 2D.

DOFs	PGMRES-BMG for Uniform K		PGMRES-BMG for Random K	
	IIPG-(β)-1	NIPG-(β)-1	IIPG-(β)-1	NIPG-(β)-1
189	6	6	6	6
1241	6	7	6	7
9009	6	7	6	7
68705	6	8	6	8

Table 7: Iteration numbers for PGMRES-BMG for IIPG$-(\beta)$ and NIPG$-(\beta)$ for 3D.

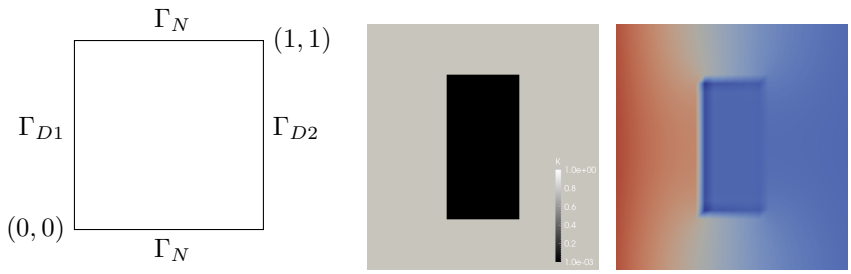
5.3. Coupled with Transport Equation. In this section, we solve a coupled problem of flow (2.1) and transport (2.3) system. First, we solve the pressure system by both EG and CG, and use the respective velocities to transport the concentration $c : \Omega \to \mathbb{R}$. In particular, the EG flow is computed using the formulas given in (3.49).

Let $C(\mathbf{x}, t)$ be the space approximation of the concentration function $c(\mathbf{x}, t)$ and the time approximation of $C(\mathbf{x}, t^n), 0 \le n \le N$ be denoted by C^n. We employ piecewise constant approximation (DG-$\mathbb{Q}^0$) for the transport system. Given C^0 , the discretized system to find $C^{n+1} \in M^0(\mathcal{T}_h)$ is given as follow:

$$
\begin{aligned}
(5.2)\quad & \left(\frac{C^{n+1}-C^n}{\delta t}, w\right)_{\mathcal{T}_h} - (C^{n+1}\mathbf{U}^{n+1}, \nabla w)_{\mathcal{T}_h} + (q^-(t^{n+1})C^{n+1}, w)_{\mathcal{T}_h} \\
& + \sum_{e\in\mathcal{E}_h}\int_e (C^{n+1})^*\mathbf{U}^{n+1}\cdot\mathbf{n}[\![w]\!]\, d\gamma + \sum_{e\in\Gamma_{\text{out}}}\int_e C^{n+1}\mathbf{U}^{n+1}\cdot\mathbf{n}w\, d\gamma \\
& = (q^+(t^{n+1}), w)_{\mathcal{T}_h} + \sum_{e\in\Gamma_{\text{in}}}\int_e c_{in}\mathbf{U}^{n+1}\cdot\mathbf{n}w\, d\gamma,\ \forall w \in M^0(\mathcal{T}_h)\times(0,\mathbb{T}],
\end{aligned}
$$

where $D_M = 0$ for simplicity, and the upwind value of concentration is

$$
(C^{n+1})^*_{|e} := \begin{cases} (C^{n+1})^+ & \text{if } \mathbf{U}^{n+1}\cdot\mathbf{n} \ge 0 \\ (C^{n+1})^- & \text{if } \mathbf{U}^{n+1}\cdot\mathbf{n} < 0. \end{cases}
$$



(a) Domain (b) K (c) Pressure

Fig. 1: Example 1. (a) Computational domain with boundary conditions (b) permeability block in the domain, and (c) computed steady state pressure value.

5.3.1. Example 1. Permeability Block. Let $\Omega = (0,1)^2$ and the permeability tensor is defined as a diagonal tensor $K = diag(k)$ with its entry $k = 10^{-3}$ in the subdomain $\Omega_c = (\frac{3}{8}, \frac{5}{8}) \times (\frac{1}{4}, \frac{3}{4})$ and $k = 1$ else where. Here $\mu = 1$ and we assume slightly compressible flow by using $c_F = 10^{-8}$. See Figure 1 for the setup. The inflow boundary condition for the transport system (2.3) is given as

$$c_{in} = 1 \quad \text{on} \quad \Gamma_{\text{in}} \times (0, \mathbb{T}],$$

and the initial conditions for the pressure and concentration is zero, i.e $c_0 = 0$ and $p_0 = 0$. The boundary conditions for the pressure system are chosen as

$$p = 1 \text{ on } \Gamma_{D1} \times [0, \mathbb{T}], \quad p = 0 \text{ on } \Gamma_{D2} \times [0, \mathbb{T}], \quad \text{and} \quad \frac{\partial p}{\partial \mathbf{n}} = 0 \text{ on } \Gamma_N \times [0, \mathbb{T}].$$

First, we compute the maximum of local conservative residual $\max_{T \in \mathcal{T}_h} \|R_{loc}\|_{L^\infty(T)}$

		CG		EG	
Cycle	δt	Dofs	R_{loc}	Dofs	R_{loc}
1	0.01	289	0.0190	145	7×10^{-8}
2	0.005	1089	0.0115	545	7×10^{-8}
3	0.0025	4225	0.0071	2113	7×10^{-8}

Table 8: Example 1. Local conservation residual values for CG and EG method. EG provides more accurate local conservation.

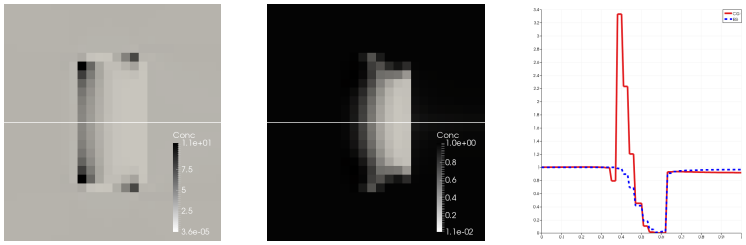
defined (3.47) by

$$R_{loc} := \int_{\partial T} \mathbf{U}^n \cdot \mathbf{n} \, ds - \int_T f \, dx + \int_T \overline{\partial}_t c_F P^n \, dx, \quad \forall T \in \mathcal{T}_h,$$

at the final computation time $\mathbb{T} = 0.1$ and compare the values between CG flow and EG flow. The meshes are composed of 145, 545, and 2113 for $\mathbb{Q}_1$-EG degree of freedoms and 289, 1089, and 4225 for $\mathbb{Q}_1$-CG degree of freedoms (Dofs). We observe in Table 8 that the EG provides more accurate local conservation as presented in Theorem 3.7 even if EG has less number of degree of freedoms compare to the CG.

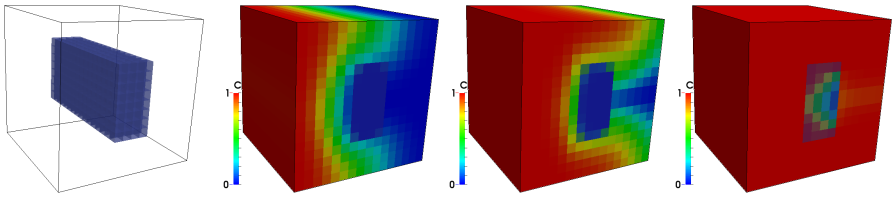
Next, we compute the same setup with more longer computation time $\mathbb{T} = 10$. At $t = 10$, we plot the value of concentration on the line $(0, 0.5) - (1, 0.5)$ which

crosses the low permeable zone, see Figure 2 (a)-(b). From the Figure 2 (c), we observe that the concentration values are significantly accumulated at the interface between the different permeabilities for the CG flux. However, the maximum value of concentration is around 1 as expected with the EG flux .



(a) Using CG flow (b) Using EG flow (c) Comparison

Fig. 2: Example 1. (a)-(b) The value of concentration at $t = 10$ for CG and EG respectively. Note that the concentration value goes up to 10 when using CG. (c) The value of concentration on the line $(0, 0.5) - (1, 0.5)$. EG preserves the expected maximum value of concentration.



(a) Permeability (b) $t = 0.5$ (c) $t = 1$ (d) $t = 10$

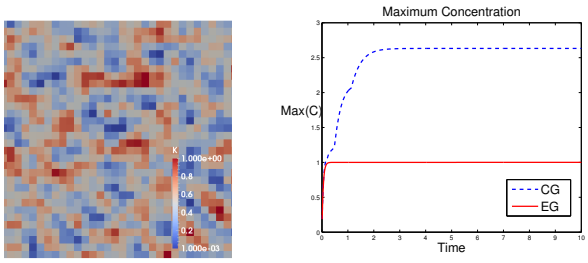
Fig. 3: Example 1. Three dimensional case. (a) The permeability block and (b)-(d) the concentration value for each time by employing the flux from EG.

We can extend this to the three dimensional setup as shown at Figure 3 (a). The permeability block is given in the middle of the domain $\Omega_c = (\frac{3}{8}, \frac{5}{8}) \times (\frac{1}{4}, \frac{3}{4}) \times (0, 1)$ with $k = 10^{-3}$ and $k = 1$ elsewhere. The Figure 3 illustrates the concentration value along each time by using the flux from locally conservative EG.

5.3.2. Example 2. Random Permeability. In this example, the diagonal entry of the permeability $K = diag(k)$ is a randomly generated piecewise constant between $k \in [0.001, 1]$ defined on $\Omega = (0, 1)^2$. See Figure 4 (a) for the setup. We have 2113 and 1089 degree of freedoms for EG and CG respectively, and the time step is $\delta t = 0.01$ with $\mathbb{T} = 10$. All other settings and parameters are same as the previous example.

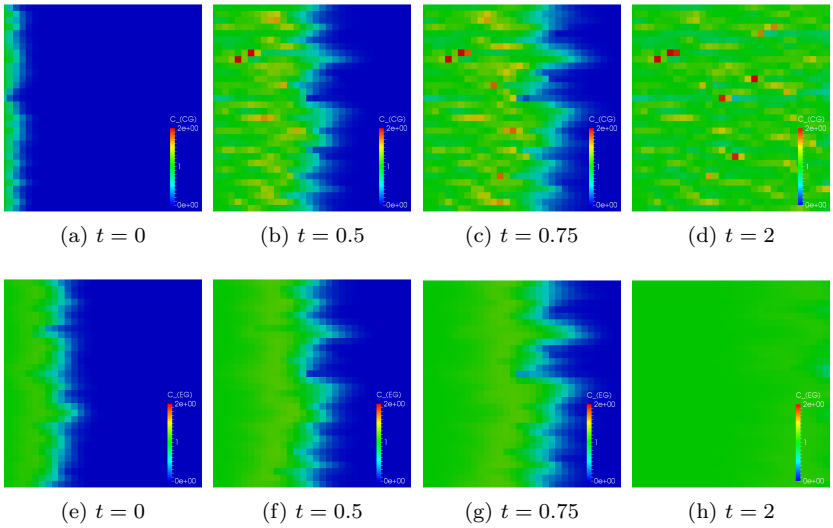
First, the Figure 4 (b) plots the maximum concentration value for the entire computation time. Note that in our case we need to preserve

$$\max_{T \in \mathcal{T}_h} \|c\|_{L^\infty(T)} \approx C_{in} = 1.$$



(a) Permeability (b) $\max_{T\in\mathcal{T}_h} \|C\|_{L^\infty(T)}$

Fig. 4: Example 2. (a) Permeability setup (b) $\max_{T\in\mathcal{T}_h} \|C\|_{L^\infty(T)}$ in the time interval.



(a) $t = 0$ (b) $t = 0.5$ (c) $t = 0.75$ (d) $t = 2$

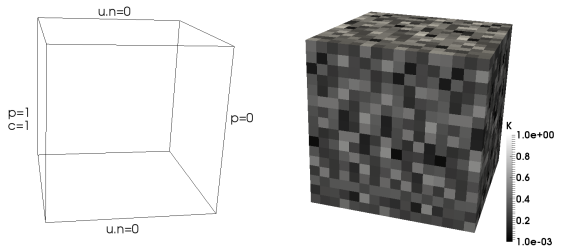
(e) $t = 0$ (f) $t = 0.5$ (g) $t = 0.75$ (h) $t = 2$

Fig. 5: Example 2. Concentration values for each time using CG flux (Top) and EG flux (Bottom).

However, the CG flux leads the overshoot for the maximum concentration value (dotted line) unlike the EG flux (solid line).

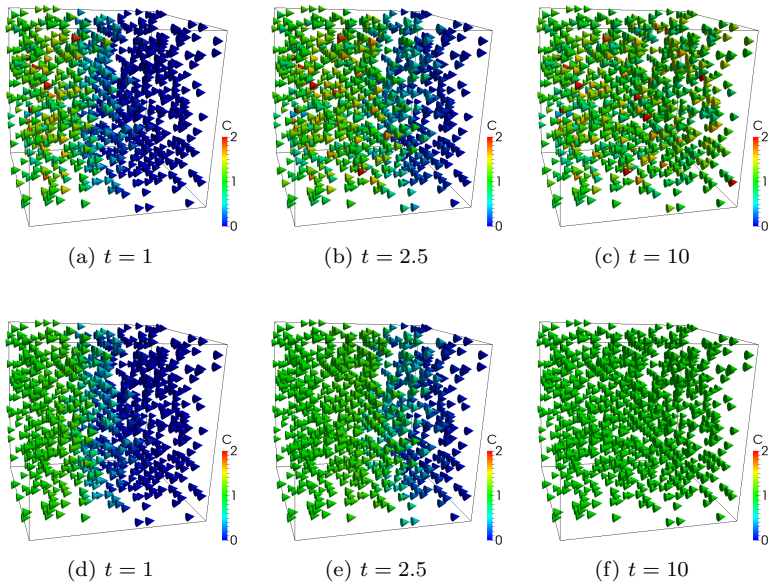
In Figure 5, we illustrate the concentration values for each time by using the CG flux (a-d) and EG flux (e-h). We observe some oscillation values when we impose non-locally conservative flux to the transport system (Figure 5 (a)-(d)). We get some overshooting of concentration values where the permeability is lower than others. However, employing locally conservative flux by EG method shows expected accurate results (Figure 5 (e)-(h)).

We also consider three dimensional case $\Omega = (0, 4)^3$ with boundary conditions as shown at Figure 6. The final computational time is $\mathbb{T} = 4$ and the time step size is $\delta t = 0.05$. Again we impose the random permeability with $k \in [0.001, 1]$ (Figure 6 (b)) and the same setup as the previous example. Similar to the two dimensional results, we also see oscillations of the concentration values when we employ non locally



(a) Domain (b) Permeability K

Fig. 6: Example 2. Domain and boundary condition for the three dimensional test case with (b) given random permeability values.



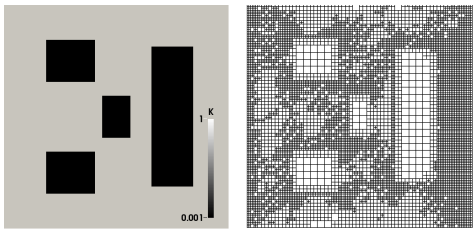
(a) $t = 1$ (b) $t = 2.5$ (c) $t = 10$

(d) $t = 1$ (e) $t = 2.5$ (f) $t = 10$

Fig. 7: Example 2. Concentration values for the each time; (a)-(c) using the flux from the CG and (d)-(f) using the flux from the EG. We observe the oscillations of concentration values using the CG flow.

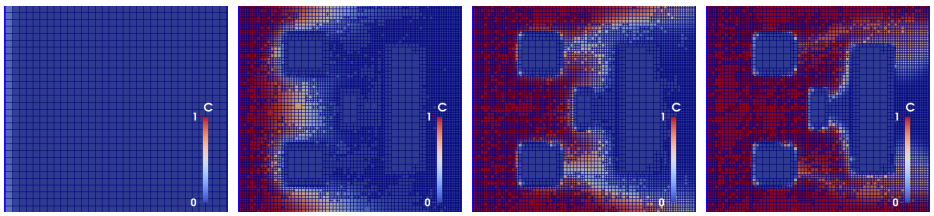
conservative fluxes to the transport system, see Figure 7.

5.3.3. Example 3. Multiple Blocks with Adaptive Finite Elements. In this section, we emphasize the dynamic local mesh refinement using the hanging nodes strategy [21, 11] . In $\Omega = (0, 1)^2$, the diagonal entry of the permeability $K = diag(k)$ is defined with its entry $k = 10^{-3}$ in the subdomain $\Omega_c = (0.2, 0.4) \times (0.65, 0.85) \cup (0.2, 0.4) \times (0.15, 0.35) \cup (0.45, 0.55) \times (0.4, 0.6) \cup (0.65, 0.85) \times (0.2, 0.8)$ and 1 else where. See Figure 8 (a) for the setup. The time step is $\delta t = 0.01$ with $\mathbb{T} = 10$. All other settings are same as the previous example. We initially refine the domain with



(a) Permeability K (b) Meshes

Fig. 8: Example 3. (a) Permeability setup (b) Mesh refinement regarding to the local conservation residual at $t = 1$.

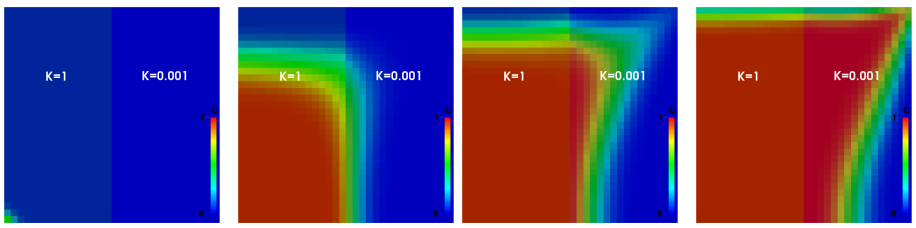


(a) $t = 0$ (b) $t = 0.5$ (c) $t = 1.$ (d) $t = 1.5$

Fig. 9: Example 3. Concentration values for each time with local mesh refinement.

1024 cells and refined more if the local conservation residual (R_{loc}) is higher than the certain criteria at each time step. To restrict the maximum number of the cells, we only refine up to the maximum refinement level 7.

The Figure 9 illustrates the value of the concentration for the each time steps with the dynamic local mesh adaptivity regarding to the local conservation residual.



(a) $t = 0$ (b) $t = 4$ (c) $t = 6$ (d) $t = 10$

Fig. 10: Example 4. Concentration values for each time with injection and extraction.

5.4. Example 4. Injection and Extraction. In this final example, we consider more real world scenario in $\Omega = (0, 1)^2$. The diagonal entry of the permeability $K = diag(k)$ is defined with its entry $k = 10^{-3}$ if $x \geq 0.5$ and $k = 1$ elsewhere. The fluid is injected on the corner $(0, 0)$ with the injection rate $q^+ = 100$, and it is

extracted on the corner $(1,1)$ with the extraction rate $q^- = -100$, with $c_b = 1$. The time step is $\delta t = 0.01$ with $\mathbb{T} = 10$. All other settings are same as the previous example. It is well known that non-locally conservative flux will result in huge oscillation at $x = 0.5$ with this setup.

In the Figure 10, we observe that the fluid smears into the low permeability zone at $t = 4$ and move towards the extraction point. Using EG we do not observe any oscillation on the interface $x = 0.5$ where we have the permeability jump.

6. Conclusion. In this paper, we focused on enriched Galerkin (EG) approximation methods for elliptic and parabolic equations and couplings with transport. The EG algorithms preserve the local conservation property and can handle non-matching grids. We developed and analyzed robust user-friendly EG solvers which can handle an arbitrary order of approximations. We provided convergence and conservation analysis for the method by employing the weighted interior penalty method. We demonstrated numerical examples considering three dimensional cases and employing the local mesh adaptivity to show that the methods perform as our theory predicts.

REFERENCES

[1] R. A. Adams. *Sobolev spaces*. Academic Press [A subsidiary of Harcourt Brace Jovanovich, Publishers], New York-London, 1975. Pure and Applied Mathematics, Vol. 65.

[2] 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(5):1749–1779 (electronic), 2001/02.

[3] B. Ayuso, I. Georgiev, J. Kraus, and L. Zikatanov. Preconditioning techniques for discontinuous Galerkin methods discretizations for linear elasticity equations. in preparation, 2009.

[4] B. Ayuso de Dios, F. Brezzi, O. Havle, and L. D. Marini. L^2-estimates for the DG IIPG-0 scheme. *Numer. Methods Partial Differential Equations*, 28(5):1440–1465, 2012.

[5] B. Ayuso de Dios, M. Holst, Y. Zhu, and L. Zikatanov. Multilevel preconditioners for discontinuous, galerkin approximations of elliptic problems, with jump coefficients. *Mathematics of Computation*, 83(287):1083–1120, 2014.

[6] I. Babuška and M. Suri. The optimal convergence rate of the p-version of the finite element method. *SIAM J. Numer. Anal.*, 24:199–238, 1987.

[7] W. Bangerth, T. Heister, G. Kanschat, and M. others. *Differential Equations Analysis Library*, 2012.

[8] D. Boffi, N. Cavallini, F. Gardini, and L. Gastaldi. Local mass conservation of stokes finite elements. *J. Sci. Comput.*, 52(2):383–400, 2012.

[9] E. Burman and P. Zunino. A domain decomposition method based on weighted interior penalties for advection-diffusion-reaction problems. *SIAM J. Numer. Anal.*, 44(4):1612–1638, 2006.

[10] C. Burstedde, L. C. Wilcox, and O. Ghattas. `p4est`: Scalable algorithms for parallel adaptive mesh refinement on forests of octrees. *SIAM Journal on Scientific Computing*, 33(3):1103–1133, 2011.

[11] C. Carstensen and J. Hu. Hanging Nodes in the Unifying Theory of a Posteriori Finite Element error control. *Computing*, 80(3):203–220, July 2007.

[12] P. Chatzipantelidis, V. Ginting, and R. Lazarov. A finite volume element method for a nonlinear elliptic problem. *Numer. Linear Algebra with Appl.*, 00(2):1–26, 2004.

[13] S. Chippada, C. Dawson, M. Martnez, and M. Wheeler. A projection method for constructing a mass conservative velocity field. *Computer Methods in Applied Mechanics and Engineering*, 157(12):1 – 10, 1998.

[14] B. Cockburn, J. Gopalakrishnan, and H. Wang. Locally conservative fluxes for the continuous galerkin method. *SIAM Journal on Numerical Analysis*, 45(4):1742–1776, 2007.

[15] C. Dawson, S. Sun, and M. F. Wheeler. Compatible algorithms for coupled flow and transport. *Comput. Methods Appl. Mech. Engrg.*, 193(23-26):2565–2580, 2004.

[16] D. A. Di Pietro and A. Ern. Analysis of a discontinuous Galerkin methods for heterogeneous diffusion problems with low-regularity solutions. *Numerical Methods for Partial Differential Equations*, 28(4):1161–1177, 2011.

[17] D. A. Di Pietro, A. Ern, and J.-L. Guermond. Discontinuous Galerkin methods for anisotropic semidefinite diffusion with advection. *SIAM J. Numer. Anal.*, 46(2):805–831, 2008.
[18] M. Dryja. On discontinuous Galerkin methods for elliptic problems with discontinuous coefficients. *Comput. Methods Appl. Math.*, 3(1):76–85, 2003.
[19] A. Ern, A. F. Stephansen, and P. Zunino. A discontinuous Galerkin method with weighted averages for advection-diffusion equations with locally small and anisotropic diffusivity. *IMA J. Numer. Anal.*, 29(2):235–256, 2009.
[20] E. Gabriel, G. E. Fagg, G. Bosilca, T. Angskun, J. J. Dongarra, J. M. Squyres, V. Sahay, P. Kambadur, B. Barrett, A. Lumsdaine, R. H. Castain, D. J. Daniel, R. L. Graham, and T. S. Woodall. Open MPI: Goals, concept, and design of a next generation MPI implementation. In *Proceedings, 11th European PVM/MPI Users' Group Meeting*, pages 97–104, Budapest, Hungary, September 2004.
[21] V. Heuveline and F. Schieweck. H1-interpolation on quadrilateral and hexahedral meshes with hanging nodes. *Computing*, 80(3):203–220, July 2007.
[22] P. Houston, C. Schwab, and E. Süli. Discontinuous *hp*-finite element methods for advection-diffusion-reaction problems. *SIAM J. Numer. Anal.*, 39(6):2133–2163 (electronic), 2002.
[23] T. J. Hughes, G. Engel, L. Mazzei, and M. G. Larson. The continuous galerkin method is locally conservative. *Journal of Computational Physics*, 163(2):467 – 488, 2000.
[24] M. Larson and A. Niklasson. A conservative flux for the continuous galerkin method based on discontinuous enrichment. *CALCOLO*, 41(2):65–76, 2004.
[25] B. Rivière, M. F. Wheeler, and V. Girault. Improved energy estimates for interior penalty, constrained and discontinuous Galerkin methods for elliptic problems. I. *Comput. Geosci.*, 3(3-4):337–360 (2000), 1999.
[26] B. Rivière, M. F. Wheeler, and V. Girault. A priori error estimates for finite element methods based on discontinuous approximation spaces for elliptic problems. *SIAM J. Numer. Anal.*, 39(3):902–931 (electronic), 2001.
[27] R. Stenberg. Mortaring by a method of J. A. Nitsche. In *Computational mechanics (Buenos Aires, 1998)*, pages CD–ROM file. Centro Internac. Métodos Numér. Ing., Barcelona, 1998.
[28] S. Sun and J. Liu. A Locally Conservative Finite Element Method based on piecewise constant enrichment of the continuous Galerkin method. *SIAM J. Sci. Comput.*, 31:2528–2548, 2009.
[29] S. Sun and M. F. Wheeler. Discontinuous galerkin methods for coupled flow and reactive transport problems. *Applied Numerical Mathematics*, 52(23):273 – 298, 2005. {ADAPT} '03: Conference on Adaptive Methods for Partial Differential Equations and Large-Scale Computation.
[30] S. Sun and M. F. Wheeler. Symmetric and nonsymmetric discontinuous Galerkin methods for reactive transport in porous media. *SIAM J. Numer. Anal.*, 43(1):195–219 (electronic), 2005.
[31] S. Sun and M. F. Wheeler. Symmetric and nonsymmetric discontinuous galerkin methods for reactive transport in porous media. *SIAM Journal on Numerical Analysis*, 43(1):195–219, 2005.
[32] S. Sun and M. F. Wheeler. Anisotropic and dynamic mesh adaptation for discontinuous galerkin methods applied to reactive transport. *Computer Methods in Applied Mechanics and Engineering*, 195(2528):3382 – 3405, 2006. Discontinuous Galerkin Methods.
[33] S. Sun and M. F. Wheeler. Projections of velocity data for the compatibility with transport. *Computer Methods in Applied Mechanics and Engineering*, 195(78):653 – 673, 2006.
[34] M. Wheeler and I. Yotov. A multipoint flux mixed finite element method. *SIAM J. Numer. Anal.*, 44(5):2082–2106, 2006.
[35] M. F. Wheeler. An elliptic collocation-finite element method with interior penalties. *SIAM J. Numer. Anal.*, 15(1):152–161, 1978.
[36] H. Wu and Y. Xiao. An unfitted *hp*-interface penalty finite element method for elliptic interface problems. *arXiv preprint arXiv:1007.2893*, 2010.
[37] J. Xu. The auxiliary space method and optimal multigrid preconditioning techniques for unstructured meshes. *Computing*, 56:215–235, 1996.
[38] J. Xu. Optimal algorithms for discretized partial differential equations. In *ICIAM 07–6th International Congress on Industrial and Applied Mathematics*, pages 409–444, 2009.
[39] J. Xu, L. Zikatanov, C. Zhang, and M. others. *FASP Auxiliary Space Preconditioning*, 2015.
[40] Q. Xu and J. S. Hesthaven. Discontinuous galerkin method for fractional convection-diffusion equations. *SIAM Journal on Numerical Analysis*, 52(1):405–423, 2014.