Superconvergence property of an over-penalized discontinuous Galerkin finite element gradient recovery method

Lunji Song a, *, Zhimin Zhang b, c

a School of Mathematics and Statistics, and Key Laboratory of Applied Mathematics and Complex Systems (Gansu Province), Lanzhou University, Lanzhou 730000, PR China
b Beijing Computational Science Research Center, Beijing 100084, PR China
c Department of Mathematics, Wayne State University, Detroit, MI 48202, USA

ARTICLE INFO

Article history:
Received 10 February 2015
Received in revised form 30 June 2015
Accepted 17 July 2015
Available online 28 July 2015

Keywords:
Discontinuous Galerkin method
Polynomial preserving recovery
Superconvergence
Gradient recovery

ABSTRACT

A polynomial preserving recovery method is introduced for over-penalized symmetric interior penalty discontinuous Galerkin solutions to a quasi-linear elliptic problem. As a post-processing method, the polynomial preserving recovery is superconvergent for the linear and quadratic elements under specified meshes in the regular and chevron patterns, as well as general meshes satisfying Condition (ε, σ). By means of the averaging technique, we prove the polynomial preserving recovery method for averaged solutions is superconvergent, satisfying similar estimates as those for conforming finite element methods. We deduce superconvergence of the recovered gradient directly from discontinuous solutions and naturally construct an a posteriori error estimator. Consequently, the a posteriori error estimator based on the recovered gradient is asymptotically exact. Extensive numerical results consistent with our analysis are presented.

© 2015 Elsevier Inc. All rights reserved.

1. Introduction

In recent years there have been superconvergence results of the gradient and gradient recovery schemes [9,12,20,23,31–33,36–38], while their contributions are based on finite element approximations and attract many researchers from the fields of modern engineering and scientific computation. The Zienkiewicz–Zhu (ZZ) error estimator [38] is referred to the superconvergence patch recovery (SPR), which is based on gradient recovery from the gradient of the finite element solution on patches in the discrete least-squares fitting sense. The robustness of the ZZ patch recovery is originated from its superconvergence under structured meshes. As a new strategy, polynomial preserving recovery (PPR) has first been introduced by Zhang in [24,34] with the use of the fitted finite element solution values to recover the gradient. The PPR keeps all known superconvergence properties of the ZZ patch recovery, out-performing the SPR in the cases of quadratic element at edge centers and linear element for the chevron mesh [36]. The PPR has superconvergence in mildly structured grids as well as anisotropic grids [34,35]. Therefore, for gradient recovery to finite element solutions, the PPR method is a good alternative.

* Corresponding author.
E-mail addresses: song@lzu.edu.cn (L. Song), zhangmath.wayne.edu (Z. Zhang).

http://dx.doi.org/10.1016/j.jcp.2015.07.036
0021-9991/© 2015 Elsevier Inc. All rights reserved.
It is well known that if the recovered quantity better approximates the exact one, then it can be used in constructing asymptotically exact \textit{a posteriori} error estimates (see [1]). The PPR becomes standard in finite element methods and has been adopted in some commercial softwares (COMSOL etc.) as a superconvergence tool.

We consider the following second-order quasi-linear elliptic problem

\[
\begin{aligned}
- \nabla \cdot (a(x, u) \nabla u) & = f(x), & \text{in } \Omega, \\
\quad u & = 0, & \text{in } \partial \Omega,
\end{aligned}
\]  

(1)

where \( \Omega \) is a bounded convex domain in \( \mathbb{R}^2 \) with a smooth boundary \( \partial \Omega \), and \( n \) is the unit outward normal vector to \( \partial \Omega \). We assume \( 0 < a_1 \leq a(x, v) \leq a_2 \), \( x \in \Omega, v \in \mathbb{R} \) for some positive constants \( a_1, a_2 \) and \( a(x, v) \in C_2^2(\Omega \times \mathbb{R}) \), where \( C_2^2(\Omega \times \mathbb{R}) \) is the space of twice continuously differentiable functions on \( \Omega \) whose first and second order derivatives are bounded in \( \Omega \times \mathbb{R} \). It holds from [15] that there exists a unique weak solution \( u \) to (1) and \( u \in C^{2+\delta} \) with \( \delta \in (0, 1) \) when \( f \in C^2(\Omega) \) and the boundary \( \partial \Omega \) is smooth. The equation (1), supplemented with the homogeneous Dirichlet boundary condition, describes an equilibrium state of a chemical species of the concentration \( u \) in a porous medium with a source term \( f(x) \).

Interior penalty discontinuous Galerkin (IPDG) methods are a powerful simulation tool for solving linear or nonlinear equations (see e.g. [4,11,14,17,18,22,25,28]). There are some primal DG versions belonged to IPDG methods (see [3,19]), such as symmetric interior penalty Galerkin (SIPG), nonsymmetric interior penalty Galerkin (NIPG), incomplete interior penalty Galerkin (IIPG) as well as its corresponding over-penalized interior penalty methods. We are interested in an over-penalized symmetric interior penalty Galerkin (OPSIPG) method [27,30] to realize a gradient recovery. One of reasons is that its penalty parameters can be bounded above rather than sufficiently large in the usual SIPG method for refined grids. The OPSIPG method we use preserves the integral terms on hybrid multiplication of jump or average of test and trial functions on interior edges. The weakly over-penalized symmetric interior penalty method (WOPSIPG) presented by Brenner in [6] ignores the hybrid multiplication terms. The OPSIPG and WOPSIPG methods produce an ill-conditioned discrete system, which results from the over-penalization terms. Fortunately, it can be remedied by a simple block-diagonal preconditioner (see [6]) and a multilevel preconditioner in [8]. Now the main question lies in how to implement the PPR technique into discontinuous Galerkin solutions under the framework of discontinuous Galerkin finite element methods.

In the present work, we aim to the PPR technique based on discontinuous Galerkin solutions and its theoretical analysis. The PPR technique is good for arbitrary order Lagrange finite elements, then for simplicity, we would focus on the linear and quadratic elements, which are widely used in practice. Several steps for the PPR are needed: we choose a patch including necessary or enough points first, and then by the fitted solution values recover the gradient, and further construct an \textit{a posteriori} error estimate in the energy norm. Due to the PPR partial to the symmetry of patches, we shall consider some specified meshes in the regular and chevron patterns, as well as general meshes satisfying Condition (\( \epsilon, \sigma \)). In case that the resultant DG system becomes ill-conditioned from over-penalized parameters, it is important to use a simple block-diagonal preconditioner remedying the problem. To the best of our knowledge, this is the first theoretical superconvergence proof for the PPR implemented on discontinuous Galerkin solutions to nonlinear elliptic problems. Furthermore, the proposed method can be used to solve time-dependent diffusion problems, e.g., the problem discussed in [13]. Our method can be applied to the spatial discretization part at each time level while maintaining the time discretization part unchanged.

The remainder of this paper is organized as follows. In Section 2 we introduce the over-penalized interior penalty discontinuous Galerkin (OPIPDG) formulas in the broken Sobolev space to approximate elliptic equations. In Section 3, we state and prove some preliminary lemmas for OPSIPG scheme analogous to those appear in the usual SIPG method. In Section 4, the gradient recovery operator will be constructed for OPSIPG solutions thereafter we prove the main results for the gradient recovery, which can be used to define an \textit{a posteriori} error estimator. In the last section, several numerical examples are given to illustrate superconvergence of the gradient recovery for linear and quadratic elements in some structured meshes as well as unstructured meshes, and also show that an \textit{a posteriori} error estimator is asymptotically exact for a conner singularity problem.

2. The over-penalized discontinuous Galerkin method

Let \( \mathcal{E}_h \) be a subdivision of \( \Omega \) into disjoint open elements such that \( \Omega = \bigcup_{i=1}^{N_h} \bar{E}_i \), where \( E_i \) is a triangle in 2D and \( N_h \) is the number of all elements. We let \( h_k := \text{diam}(\bar{E}_k) \) and \( h := \max_{E \in \mathcal{E}_h} h_k \). It is assumed that the family of subdivisions \( \mathcal{E}_h \) is shape regular and each element \( E \in \mathcal{E}_h \) shall be an affine image of a standard reference element. Assume that the mesh \( \mathcal{E}_h \) is quasi-uniform: for all \( E_i \in \mathcal{E}_h \), there exists a constant \( \tau > 0 \), independent of \( h \), such that

\[
h \leq \tau h_i,
\]

(2)

where \( \rho_i \) denotes the diameter of the largest circle inscribed in \( E_i \). We introduce the set of all edges of the mesh \( \mathcal{E}_h \) by

\[
\mathcal{F}_h := \{ e_1, e_2, \cdots, e_{P_h}, e_{P_h+1}, \cdots, e_{M_h} \},
\]

where

\[
\begin{aligned}
\quad e_i & \subset \Omega, \quad & \text{if } 1 \leq i \leq P_h, \\
\quad e_i & \subset \partial \Omega, \quad & \text{if } P_h + 1 \leq i \leq M_h.
\end{aligned}
\]
On each edge $e_i$ ($1 \leq i \leq M_h$) of $\mathcal{E}_h$, we fix a unit outer normal vector $\mathbf{n}_i$:

- the unit normal vector on $e_i$, pointing from $E_k$ to $E_j$, if $e_i = \partial E_k \cap \partial E_j$,
- the unit normal vector on $e_i$, pointing outward of $\Omega$, if $P_h + 1 \leq i \leq M_h$.

then we denote the average on $e_i$ and the jump across $e_i$ below: For $v \in H^1(\mathcal{E}_h)$, $s > \frac{1}{2}$,

$$
\{v\}_{e_i} := \left\{ \begin{array}{ll}
\frac{1}{2}(v|_{E_k})_{|e_i} + \frac{1}{2}(v|_{E_j})_{|e_i}, & \text{if } e_i = \partial E_j \cap \partial E_k, 1 \leq i \leq P_h, \\
(v|_{E_k})_{|e_i}, & \text{if } e_i = \partial E_k \cap \partial \Omega, P_h + 1 \leq i \leq M_h,
\end{array} \right.
\]

$$

$$
[v]_{e_i} := \left\{ \begin{array}{ll}
(v|_{E_k})_{|e_i} - (v|_{E_j})_{|e_i}, & \text{if } e_i = \partial E_j \cap \partial E_k, 1 \leq i \leq P_h, \\
(v|_{E_k})_{|e_i}, & \text{if } e_i = \partial E_k \cap \partial \Omega, P_h + 1 \leq i \leq M_h.
\end{array} \right.
\]

For brevity, we drop the subscript $e_i$ of these two operators throughout this paper. And a definition for general meshes is given as follows.

**Definition 2.1.** Triangulation $\mathcal{E}_h$ satisfies Condition ($\epsilon$, $\sigma$) if $\mathcal{E}_h$ can be separated into two parts $\mathcal{E}_h = \mathcal{E}_{0,h} \cup \mathcal{E}_{1,h}$, $\bigcup_{K \in \mathcal{E}_{1,h}} \tilde{K} = \Omega_{1,h}$, $\Omega = \Omega_{0,h} \cup \Omega_{1,h}$, such that the following conditions exist:

1. Any two triangles that share a common edge in $\Omega_{0,h}$ form a convex quadrilateral which is an $\epsilon$-perturbation from a parallelogram.
2. $\Omega_{1,h}$ has a small measure: $|\Omega_{1,h}| \leq 0(h^3), \sigma > 0$.

Throughout the paper, $W^{k,p}(\omega)$ will signify the usual Sobolev space on polygonal Lipschitz domain $\omega$, of differentiability index $k$ and integrability index $p$, equipped with the norm $\|\cdot\|_{W^{k,p}(\omega)}$ and seminorm $|\cdot|_{W^{k,p}(\omega)}$. When $p = 2$, we shall write $W^{2,2}(\omega) = H^2(\omega)$ and suppress the index $p$ in the notation of the norm and seminorm, writing $\|\cdot\|_{k,\omega}$ and $|\cdot|_{k,\omega}$, respectively. Note that, by default, $H^0(\omega)$ denotes $L^2(\omega)$ with the $L^2$ inner product $\langle \cdot, \cdot \rangle$ and the standard $L^2$-norm $\|\cdot\|_{L^2(\omega)}$, or equivalently, the standard $L^\infty$-norm $\|\cdot\|_{\infty, \omega}$ on $\omega$. Then we assign to the subdivision $\mathcal{E}_h$ the broken Sobolev space of real orders $s$:

$$
H^s(\mathcal{E}_h) = \left\{ v \in L^2(\Omega) : v|_{E_i} \in H^s(E_i), 1 \leq i \leq N_h \right\},
$$

equipped with the broken Sobolev space norm,

$$
\|v\|_s := \left( \sum_{i=1}^{N_h} \|v\|^2_{L^2(\Omega)} \right)^{\frac{1}{2}}.
$$

We also introduce a space of test functions

$$
\mathcal{D}_r(\mathcal{E}_h) = \left\{ v \in L^2(\Omega) : v|_{E_i} \in \mathbb{P}_r(E_i), 1 \leq r \leq 2, 1 \leq i \leq N_h \right\},
$$

where $\mathbb{P}_r(E_i)$ denotes the set on $E_i$ of all polynomials of (total) degree at most $r$ on $E_i$. It is clear that $\mathcal{D}_r(\mathcal{E}_h) \subseteq H^1(\Omega)$. We thus introduce the interior super penalty term $J^\beta_r : \mathcal{D}_r(\mathcal{E}_h) \times \mathcal{D}_r(\mathcal{E}_h) \to \mathbb{R}$ in the form:

$$
J^\beta_r(v, w) = \sum_{k=1}^{M_h} \frac{\sigma_k}{|e_k|^2} \int_{e_k} |v| |w| ds, \quad \beta > 1,
$$

which penalizes the jump of the functions across the edges $e_k$, $1 \leq k \leq M_h$. Here the penalty parameter $\sigma_k$ is a nonnegative real number to be chosen and $|e_k|$ is the Lebesgue measure of the edge $e_k$. We also define the energy norm on $\mathcal{D}_r(\mathcal{E}_h)$ throughout this paper:

$$
\|v\|_{DG} = \left( \sum_{k=1}^{N_h} \|\nabla v\|^2_{L^2(E_k)} + J^\beta_r(v, v) \right)^{\frac{1}{2}}, \quad \forall v \in \mathcal{D}_r(\mathcal{E}_h).
$$

Aiming to study the strong solution $u \in C^2(\overline{\Omega} \times [0, T])$, satisfying the regularity conditions, of the problem (1), we proceed element by element as appears in [27]. The weak formulation of the problem (1) reads as follows: Find $u \in H^s(\mathcal{E}_h)$, $s \geq 2$, such that

$$
A_2(u; u, v) = \langle f, v \rangle, \quad \forall v \in H^s(\mathcal{E}_h),
$$

where $A_2(\rho; v, w)$ is bilinear in the last two terms:
\[ a_j(\rho; \nu, w) = \sum_{j=1}^{N_h} \int a(\rho)\nabla \nu \cdot \nabla w \, dx - \sum_{k=1}^{M_h} \int (a(\rho)\nabla \nu \cdot n_k)[w] \, ds \]
\[ + \lambda \sum_{k=1}^{M_h} \int (a(\rho)\nabla w \cdot n_k)[v] \, ds + \int_{\Omega_1} f_j^w(v, w), \quad \nu, w \in H^2(\mathcal{E}_h). \]  

(7)

The OPIDPG approximation of (1) is to find \( u_h \in \mathcal{D}_r(\mathcal{E}_h) \) such that
\[ A_j(u_h; u_h, v_h) = (f, v_h), \quad \forall v_h \in \mathcal{D}_r(\mathcal{E}_h). \]  

(8)

Moreover, depending on the coefficient \( \lambda \), the discontinuous Galerkin method (8) is referred to OPSIPG if \( \lambda = -1 \); OPNIPG if \( \lambda = 1 \); or OPIPG if \( \lambda = 0 \). In the rest of the paper, we will use the parameter \( \lambda = -1 \) in \( A_j \). For the choice of penalty parameters \( \sigma_k \) of these discontinuous formulations, the reader is referred to Georgoulis and Süli [16], Rivi ère et al. [28], Song [29], and the references therein. By analogous arguments as in [18] and Brouwer’s fixed theorem, the consistence of (6) and the existence and uniqueness of the OPSIPG scheme (8) can be shown.

3. Some estimates of the OPSIPG scheme

In this section, before embarking on gradient recovery and its superconvergence analysis, we state and prove some preliminary results of (8) with \( \lambda = -1 \). And we denote by \( C \) a generic positive constant. From (8), the OPSIPG approximation of \( u \) is to find \( u_h \in \mathcal{D}_r(\mathcal{E}_h) \) such that
\[ A_{-1}(u_h; u_h, v_h) = (f, v_h), \quad \forall v_h \in \mathcal{D}_r(\mathcal{E}_h). \]  

(9)

There exists a positive constant \( C_{\Omega, \tau} \) depending on \( \Omega \) and \( \tau \) in [29, Lemma 3.4] such that
\[ \sum_{k=1}^{M_h} \| e_k \|_{0, e_k}^2 \leq C_{\Omega, \tau} \| \nabla v \|_{0, \Omega}^2, \quad \forall v \in H^1(\mathcal{E}_h). \]  

(10)

Analogously to the proof of Lemma 3.5 in [29], we note that \( A_{-1} \) is coercive. Indeed, if a positive constant \( \delta \) is such that
\[ \frac{a_1}{C_{\Omega, \tau}} > \delta > \min_k \left\{ \frac{\sigma_k^2}{|e_k|^2} \right\} \]  

and the penalty parameters \( \sigma_k \) satisfy \( \sigma_k > C_{\Omega, \tau} \frac{a_1}{|e_k|^2} \) for any \( v, \rho \in \mathcal{D}_r(\mathcal{E}_h) \),
\[ A_{-1}(\rho; v, v) \geq \alpha_0 \left( \| \nabla v \|_{0, \Omega}^2 + \int_{\Omega_1} f_j^w(v, v) \right), \]  

(11)

where \( \alpha_0 = \min_k \left\{ \frac{a_1 - C_{\Omega, \tau} \delta}{\sigma_k} \right\} \). Thus, in OPSIPG scheme, the penalty parameters \( \sigma_k \) may be small quantities on edges, due to the ratio \( \frac{a_1}{\sigma_k} \) or \( \frac{\sigma_k^2}{|e_k|^2} \) multiplying with a small factor \( |e_k|^2 \). The choice of penalty parameters is significantly different from that in the usual SIPG scheme. For simplicity, we may take small constant penalty parameters for OPSIPG scheme. Based on the estimate (11), then we have the following lemma for (8).

**Lemma 3.1.** Assume that \( \frac{a_1}{C_{\Omega, \tau}} > \delta > \min_k \left\{ \frac{\sigma_k^2}{|e_k|^2} \right\} \) and \( \{ \sigma_k \} > C_{\Omega, \tau} \frac{a_1}{|e_k|^2} \) for any \( v, \rho \in \mathcal{D}_r(\mathcal{E}_h) \), then it holds that
\[ A_{-1}(\rho; v, v) \geq \alpha_0 \| \nabla v \|_{0, \Omega}^2, \quad \forall \rho, v \in \mathcal{D}_r(\mathcal{E}_h). \]  

(12)

The following continuity lemma can be proven similarly as in [29].

**Lemma 3.2.** There exists a constant \( \beta_0 > 0 \) such that
\[ A_{-1}(\rho; v, u) \leq \beta_0 \| v \|_{DG} \| u \|_{DG}, \quad \forall \rho, \nu, \omega \in \mathcal{D}_r(\mathcal{E}_h). \]  

(13)

By Brouwer’s fixed point theorem, the existence of discrete solution of SIPG scheme has been shown in [18] for sufficiently small \( h \), so does OPSIPG scheme. We need the following a priori error estimates for solutions of OPSIPG scheme (9) in the broken \( H^1 \) and \( L^2 \) norms, which can be derived analogously from the proofs in [18, Theorems 4.9 and 4.10] and [7, Theorem 2.6].

**Lemma 3.3.** Let \( u \in W^{1, \infty}(\Omega) \cap H^{1+\epsilon}(\Omega) \) and \( u_h \in \mathcal{D}_r(\mathcal{E}_h) \) be the solutions of (6) and (9), respectively. For sufficiently small \( h \) and for any \( \beta \geq 1 \), there exists a positive constant \( C \) independent of \( h \) such that
\[ \| u - u_h \|_{DG} \leq Ch^\beta \| v \|_{\beta+1}, \quad \| u - u_h \|_{L^2(\mathcal{E}_h)} \leq Ch^{\beta+1} \| v \|_{\beta+1}. \]  

(14)
The Aubin–Nitsche lifting technique is well suited to the analysis of the DG method for linear problems, since OPSIPG scheme is symmetric. But for the quasilinear elliptic equation, we will use the nonlinear elliptic projection \( \pi_h : H^2(\Omega) \cap H^1_0(\Omega) \to \mathcal{D}_h(\mathcal{E}_h) \) of \( u \) defined by
\[
A_{-1}(u; u, v_h) = A_{-1}(u; \pi_h u, v_h), \quad \forall v_h \in \mathcal{D}_h(\mathcal{E}_h).
\]
(15)

For any \( u \in H^{r+1}(\Omega) \cap H^1_0(\Omega) \), the Lax–Milgram Theorem leads to the uniqueness and existence of \( \pi_h u \) in the linear variational form (15). By a minor modification of Theorem 4.5 in [2], we recover an approximation property of \( \pi_h u \) in the following lemma.

**Lemma 3.4.** There exists a positive constant \( C \) independent of \( h \) such that
\[
\| u - \pi_h u \|_{L^2(\mathcal{E}_h)} + h\| u - \pi_h u \|_{DG} \leq C(u)h^{r+1},
\]
for \( u \in H^{r+1}(\Omega) \cap H^1_0(\Omega) \).

We apply analogously the proof of superconvergence of an elliptic projection in [4, Theorem A.1] to the projection \( \pi_h \) in the broken \( H^1 \)-norm, then the following lemma still holds.

**Lemma 3.5.** Let \( u \in W^{r+1, \infty}(\Omega) \) and \( u_h \in \mathcal{D}_h(\mathcal{E}_h) \) be the solution of (6) and (8), respectively. For sufficiently small \( h \), there exists a positive constant \( C \) independent of \( h \) such that
\[
\| u_h - \pi_h u \|_{DG} \leq C(u)h^{r+1}.
\]

Let \( V_h \) be a space of continuous piecewise polynomials of order \( r \) \((r = 1, 2)\) associated with \( \mathcal{E}_h \). Let \( z \) be any interior node of the Lagrange finite element space \( P_r \), associated with triangulation \( \mathcal{E}_h \) and let \( \mathcal{T}_z \) be the set of the triangles in \( \mathcal{E}_h \) that share the node \( z \) in their closures. A linear map \( \mathcal{R} : \mathcal{D}_h(\mathcal{E}_h) \to V_h \) can be constructed by averaging techniques [5,17], i.e., for \( v \in \mathcal{D}_h(\mathcal{E}_h) \),
\[
\mathcal{R}v = \frac{1}{|\mathcal{T}_z|} \sum_{E \in \mathcal{E}_h} v|_{E}(z),
\]
where \(|\mathcal{T}_z|\) is the cardinality of triangles in \( \mathcal{T}_z \). For all boundary nodes \( z \), we set \( \mathcal{R}v(z) = 0 \). For uniformly simplicially reducible meshes or quasiuniform quadrilateral meshes obtained by hierarchical refinement, the following result provides a link between discontinuous piecewise polynomial functions and functions in \( H^1_0(\Omega) \) (see [25]).

**Lemma 3.6.** There exists a constant \( C \), independent of \( h \), and a linear operator \( \mathcal{R} : \mathcal{D}_h(\mathcal{E}_h) \to H^1_0(\Omega) \) such that, for all \( u_h \in \mathcal{D}_h(\mathcal{E}_h) \) and \( s \in [0, 1] \),
\[
\| \nabla^s(u_h - \mathcal{R}u_h) \|_{L^2(\Omega)} \leq Ch^{\frac{1}{2} - s}\left( \sum_{e \in \Gamma_{inj\cup \partial \Omega}} \left\| u_h \right\|_{L^2(e)} \right)^{\frac{1}{2}},
\]
where \( \nabla^0 = id \) and \( \nabla^1 = \nabla \).

4. The gradient recovery operator

4.1. Construction of DG gradient recovery and its primal properties

We introduce a gradient recovery operator \( G_h : \mathcal{D}_h(\mathcal{T}_h) \to V_h \times V_h \). For any \( r \)-th order finite elements, all we need is to define \( G_h u_h \) at each node \( z_i \) of the triangulation \( \mathcal{E}_h \):
\[
G_h u_h (z_i) = \sum_j C_{ij} u_h (z_{ij}), \quad \sum_j C_{ij} = 0,
\]
where \( C_{ij} \) are coefficients of some finite difference schemes and \( |z_{ij}| \) are all local points on a patch of elements around \( z_i \). In some special situations, we are referred to [36] for the choices of \( C_{ij} \). It means that the recovered gradient at \( z_i \) can be regarded as a linear combination of neighbor values of discontinuous finite element solutions [30].

First, we present definition of the PPR for discontinuous approximations. By \( \mathcal{N}_h \) we denote the set of all interior nodes in \( \mathcal{E}_h \). Let \( z \in \mathcal{N}_h \) be a mesh vertex and let \( \mathcal{E}_{h,z} \) denote a patch of mesh elements around \( z \). Set \( p_z \in P_{r+1} \) be the polynomial that best fits discontinuous solutions \( u_h \) at the mesh nodes in \( \mathcal{E}_{h,z} \) in the local discrete least-squares sense:
\[
\sum_{z_i \in \mathcal{N}_h \cap \mathcal{E}_{h,z}} |(u_h - p_z)(z_i)|^2 = \min_{p \in P_{r+1}(\mathcal{E}_{h,z})} \sum_{z_i \in \mathcal{N}_h \cap \mathcal{E}_{h,z}} |(u_h - p)(z_i)|^2.
\]

(20)
Here \( p_z \) is called the least-squares polynomial approximation (LSPA) of \( u_h \) at \( z \). Then it is well defined by

\[
(G_h u_h)(z) \triangleq \nabla p_z(z).
\]

We denote by \( N_z \) the number of mesh nodes in the patch \( E_{h,z} \). For an internal mesh vertex \( z \) and an order \( r \), there are \( N_z \) \((\geq m := \frac{(r+2)(r+3)}{2})\) points required in an element patch \( E_{h,z} \) including the mesh vertex \( z \) itself. To fit a polynomial of degree \( r + 1 \), in the least-squares sense, we select points distributed around \( z \) on the ball \( B_h(z) = \{ x \in E_h : |x - z| \leq h \} \). If the number of points (including \( z \)) is less than \( m \), we search further points and proceed this process on a larger circle until more than or identical to \( m \) points being chosen. Then the patch \( E_{h,z} \) is well defined and must have at least \( m \) points distributed around \( z \) in a way that leads to a unique \( p_z \). Next, to define \( E_{h,z} \) at a boundary mesh vertex \( z \), we set

\[
E_{h,z} \triangleq E_{h,z_0} \cup L_{z,n_0},
\]

where \( z_0 \) is the closest internal vertex to \( z \) and \( L_{z,n_0} \) is the union of mesh elements in the first \( n_0 \in \mathbb{Z}^+ \) layers around \( z \) including the internal mesh vertex \( z_0 \). This definition ensures the uniqueness of \( p_z \) as shown in Lemma 3.6 of [36].

Let \( h_z \) be the length of the longest edge attached to \( z \). Taking the local coordinates \((x, y)\) with \( z \) as the origin, the fitting polynomial is

\[
p_{r+1}(x, y; z) = P^T a = (1, x, y, \ldots, x^{r+1}, y^{r+1}, y^2, \ldots, y^{r+1}) a
\]

with \( a^T = (a_1, a_2, \ldots, a_m) \). With a scaling augmeny by \( h = h_z \), set

\[
\hat{p}_{r+1}(\xi, \eta) = \hat{P}^T \hat{a},
\]

the fitting polynomial becomes

\[
\hat{p}_{r+1}(\xi, \eta) = \hat{P}^T \hat{a},
\]

where \( \hat{a}^T = (a_1, h a_2, h a_3, \ldots, h^{k+1} a_m) \). The coefficient vector \( \hat{a} \) is determined by the linear system

\[
A^T \hat{a} = A^T b,
\]

where \( b^T = (u_h(z_1), u_h(z_2), \ldots, u_h(z_m)) \) and

\[
A = \begin{pmatrix}
1 & \xi_1 & \eta_1 & \cdots & \eta_1^{r+1} \\
1 & \xi_2 & \eta_2 & \cdots & \eta_2^{r+1} \\
\vdots & \vdots & \vdots & \ddots & \vdots \\
1 & \xi_n & \eta_n & \cdots & \eta_n^{r+1}
\end{pmatrix}.
\]

The uniqueness condition for the linear system (22) is \( \text{Rank}(A) = m \), so it holds when \( n \leq m \) and all the \( n \) sampling points are not on the same conic curve for the linear element [24]. For linear or quadratic discontinuous/continuous finite elements in the regular and chevron patterns, the reader is referred to [30,36] about the weights of recover gradient on vertices, edge nodes and internal nodes for more details.

Especially, if one takes the average of the discontinuous Galerkin solution at each node, i.e., for any a node \( z \in N_h \), the average \( \bar{u}_h(z) := \mathcal{R} u_h(z) \), then \( G_h \bar{u}_h(z) \) is the same as in [36]. It is well known that \( G_h \bar{u}_h \) has some important properties:

1. \( |G_h \bar{u}_h|_{L^2(\Omega)} \leq C|\nabla \bar{u}_h|_{L^2(\Omega)} \), \( \forall \bar{u}_h \in V_h \);
2. If \( p \in \mathbb{P}_{r+1}(\omega_{k2}) \), then \( (G_h p)(z) = \nabla p(z), \forall v \in N_h \).

where \( p(z) \) is the Lagrangian interpolation of \( p, \omega_{k2} \) is an element patch associated with node \( z \) and \( \omega_{k2} \) denotes the union of element patches associated with three vertices of element \( r \). We have

\[
|\nabla u - G_h u_i|_{L^2(\Omega)} \leq Ch^{r+1} |u|_{r+2, \Omega}, \quad r = 1, 2.
\]

We need the following boundedness assumption on \( G_h \): When there are no two adjacent angles on an element patch adding up to exceed \( \pi \), it is assumed that

\[
|G_h v|_{L^2(E_h)} \leq \|v\|_{DG}, \quad \forall v \in D_t(E_h).
\]

Remark 1. As a post-processing technique, PPR has been implemented by COMSOL Multiphysics [39,40] for any order finite elements. Therefore, PPR for higher-order elements can also be applied to problems with corner singularities similar to what we have done for linear element. However, theoretical analysis for higher-order elements would be much involved.

The difference between \( G_h u_h \) and \( G_h \bar{u}_h \) has been presented in the following lemma.

Lemma 4.1. Consider an interior patch \( \omega_2 \subset \subset \Omega_{2,h} \subset \Omega \). It holds that

\[
|G_h (u_h - \bar{u}_h)|_{W^{r-\beta,p}(\Omega_{2,h})} \lesssim |\ln h| h^{r+\beta-1} \|u\|_{W^{r+1,\infty}(\Omega)},
\]

(25)
Proof. Setting \( \omega_2 \) define a subdomain including a patch with a node \( z \) as the center and a vector function \( \mathbf{w} \in C_0^\infty(\omega_2)^2 \) with \( \|\mathbf{w}\|_{H^1(\omega_2)} = 1 \), we have

\[
(G_h(u_h - \bar{u}_h), \mathbf{w}) = (u_h - \bar{u}_h, G_h^* \mathbf{w}) \\
\leq \|u_h - \bar{u}_h\|_{\infty,(\omega_2 + h)} \|G_h^* \mathbf{w}\|_{L^1(\omega_2 + h)} \\
\leq \|u_h - \bar{u}_h\|_{\infty,(\omega_2 + h)},
\]

where \( \omega_2 + h \) defines a subdomain that stretches out \( h \) from \( \omega_2 \) and \( \|G_h^* \mathbf{w}\|_{L^1(\omega_2 + h)} \) and \( \|\mathbf{w}\|_{H^1(\omega_2 + h)} \) are bounded uniformly with respect to \( h \). From the discrete Sobolev inequality [5] for piecewise polynomial functions \( w_h \) with respect to \( \mathcal{E}_h \), we have

\[
\|w_h\|_{L^\infty(\Omega)}^2 \leq (1 + |\ln h|) \left( \sum_{E \in \mathcal{E}_h} \|w_h\|_{H^1(E)}^2 + \sum_{e \in \Gamma_{	ext{int}}} |e|^{-1} \|w_h\|_{L^2(e)}^2 \right).
\]

Taking \( w_h = u_h - \bar{u}_h \) leads to

\[
\|u_h - \bar{u}_h\|_{L^\infty(\Omega)}^2 \leq (1 + |\ln h|) \left( \sum_{E \in \mathcal{E}_h} \|u_h - \bar{u}_h\|_{H^1(E)}^2 + \sum_{e \in \Gamma_{	ext{int}}} |e|^{-1} \|u_h\|_{L^2(e)}^2 \right),
\]

which gives by Lemma 3.6 and the continuity of \( u \)

\[
\|u_h - \bar{u}_h\|_{L^\infty(\Omega)}^2 \lesssim |\ln h| \int_{\Gamma_{\text{int}} \cup \partial \Omega} h^{-1}|u_h|^2\,ds \\
\lesssim |\ln h| \int_{\Gamma_{\text{int}} \cup \partial \Omega} h^{-1}|u - u_h|^2\,ds \\
\lesssim |\ln h| h^{\beta - 1} \|u - u_h\|^2_{DG} \\
\lesssim |\ln h|h^{r+\beta - 1} \|u\|^2_{W^{r+1,\infty}(\Omega)}. \quad \square
\]

4.2. The main results for the DG gradient recovery

The analysis of averaging techniques plays an important role in a posteriori error control in [10]. Due to its importance, we shall first prove the superconvergence between the gradient and the recovery gradient from the averaged values of OPSIPG solutions under reasonable smoothness conditions. We decompose

\[
\nabla u - G_h \bar{u}_h = (\nabla u - G_h u_I) + G_h (u_I - \bar{u}_h),
\]

where \( u_I \in V_h \) is the interpolation of \( u \) and \( V_h \subset H^1(\Omega) \) represents the \( C^0 \) finite element space consisting of piecewise polynomials of order \( r \) on \( \mathcal{E}_h \).

First, applying (23) to the first term on the RHS of (28) gives

\[
\|\nabla u - G_h u_I\|_{L^2(\Omega)} \lesssim h^{r+1} |u|_{r+2, \Omega}.
\]

To bound the second term on the right hand side of (28), it follows from the triangle inequality

\[
\|G_h (u_I - \bar{u}_h)\|_{L^2(\mathcal{E}_h)} \lesssim \|\nabla (u_I - \pi_h u)\|_{L^2(\mathcal{E}_h)} + \|\nabla (\pi_h u - \bar{u}_h)\|_{L^2(\mathcal{E}_h)},
\]

where the boundedness of \( G_h \) has been used. To estimate the RHS terms of (30), we will apply the following critical lemma.

Lemma 4.2. Let \( u \in W^{r+1,\infty}(\Omega) \cap H^{r+2}(\Omega) \cap H^1_0(\Omega) \) and \( u_I \) be the solution of (6) and the Lagrange interpolation of \( u \) in \( V_h \), respectively. Assume that \( \mathcal{E}_h \) satisfies Condition \((\epsilon, \sigma)\), the following estimates hold.

(1) If \( r = 1 \), it holds that

\[
\sum_{j=1}^{N_h} \int_{\mathcal{E}_j} \nabla e_I \cdot \nabla v_h \, dx \lesssim E_{\beta}^1 \|v_h\|_{DG}, \quad \forall v_h \in D_1(\mathcal{E}_h)
\]

where

\[
E_{\beta}^1 := h^{3/2}(\|u\|_{H^3(\Omega_0, h)} + \epsilon) + h \epsilon (\|u\|_{H^2(\Omega_0, h)} + \epsilon) + h^{1+\sigma/2} |u|_{W^{2, \infty}(\Omega)} \\
+ h^{\beta+1/2} |u|_{W^{2, \infty}(\Omega)}.
\]
(2) If \( r = 2 \), it holds that
\[
\sum_{j=1}^{N_h} \int_{E_j} \nabla e_1 \cdot \nabla v_h \, dx \lesssim E_{2h}^2 \| v_h \|_{DG}, \quad \forall v_h \in \mathcal{D}_2(E_h)
\]
where
\[
E_{2h}^2 := h^{5/2}(\| u \|_{H^3(\Omega_{0,h})} + \epsilon) + h^2\epsilon(\| u \|_{H^3(\Omega_{0,h})} + \epsilon) + h^{2+\sigma/2} |u|_{W^{3,\infty}(\Omega)} + h^{\sigma+3} \| u \|_{W^{3,\infty}(\Omega)}.
\]

**Proof.** Let \( Q_h \) be the \( L^2 \) projection onto \( V_h \) such that
\[
(w_h, Q_h v_h) = (w_h, v_h), \quad \forall v_h \in \mathcal{D}_h(E_h), \quad w_h \in V_h.
\]
Due to \( v_h - Q_h v_h \mid_{L^2(\Omega_h)} \leq \| v_h - \mathcal{R} v_h \|_{L^2(\Omega_h)} \), it holds from Lemma 3.6
\[
\| v_h - Q_h v_h \|_{L^2(\Omega_h)} \lesssim h^{1/2} \left( \sum_{j=1}^{M_h} [v_h]^2_{L^2(\Omega_j)} \right)^{1/2}, \tag{31}
\]
and analogously, by the inverse inequality,
\[
\| \nabla (v_h - Q_h v_h) \|_{L^2(\Omega_h)} \lesssim h^{-1/2} \left( \sum_{j=1}^{M_h} [v_h]^2_{L^2(\Omega_j)} \right)^{1/2}. \tag{32}
\]
Notice that \( v_h \in \mathcal{D}_h(E_h) \) does not apply since it is discontinuous and is not included in the continuous finite element space \( V_h \). It is useful and applicable to analyze the error of the \( L^2 \) projection by a decomposition of \( v_h \) and the inverse inequality. We have
\[
\sum_{j=1}^{N_h} \int_{E_j} \nabla (u_j - u) \cdot \nabla v_h \, dx \\
\leq \sum_{j=1}^{N_h} \int_{E_j} \nabla (u_j - u) \cdot Q_h (\nabla v_h) \, dx \\
+ \sum_{j=1}^{N_h} \int_{E_j} \nabla (u_j - u) \cdot (\nabla v_h - Q_h (\nabla v_h)) \, dx \\
\lesssim \sum_{j=1}^{N_h} \int_{E_j} \nabla (u_j - u) \cdot Q_h (\nabla v_h) \, dx \\
+ h^{r+1/2} \| u \|_{W^{r+1,\infty}(\Omega)} \left( \sum_{j=1}^{M_h} \frac{\sigma_0}{|e_j|^\beta} [v_h]^2_{L^2(\Omega_j)} \right)^{1/2}, \tag{33}
\]
where for the last inequality we have used (32). According to the analysis in [21,35], we can directly use the well-known superconvergence results to estimate the first term in the right-hand side of (33). The case of linear elements \( (r = 1) \) employs these conclusions [35, (2.10) and (2.12)], while the results [21, (2.5) and (2.7)] have been applied to the case of quadratic elements \( (r = 2) \). Therefore, inserting these known results to (33) completes the proof of this lemma. \( \square \)

**Theorem 4.3.** For \( \lambda = -1 \), let \( u \in W^{r+1,\infty}(\Omega) \cap H^{r+2}(\Omega) \cap H^1_0(\Omega) \) and \( u_h \) be the solution of (6) and (8), respectively. Assume that \( \tilde{u}_h \) is a continuous approximation of \( u_h \), i.e., \( \tilde{u}_h = \mathcal{R} u_h \) and \( E_h \) satisfies Condition \((\epsilon, \sigma)\), then there is a constant \( C \) independent of \( h \) and \( r \) such that
\[
\| \nabla u - G_h \tilde{u}_h \|_{L^2(\Omega)} \leq C (h^{r+1} \| u \|_{H^{r+2}(\Omega)} + h^{r+\sigma/2} \| u \|_{H^{r+1}(\Omega)} + E_{2h}^2), \tag{34}
\]
where the terms \( E_{2h}^2 \) are specified as follows.
(1) If \( r = 1 \), we have
\[
E^1_\beta = h^{2/2} (\| u \|_{H^3(\Omega)} + \epsilon) + h \epsilon (\| u \|_{H^2(\Omega)}) + h^{1+\sigma/2} |u|_{W^{2,\infty}(\Omega)}
\]
\[
+ h^{r+\sigma} \| u \|_{W^{2,\infty}(\Omega)}.
\]

(2) If \( r = 2 \), we have
\[
E^2_\beta = h^{5/2} (\| u \|_{H^4(\Omega)} + \epsilon) + h^{2} \epsilon (\| u \|_{H^3(\Omega)} + \epsilon) + h^{2+\sigma/2} |u|_{W^{3,\infty}(\Omega)}
\]
\[
+ h^{r+\sigma} \| u \|_{W^{3,\infty}(\Omega)}.
\]

**Proof.** By the above boundedness property (1) of \( G_h \) and (17), under the assumption that \( u \in W^{r+2,\infty} \cap H^1_0(\Omega) \), we have
\[
\| \nabla (u - \pi_h u) \|_{L^2(\mathcal{E}_h)} = \sum_{E \in \mathcal{E}_h} \| \nabla (u - \pi_h u) \|_{L^2(E)} \leq \| u_1 - \pi_h u \|_{DG}.
\] (35)

Note that
\[
A_{-1}(u; u_1, v_h) = A_1(u; u_1 - \pi_h u, v_h)
\]
\[
= A_{-1}(u; u_1 - \pi_h u, v_h)
\]
where we have used the definition of \( \pi_h \) for the second identity and Lemma 3.2 for the last inequality. Moreover, applying Lemma 3.1 in (36) with \( v_h = u_1 - \pi_h u \) taken also gives
\[
A_{-1}(u; u_1 - \pi_h u, v_h) \geq \alpha_0 \| u_1 - \pi_h u \|_{DG} \| v_h \|_{DG},
\] (36)

therefore, it follows from (36) and (37)
\[
\| u_1 - \pi h u \|_{DG} \leq \frac{A_{-1}(u; u_1 - u, v_h)}{\alpha_0 \| v_h \|_{DG}}.
\] (38)

Let \( \overline{a(u)} \) be a piecewise constant function \( \overline{a(u)} = \frac{1}{|E|} \int_E a(x, u) dx \) on each element \( E \in \mathcal{E}_h \). Then we have \( 0 < a_1 < \overline{a(u)} < a_2 \) on \( E \). It follows from the assumption of \( a(\cdot) \) that
\[
|a(u) - \overline{a(u)}| \leq Ch.
\] (39)

We define \( \tilde{A}_{-1}(u; u_1 - u, v_h) \) replacing the term \( a(u) \) by \( \overline{a(u)} \) in \( A_{-1}(u; u_1 - u, v_h) \), then it is easy to show that
\[
|A_{-1}(u; u_1 - u, v_h) - \tilde{A}_{-1}(u; u_1 - u, v_h)| \lesssim h \| u_1 - u \|_{H^1(\Omega)} \| v_h \|_{DG}.
\] (40)

Therefore, we shift our analysis to \( \tilde{A}_{-1}(u; u_1 - u, v_h) \). According to the formula of \( A_{-1}(u; v, w) \), we can write
\[
\tilde{A}_{-1}(u; u_1 - u, v_h) = \sum_{j=1}^{N_h} \int_{E_j} \overline{a(u)} \nabla (u_1 - u) \cdot \nabla v_h \; dx - \sum_{k=1}^{M_h} \int_{E_k} [\overline{a(u)} \nabla (u_1 - u) \cdot \mathbf{n}_k] [v_h] \; ds
\]
\[
:= I_1 + I_2.
\] (41)

By the trace lemma and Cauchy–Schwarz inequality, \( I_2 \) is bounded
\[
I_2 \leq \left( \frac{C}{\sigma_0} \sum_{k=1}^{M_h} |e_k| \beta |u|_{H^{r+1}(E_k)}^2 \right) \frac{1}{2} J^\sigma_h (v_h, v_h)^{1/2}
\]
\[
\leq \left( \frac{C}{\sigma_0} \sum_{k=1}^{M_h} h^{r+1+\beta} |u|_{H^{r+1}(E_k)}^2 \right) \frac{1}{2} J^\sigma_h (v_h, v_h)^{1/2}
\]
\[
\lesssim h^{r+1+\beta} \| u \|_{H^{r+1}(\Omega)} J^\sigma_h (v_h, v_h)^{1/2},
\] (42)

where \( \sigma_0 \) means the maximum penalty parameter of \( |\sigma_k|_{\mathcal{E}_k} \). Taking \( \beta > 1 \) in (38) for \( r = 1 \), or 2, applying Lemma 4.2 to the term \( I_1 \) and using (41)–(42), we get the following result for \( \| u_1 - \pi_h u \|_{DG} \):
\[
\| u_1 - \pi_h u \|_{DG} \lesssim E^r_\beta + \left( h^{r+1} + R^{r+\frac{\beta-1}{2}} \right) \| u \|_{H^{r+1}(\Omega)}.
\] (43)
Here, the superconvergence of the error between \( u_t \) and \( \pi_h u \) depends on \( \beta \). Note that (43) shows superconvergence between \( u_t \) and \( \pi_h u \).

Next, we bound the second term on the right hand side of (30). This term can be decomposed into the two parts \( \| \nabla (\pi_h u - u_h) \|_{L^2(E_h)} \) and \( \| \nabla (u_h - \bar{u}_h) \|_{L^2(E_h)} \). For the first part, we have

\[
\| \nabla (\pi_h u - u_h) \|_{L^2(E_h)} \leq \| \pi_h u - u_h \|_{DG} \lesssim h^{r+1} \| u \|_{H^{r+1}(\Omega)},
\]

where Lemma 3.5 has been used for the last inequality. For the second part \( \| \nabla (u_h - \bar{u}_h) \|_{L^2(E_h)} \), it is observed from Lemma 3.6 that

\[
\| \nabla (u_h - \bar{u}_h) \|_{L^2(E_h)} \lesssim \left( \sum_{e \in E_{\text{int}}} \| e \|_{\text{int}} \| u_h \|_{L^2(e)} \right)^{1/2} \lesssim \frac{h^{\beta-1}}{\sigma_0^2} \| f^\beta \|_{L^2(E_h)} \| u - u_h \|_{DG}.
\]

From (44)–(45) we bound

\[
\| \nabla (\pi_h u - \bar{u}_h) \|_{L^2(E_h)} \lesssim h^{r+1} \| u \|_{H^{r+1}(\Omega)} + \frac{h^{\beta-1}}{\sigma_0^2} \| u - u_h \|_{DG}.
\]

Finally, applying Lemma 4.2 and inserting (29)–(30), (35), (43) and (46) into (28), we obtain (34) and complete the proof. \( \square \)

Remark 2. From Theorem 4.3, the superconvergence result holds under the conditions that \( \beta > 1 \), \( \sigma > 0 \) and \( \epsilon =: O(h^{\alpha}) \), for any \( \alpha > 0 \). By the definition of Condition (\( \epsilon, \sigma \)), a convex quadrilateral formed by any two triangles at most has a perturbation of \( O(h^{\alpha}) \) from a parallelogram in \( \Omega_{0,h} \). The requirement \( \beta > 1 \) means over-penalized DG methods to be a good alternative. However, the theoretical proof above cannot be applied directly to the corresponding SIPG scheme with \( \beta = 1 \), which is left behind in the near future as an open problem.

Further, we derive the superconvergence between the gradient and the recovery gradient from OPSIPG solutions under suitable conditions. As shown in the proof of Theorem 4.3, some above estimates are very useful to prove superconvergence of \( \| \nabla u - G_h u_h \|_{L^2(\Omega)} \) for the DG gradient recovery \( G_h u_h \). Decompose the term \( \nabla u - G_h u_h \) by

\[
\nabla u - G_h u_h = (\nabla u - G_h u_t) + G_h (u_t - \pi_h u) + G_h (\pi_h u - u_h).
\]

Then with the use of (24), we get the following main superconvergence theorem from (29), (43) and Lemma 3.5.

**Theorem 4.4.** For \( \lambda = -1 \), let \( u \in W^{r+1,\infty}(\Omega) \cap H^{r+2}(\Omega) \cap H^{1,1}(\Omega) \) and \( u_h \) be the solution of (6) and (8), respectively. Assume that \( E_h \) satisfies Condition (\( \epsilon, \sigma \)) with \( \sigma > 0 \) and \( \epsilon =: O(h^{\alpha}) \), then there is a constant \( C \) independent of \( h \) and \( r \) such that

\[
\| \nabla u - G_h u_h \|_{L^2(\Omega)} \leq C \left( h^{r+1} \| u \|_{H^{r+2}(\Omega)} + h^{r+\frac{\beta-1}{2}} \| u \|_{H^{r+1}(\Omega)} + E_r^\beta \right),
\]

where \( E_r^\beta \) \( (r = 1, 2) \) are specified as in Theorem 4.3 and \( \beta > 1 \).

4.3. An a posteriori error estimator

We naturally define the global error estimator for discontinuous numerical solutions by

\[
\Pi_h := \| G_h u_h - \nabla u_h \|_{L^2(\Omega)}.
\]

Without a loss of generality, the small quantity \( \epsilon \) in Condition (\( \epsilon, \sigma \)) is assumed to be of \( O(h^{\alpha}) \) with \( \alpha > 0 \). Then we get the following asymptotically exact result motivated by Theorem 5.1 in [33].

**Theorem 4.5.** Under the assumptions of Theorem 4.4, and if there exists a constant \( C(u) > 0 \) such that

\[
\| \nabla (u - u_h) \|_{L^2(\Omega)} \geq C(u) h^r, \quad r = 1, 2.
\]

Then it holds

\[
\left| \frac{\Pi_h}{\| \nabla (u - u_h) \|_{L^2(\Omega)}} - 1 \right| \lesssim h^\rho, \quad \rho = \min \left\{ \frac{1}{2} \cdot \frac{\sigma}{2}, \frac{\beta - 1}{2} \right\}.
\]

**Proof.** Applying Theorem 4.4 and (49) completes the proof. \( \square \)
5. Numerical tests

In this section, we carry out a series of numerical experiments to validate our code and to illustrate superconvergence of the DG gradient recovery method. We shall focus on the superiority of the recovery method over the OPSIPG method by comparing the two facets: on the one hand, a linear DG finite element on uniform triangular meshes of the regular and chevron patterns (see Figs. 1–2), as well as on unstructured mesh (see Fig. 3); on the other hand, a quadratic element on the uniform triangular mesh of the regular and chevron patterns. Moreover, we will consider a milestone problem with a corner singularity to verify if the a posteriori error estimator $\Pi_h$ is asymptotically exact. As a post-processing technique, the PPR can be applied to any unstructured meshes. However, theoretical results are only available for very limited structured cases.

Case 1. We consider a nonlinear elliptic equation with a homogeneous boundary condition on the unit square $[0, 1]^2$ satisfying $a(u) = u^2 + 1$ and the exact solution

$$u(x, y) = x(1 - x)y(1 - y).$$

(51)

Case 2. The second example to be used is the nonlinear elliptic equation with $a(u) = u^2 + 1$ and a homogeneous boundary condition on the unit square $\Omega = [0, 1]^2$. The exact solution is chosen to be

$$u(x, y) = \sin(\pi x) \sin(\pi y).$$

(52)

We define $\| \nabla \cdot \|_{L^2(\Omega_{in})}$ as a discrete $H^1$ semi-norm in an interior region $\Omega_{in}$. An initial mesh for the linear elements in Cases 1 and 2 is produced by dividing the unit square into $2 \times 2$ uniform squares and only decomposing each subsquare into two triangles in the regular and chevron patterns. Then we intend to perform the numerical computation on multi-level meshes based on a similar process of uniform refinement, which preserves many advantages of initial meshes.
First, we present a numerical comparison on $H^1$-seminorm relative errors in Fig. 4 for OPSIPG scheme ($\beta > 1$), including SIPG scheme ($\beta = 1$), with different values of $\beta$ and the same penalty parameters $\sigma_r$ taken. The illustrations imply that the SIPG and OPSIPG schemes have almost the same convergence order of $H^1$-seminorm errors and their gradient recovery superconverges with different values of $\beta$. Numerically, the PPR technique can be applied to SIPG and OPSIPG schemes to realize superconvergence. Consequently, we may choose $\beta = 3, \sigma_e = 1/19$ in the over-penalized term and the relative error norms for the following numerical computation.

In the regular pattern, Figs. 5–6 illustrate the performance of the OPSIPG gradient recovery method on the $H^1$-seminorm relative errors for Cases 1 and 2. They compare a larger than second order convergence rate of the PPR to a first-order convergent rate of the DG method without using recovery technique. Some $H^1$-seminorm errors and convergence rates are listed in Table 1. In the chevron pattern, the PPR for Case 1 performs well with a superconvergence rate of order 1.81 in the inner region as in Fig. 7, while it also provides a superconvergent recovery of order 1.78 in $\Omega_{in}$ in Fig. 8 for Case 2. And Table 2 shows $H^1$-seminorm errors and convergence rates. For the unstructured meshes (see Fig. 3), the PPR for Case 1 performs well with a superconvergence rate of order 1.96 in Fig. 9, and it has a superconvergent recovery of order 1.85 in Fig. 10 for Case 2 and the convergence rates are listed in Table 3. In Figs. 11–12, we also illustrate the errors of the PPR in 3D based on SIPG and OPSIPG schemes, respectively.

We also consider the quadratic elements in the regular and chevron patterns with an inner region $[0.1, 0.9]^2$ for Case 2. Figs. 13 and 14 indicate more than third order convergence rates of the recovered gradient for Case 2. We also list the errors and convergence rates in Tables 4 and 5 for the regular and chevron patterns, respectively. Applying the quadratic elements in unstructured meshes, with a smaller inner region $[0.15, 0.85]^2$ for Cases 1–2, one observes from Figs. 15–16 that compared to the original gradients, the recovered gradients from the OPSIPG method are superconvergent. Table 6 shows the errors and convergence rates in the unstructured meshes for Cases 1–2.
Table 1  
Convergence rates in the regular pattern by linear elements and the initial mesh size $h = 1.25e-1$.  

<table>
<thead>
<tr>
<th>Regular</th>
<th>Mesh</th>
<th>$\frac{|\mathbf{V}u - \mathbf{V}u_h|<em>{L^2(\Omega)}}{|\mathbf{V}u - \mathbf{V}u_h|</em>{L^2(\Omega)}}$</th>
<th>Order</th>
<th>$\frac{|\mathbf{V}u - \mathbf{V}u_h|<em>{L^2(\Omega)}}{|\mathbf{V}u - \mathbf{V}u_h|</em>{L^2(\Omega)}}$</th>
<th>Order</th>
</tr>
</thead>
<tbody>
<tr>
<td>Case 1</td>
<td>$h$</td>
<td>4.030222e-1</td>
<td></td>
<td>1.4401</td>
<td>1.884911e-1</td>
</tr>
<tr>
<td></td>
<td>$h/2$</td>
<td>1.485308e-1</td>
<td></td>
<td>0.9293</td>
<td>7.438103e-2</td>
</tr>
<tr>
<td></td>
<td>$h/4$</td>
<td>7.799522e-2</td>
<td></td>
<td>1.0070</td>
<td>1.022604e-2</td>
</tr>
<tr>
<td></td>
<td>$h/8$</td>
<td>3.880853e-2</td>
<td></td>
<td>0.7966</td>
<td>2.874884e-3</td>
</tr>
<tr>
<td>Case 2</td>
<td>$h$</td>
<td>3.425271e-1</td>
<td></td>
<td>1.0352</td>
<td>1.931751e-1</td>
</tr>
<tr>
<td></td>
<td>$h/2$</td>
<td>1.624057e-1</td>
<td></td>
<td>0.7966</td>
<td>3.705632e-2</td>
</tr>
<tr>
<td></td>
<td>$h/4$</td>
<td>7.924589e-2</td>
<td></td>
<td>0.7966</td>
<td>9.829886e-3</td>
</tr>
<tr>
<td></td>
<td>$h/8$</td>
<td>4.027035e-2</td>
<td></td>
<td>0.7966</td>
<td>2.406333e-3</td>
</tr>
</tbody>
</table>

Table 2  
Convergence rates in the chevron pattern by linear elements and the initial mesh size $h = 2.5e-1$.  

<table>
<thead>
<tr>
<th>Chevron</th>
<th>Mesh</th>
<th>$\frac{|\mathbf{V}u - \mathbf{V}u_h|<em>{L^2(\Omega)}}{|\mathbf{V}u - \mathbf{V}u_h|</em>{L^2(\Omega)}}$</th>
<th>Order</th>
<th>$\frac{|\mathbf{V}u - \mathbf{V}u_h|<em>{L^2(\Omega)}}{|\mathbf{V}u - \mathbf{V}u_h|</em>{L^2(\Omega)}}$</th>
<th>Order</th>
</tr>
</thead>
<tbody>
<tr>
<td>Case 1</td>
<td>$h$</td>
<td>7.610685e-1</td>
<td></td>
<td>0.9056</td>
<td>4.346563e-1</td>
</tr>
<tr>
<td></td>
<td>$h/2$</td>
<td>4.062792e-1</td>
<td></td>
<td>1.1695</td>
<td>1.663987e-1</td>
</tr>
<tr>
<td></td>
<td>$h/4$</td>
<td>1.521582e-1</td>
<td></td>
<td>0.9633</td>
<td>5.056267e-2</td>
</tr>
<tr>
<td></td>
<td>$h/8$</td>
<td>7.803728e-2</td>
<td></td>
<td>0.9633</td>
<td>9.779187e-3</td>
</tr>
<tr>
<td>Case 2</td>
<td>$h$</td>
<td>6.923836e-1</td>
<td></td>
<td>0.9209</td>
<td>3.550487e-1</td>
</tr>
<tr>
<td></td>
<td>$h/2$</td>
<td>3.657054e-1</td>
<td></td>
<td>0.9209</td>
<td>1.453494e-1</td>
</tr>
<tr>
<td></td>
<td>$h/4$</td>
<td>1.626811e-1</td>
<td></td>
<td>1.1686</td>
<td>4.534501e-2</td>
</tr>
<tr>
<td></td>
<td>$h/8$</td>
<td>7.926170e-2</td>
<td></td>
<td>1.0137</td>
<td>9.507066e-3</td>
</tr>
</tbody>
</table>

Table 3  
Convergence rates in the unstructured meshes by linear elements and the initial mesh size $\max(h) = 1.2e-1$.  

<table>
<thead>
<tr>
<th>Unstructured</th>
<th>Mesh</th>
<th>$\frac{|\mathbf{V}u - \mathbf{V}u_h|<em>{L^2(\Omega)}}{|\mathbf{V}u - \mathbf{V}u_h|</em>{L^2(\Omega)}}$</th>
<th>Order</th>
<th>$\frac{|\mathbf{V}u - \mathbf{V}u_h|<em>{L^2(\Omega)}}{|\mathbf{V}u - \mathbf{V}u_h|</em>{L^2(\Omega)}}$</th>
<th>Order</th>
</tr>
</thead>
<tbody>
<tr>
<td>Case 1</td>
<td>$h$</td>
<td>3.439983e-1</td>
<td></td>
<td>1.1455</td>
<td>1.165086e-1</td>
</tr>
<tr>
<td></td>
<td>$h/2$</td>
<td>1.554941e-1</td>
<td></td>
<td>1.0601</td>
<td>5.063917e-2</td>
</tr>
<tr>
<td></td>
<td>$h/4$</td>
<td>7.457448e-2</td>
<td></td>
<td>1.0199</td>
<td>1.243027e-2</td>
</tr>
<tr>
<td></td>
<td>$h/8$</td>
<td>3.677619e-2</td>
<td></td>
<td>1.0199</td>
<td>1.976579e-3</td>
</tr>
<tr>
<td>Case 2</td>
<td>$h$</td>
<td>3.312666e-1</td>
<td></td>
<td>1.1353</td>
<td>1.954807e-1</td>
</tr>
<tr>
<td></td>
<td>$h/2$</td>
<td>1.508032e-1</td>
<td></td>
<td>1.0622</td>
<td>7.786595e-2</td>
</tr>
<tr>
<td></td>
<td>$h/4$</td>
<td>7.221908e-2</td>
<td></td>
<td>1.0243</td>
<td>1.954964e-2</td>
</tr>
<tr>
<td></td>
<td>$h/8$</td>
<td>3.550672e-2</td>
<td></td>
<td>1.0243</td>
<td>4.280404e-3</td>
</tr>
</tbody>
</table>

Table 4  
Convergence rates in the regular pattern by quadratic elements and the initial mesh size $h = 2.5e-1$.  

<table>
<thead>
<tr>
<th>Regular</th>
<th>Mesh</th>
<th>$\frac{|\mathbf{V}u - \mathbf{V}u_h|<em>{L^2(\Omega)}}{|\mathbf{V}u - \mathbf{V}u_h|</em>{L^2(\Omega)}}$</th>
<th>Order</th>
<th>$\frac{|\mathbf{V}u - \mathbf{V}u_h|<em>{L^2(\Omega)}}{|\mathbf{V}u - \mathbf{V}u_h|</em>{L^2(\Omega)}}$</th>
<th>Order</th>
</tr>
</thead>
<tbody>
<tr>
<td>Case 2</td>
<td>$h$</td>
<td>3.083754e-1</td>
<td></td>
<td>1.6118</td>
<td>1.097187e-1</td>
</tr>
<tr>
<td></td>
<td>$h/2$</td>
<td>1.088967e-1</td>
<td></td>
<td>3.5900</td>
<td>1.957182e-2</td>
</tr>
<tr>
<td></td>
<td>$h/4$</td>
<td>8.501780e-3</td>
<td></td>
<td>2.2659</td>
<td>1.238307e-3</td>
</tr>
<tr>
<td></td>
<td>$h/8$</td>
<td>1.767713e-3</td>
<td></td>
<td>2.4148</td>
<td>7.733422e-5</td>
</tr>
</tbody>
</table>

Table 5  
Convergence rates in the chevron pattern by quadratic elements and the initial mesh size $h = 2.5e-1$.  

<table>
<thead>
<tr>
<th>Chevron</th>
<th>Mesh</th>
<th>$\frac{|\mathbf{V}u - \mathbf{V}u_h|<em>{L^2(\Omega)}}{|\mathbf{V}u - \mathbf{V}u_h|</em>{L^2(\Omega)}}$</th>
<th>Order</th>
<th>$\frac{|\mathbf{V}u - \mathbf{V}u_h|<em>{L^2(\Omega)}}{|\mathbf{V}u - \mathbf{V}u_h|</em>{L^2(\Omega)}}$</th>
<th>Order</th>
</tr>
</thead>
<tbody>
<tr>
<td>Case 2</td>
<td>$h$</td>
<td>3.160596e-1</td>
<td></td>
<td>1.2831</td>
<td>1.599464e-1</td>
</tr>
<tr>
<td></td>
<td>$h/2$</td>
<td>1.298687e-1</td>
<td></td>
<td>3.5897</td>
<td>2.492384e-2</td>
</tr>
<tr>
<td></td>
<td>$h/4$</td>
<td>1.078681e-2</td>
<td></td>
<td>3.5897</td>
<td>1.627671e-3</td>
</tr>
<tr>
<td></td>
<td>$h/8$</td>
<td>2.022879e-3</td>
<td></td>
<td>3.5897</td>
<td>1.681631e-4</td>
</tr>
</tbody>
</table>
Fig. 9. Linear elements in the unstructured meshes, $\|\nabla u - G_h u_h\|_{L^2(\Omega_1)} / \|\nabla u\|_{L^2(\Omega_1)}$ for Case 1.

Fig. 10. Linear elements in the unstructured meshes, $\|\nabla u - G_h u_h\|_{L^2(\Omega_1)} / \|\nabla u\|_{L^2(\Omega_1)}$ for Case 2.

Fig. 11. Error $\nabla u - G_h u_h$ in 3D by SIPG gradient recovery method, linear elements in the unstructured meshes in $\Omega_1$ := $[0.15, 0.85]^2$ for Case 2.

Table 6

<table>
<thead>
<tr>
<th>Unstructured</th>
<th>Mesh</th>
<th>$\frac{|\nabla u - G_h u_h|<em>{L^2(\Omega_1)}}{|\nabla u|</em>{L^2(\Omega_1)}}$</th>
<th>Order</th>
<th>$\frac{|\nabla u - G_h u_h|<em>{L^2(\Omega_1)}}{|\nabla u|</em>{L^2(\Omega_1)}}$</th>
<th>Order</th>
</tr>
</thead>
<tbody>
<tr>
<td>Case 1</td>
<td>$h$</td>
<td>1.409106e−1</td>
<td>5.96</td>
<td>7.933208e−2</td>
<td>7.96</td>
</tr>
<tr>
<td></td>
<td>$h/2$</td>
<td>2.446072e−3</td>
<td>5.84</td>
<td>5.703833e−4</td>
<td>7.11</td>
</tr>
<tr>
<td></td>
<td>$h/4$</td>
<td>4.179735e−4</td>
<td>5.54</td>
<td>6.599119e−5</td>
<td>6.59</td>
</tr>
<tr>
<td>Case 2</td>
<td>$h$</td>
<td>3.831871e−2</td>
<td>5.34</td>
<td>2.536329e−2</td>
<td>5.34</td>
</tr>
<tr>
<td></td>
<td>$h/2$</td>
<td>2.689413e−3</td>
<td>3.83</td>
<td>6.226057e−4</td>
<td>5.26</td>
</tr>
<tr>
<td></td>
<td>$h/4$</td>
<td>5.695356e−4</td>
<td>2.24</td>
<td>7.518793e−5</td>
<td>4.04</td>
</tr>
</tbody>
</table>

Case 3. The last example is a singularity problem setting in the L-shaped domain $\Omega = (0, 1)^2 \setminus [1/2, 1]^2$ with $a(u) = 3$. Under a polar coordinate system $(r, \theta)$ with the origin $(1/2, 1/2)$, the solution is

$$u(r) = \frac{1}{3} r^2 \sin \left( \frac{2\theta + 5\pi}{3} \right), \quad \frac{\pi}{2} \leq \theta \leq 2\pi.$$  

Note that the solution in Case 3 has a corner singularity at the node $(1/2, 1/2)$ as well as the other five vertices of the L-shaped domain. Delaunay meshes generated by a software DISTMESH [26] have been considered as an adaptive mesh to simulate the convergence behavior. The domain is decomposed into two parts: one is the interior subdomain $\Omega_{in} := [0.08, 0.92] \times [0.08, 0.42] \cup [0.08, 0.42] \times [0.42, 0.92]$, and the other one is the boundary layer $\Omega_{ext} = \Omega \setminus \Omega_{in}$ with a width of
With the use of the gradient recovery method on a locally refined mesh (see Fig. 17), we observe in Fig. 18 that \( \| \nabla u - G_h u_h \|_{L^2(\Omega_h)} \) is superconvergent, and the term \( \| G_h u_h - \nabla u_h \|_{L^2(\Omega_h)} / \| \nabla u - \nabla u_h \|_{L^2(\Omega_h)} \) approaches 1 at the slope \(-2.47\) in the interior subdomain. In Table 7, we present the absolute errors \( \| \nabla (u - u_h) \|_{L^2(\Omega_h)} \), \( \| \nabla u - G_h u_h \|_{L^2(\Omega_h)} \), and convergence orders with respect to DoFs on Delaunay meshes with linear elements. Consequently, the gradient recovery operator \( G_h \) is effective and the \textit{a posteriori} error estimator \( \Pi_h \) is asymptotically exact under mildly structured grids. It is observed that the polynomial preserving recovery technique provides superconvergent recovered gradients for discontinuous solutions to the OPSIPG method.

6. Conclusions

In this work we employed the PPR to realize the gradient recovery technique based on the over-penalized interior penalty discontinuous Galerkin method for solving the second-order nonlinear elliptic problem. Superconvergence of the PPR implemented on the OPSIPG method has been rigorously proved. In this work, the gradient recovery method has superconvergence for linear and quadratic elements on the structured meshes in the regular and chevron patterns, even on the unstructured meshes. In addition, the gradient recovery method for the linear elements is well-performed for the corner
Fig. 15. Quadratic elements in unstructured meshes, $\|\nabla u - G_h u_h\|_{L^2(\Omega_1)} / \|\nabla u\|_{L^2(\Omega_1)}$, $\Omega_1 := [0.15, 0.85]^2$ for Case 1.

Fig. 16. Quadratic elements in unstructured meshes, $\|\nabla u - G_h u_h\|_{L^2(\Omega_1)} / \|\nabla u\|_{L^2(\Omega_1)}$, $\Omega_1 := [0.15, 0.85]^2$ for Case 2.

Fig. 17. A locally refined mesh in Case 3.

Fig. 18. Errors $\|\nabla u - G_h u_h\|_{L^2(\Omega_1)}$, $\|\nabla u - V_h u_h\|_{L^2(\Omega_1)}$, and the ratio versus the number of DoFs for Case 3, linear elements in locally refined meshes.

Table 7
Convergence orders with respect to DoFs in Delaunay meshes by linear elements.

<table>
<thead>
<tr>
<th>Case</th>
<th>DoFs</th>
<th>$|\nabla(u - u_h)|_{L^2(\Omega_1)}$</th>
<th>Order</th>
<th>$|\nabla(u - G_h u_h)|_{L^2(\Omega_1)}$</th>
<th>Order</th>
<th>The ratio</th>
</tr>
</thead>
<tbody>
<tr>
<td>3</td>
<td>822</td>
<td>6.331477e-2</td>
<td></td>
<td>4.732690e-2</td>
<td></td>
<td>1.932319e-1</td>
</tr>
<tr>
<td></td>
<td>1284</td>
<td>3.657636e-2</td>
<td>-1.2303</td>
<td>2.589395e-2</td>
<td>-1.3522</td>
<td>2.272679e-1</td>
</tr>
<tr>
<td></td>
<td>2949</td>
<td>2.404580e-2</td>
<td>-0.5044</td>
<td>5.423565e-3</td>
<td>-1.8001</td>
<td>1.791533e-2</td>
</tr>
<tr>
<td></td>
<td>5376</td>
<td>1.810957e-2</td>
<td>-0.4722</td>
<td>1.642915e-3</td>
<td>-1.9889</td>
<td>4.456809e-3</td>
</tr>
<tr>
<td></td>
<td>11997</td>
<td>1.216340e-2</td>
<td>-0.4958</td>
<td>7.442289e-4</td>
<td>-0.9865</td>
<td>3.541773e-4</td>
</tr>
</tbody>
</table>

singularity problem with the inhomogeneous Dirichlet boundary condition, illustrating that the *a posteriori* error estimator based on the recovered gradient is asymptotically exact and efficient. Following the preconditioning technique in [5], it is straightforward to remedy the ill-conditioned system. However, related questions that remain open comprise construction of an optimal preconditioner of the resulting systems, and superconvergence analysis of the PPR for IPDG solutions in general meshes.

Acknowledgements

The first author of this work was partially supported by the National Natural Science Foundation of China (Grants 11101196 and 11471150) and by the Natural Science Foundation of Gansu Province, China (Grant 145RJZA046). The second author was supported in part by the National Natural Science Foundation of China (Grants 11471031 and 91430216), and the US National Science Foundation (Grant DMS-1419040).
References

[40] Introduction to COMSOL Multiphysics Version 5.1, March 2015, p. 46.