DISCRETE AND CONTINUOUS
DYNAMICAL SYSTEMS SERIES B
Volume **26**, Number **5**, May **2021**

doi:10.3934/dcdsb.2020184

pp. **2411–2428**

A NEW OVER-PENALIZED WEAK GALERKIN METHOD. PART I: SECOND-ORDER ELLIPTIC PROBLEMS

Kaifang Liu and Lunji Song*

School of Mathematics and Statistics
and Key Laboratory of Applied Mathematics and Complex Systems (Gansu Province)
Lanzhou University, Lanzhou 730000, China

Shan Zhao

Department of Mathematics, University of Alabama
Tuscaloosa, AL 35487, USA

(Communicated by Zhimin Zhang)

Abstract. In this article, we propose a new over-penalized weak Galerkin (OPWG) method with a stabilizer for second-order elliptic problems. This method employs double-valued functions on interior edges of elements instead of single-valued ones and elements $(\mathbb{P}_k, \mathbb{P}_k, [\mathbb{P}_{k-1}]^d)$, or $(\mathbb{P}_k, \mathbb{P}_{k-1}, [\mathbb{P}_{k-1}]^d)$, with dimensions of space $d = 2, 3$. The method is absolutely stable with a constant penalty parameter, which is independent of mesh size and shape-regularity. We prove that for quasi-uniform triangulations, condition numbers of the stiffness matrices arising from the OPWG method are $O(h^{-\beta_0(d-1)-1})$, β_0 being the penalty exponent. Therefore we introduce a new *mini-block diagonal* preconditioner, which is proven to be theoretically and numerically effective in reducing the condition numbers of stiffness matrices to the magnitude of $O(h^{-2})$. Optimal error estimates in a discrete H^1-norm and L^2-norm are established, from which the optimal penalty exponent can be easily chosen. Several numerical examples are presented to demonstrate flexibility, effectiveness and reliability of the new method.

1. **Introduction.** The weak Galerkin finite element method (WG-FEM) was first introduced by Wang and Ye [17] for solving general second order elliptic equations, and a computational study was conducted in [5], illustrating that the WG method is efficient and reliable for scientific computing. In general, the WG method refers to a finite element one in which true solutions are approximated by weak functions consisting of two parts, defined in the interior and on the edges of each element respectively, while differential operators are approximated by weak ones.

Two kinds of elements are commonly employed in the WG-FEM. The first kind of elements include Raviart-Thomas element (RT) [12], which can be written as $(\mathbb{P}_k, \mathbb{P}_k, RT_k)$, and Brezzi-Douglas-Marini element (BDM) [2], written as

2020 *Mathematics Subject Classification.* Primary: 65N15, 65N30; Secondary: 35j50.

Key words and phrases. Weak Galerkin, over-penalized, second order elliptic problems, finite element method, preconditioning, discrete weak gradient.

* Corresponding author: Lunji Song.

$(\mathbb{P}_k, \mathbb{P}_{k+1}, [\mathbb{P}_{k+1}]^d)$, where the three parts represent polynomial spaces for weak functions and weak gradient, respectively. These elements naturally possess weak continuity between interior part and boundary part of weak functions within each element, without requirement for stabilization. However, the limitation to simplex partitions makes them less attractive. The second kind of elements are $(\mathbb{P}_k, \mathbb{P}_k, [\mathbb{P}_{k-1}]^d)$ and $(\mathbb{P}_k, \mathbb{P}_{k-1}, [\mathbb{P}_{k-1}]^d)$ with an integer $k \geq 1$. Compared to RT and BDM elements, elements of the second kind need a stabilizer for variational schemes to ensure the weak continuity, which brings great convenience and flexibility [8] in mesh generation and assembly of stiffness matrix. In [20], a WG scheme with boundary continuity has been proposed for second-order elliptic problems, where degrees of freedom can be greatly reduced to magnitude of standard conforming Galerkin method by using the Schur complement. With these progress, the WG-FEM becomes more powerful and has been developed for solving many applications, such as elliptic interface problems, Stokes, Helmholtz, Maxwell etc [6, 19, 7, 10, 16]. Very recently, to overcome the difficulties on low regularity solutions, we also presented a relaxed WG method for solving elliptic and elliptic interface problems and analyzed L^p error estimates [14, 15].

Traditionally, the WG-FEM is defined based on the weak functions in which the edge part is single-valued on each interior edge. Recently, a penalized WG method based on RT elements for solving second-order elliptic problems is proposed and analyzed in [4], where the shape functions along interior edges are double-valued. The main idea of Ref. [4] is to integrate WG-FEM with discontinuous Galerkin (DG) methods [13] in a framework, giving rise to an over-penalized WG (OPWG) method. The OPWG method can inherit some advantages of DG methods. According to theoretical analysis and numerical results, it has succeeded to a large extent. It makes stable approximation, adaptive meshes and *hp* error analysis possible, and the OPWG method is very promising in parallel computing and in discretizations of other PDEs with jump solutions. It is observed that linear systems from the penalized WG method in [4] have to overcome increasing condition numbers of stiffness matrices. Although those condition numbers numerically have been settled by using a simple, easy-invertible preconditioner, its analysis is still open.

In this paper, we generalize the idea of penalized WG method in [4] and adapt it to elements of the second kind $(\mathbb{P}_k, \mathbb{P}_{k-1}, [\mathbb{P}_{k-1}]^d)$. There are three major objectives for this work. The first is to extend the limited range of the OPWG method in weak finite element spaces and element shapes. Since there are rather difference between the two kinds of elements, for completeness, we present the error estimates in details. Moreover, the new choice of polynomial combination will greatly increase the geometry flexibility and facilitate the mesh generation. The second is to design a more effective *mini-block* diagonal preconditioner of stiffness matrices whenever the penalty exponent β_0 increases. We also intend to show a theoretical estimate on preconditioning effect, which can be modified without any difficulty to analyze the conditioning of original WG method [8] and the effect of the preconditioner introduced in [4]. Finally, to reduce the size of stiffness matrices, a Schur complement formulation will be developed, which can eliminate degrees of freedom corresponding to shape functions defined in interior of elements. Thus, one can expect that for simplex meshes, the new OPWG method with $(\mathbb{P}_k, \mathbb{P}_{k-1}, [\mathbb{P}_{k-1}]^d)$ element has the same or less degrees of freedom compared to the primal interior penalty discontinuous Galerkin methods [13, 1]. Furthermore, this reducing process can be done elementwisely and simultaneously, which could be easily to be parallelized.

Here we consider a second-order elliptic problem with nonhomogeneous Dirichlet boundary condition:

$$-\nabla \cdot A\nabla u = f, \quad \text{in } \Omega, \tag{1}$$

$$u = g, \quad \text{on } \partial\Omega, \tag{2}$$

where Ω is a polygonal or polyhedral domain in $\mathbb{R}^d$ $(d = 2, 3)$, $f \in L^2(\Omega)$ and A is a symmetric positive definite matrix-valued function on Ω, i.e., there exist two positive numbers $\lambda_1, \lambda_2 > 0$ such that

$$\lambda_1 \xi^t \xi \le \xi^t A \xi \le \lambda_2 \xi^t \xi, \quad \forall\, \xi \in \mathbb{R}^d,$$

where ξ is a column vector and ξ^t means the transpose of ξ.

The rest of this paper is organized as follows. In Section 2, we recall the definition of weak gradient, and define the jumps of weak functions as well as some projection operators. In Section 3, we present an OPWG method. To handle the problem of condition number, an easy-implement and efficient preconditioner is introduced in Section 4 and theoretical analysis is given to dramatically reduce the condition number of the stiffness matrices arising from the OPWG method to $O(h^{-2})$. In Section 5, *a priori* error estimates in energy norm and L^2- norm are derived rigorously. A practical Schur complement formulation is developed in Section 6. In the last section, some numerical experiments are conducted.

2. **Notations and preliminaries.** We begin with some basic notations and definitions. Let $\mathcal{T}_h$ be a shape regular partition of the domain Ω defined in [18, Assumptions **A1-A4**], $\mathcal{E}$ be all edges (or faces), and $\mathcal{E}_{\mathcal{I}}$ be all interior edges. We denote by h_K diameter for every $K \in \mathcal{T}_h$ and denote by $h = \max_{K\in\mathcal{T}_h} h_K$ mesh size for $\mathcal{T}_h$. For a given partition $\mathcal{T}_h$, let $\mathbb{P}_k(K)$ be the space of piecewise polynomials of degree no more than k on $\mathcal{T}_h$ in each $K \in \mathcal{T}_h$. Similarly, $\mathbb{P}_k(e)$ denotes the space of piecewise polynomials of degree no more than k on $e \in \mathcal{E}$.

First, we define local weak function space as follows

$$W_{k,j}(K) = \{v = (v_0, v_b) : v_0 \in \mathbb{P}_k(K), v_b \in \mathbb{P}_j(e), e \in \partial K\},$$

where we assume $j = k$, and the case of $j = k - 1$ can be analyzed similarly.

Analogous to the definition in [17], we define a local discrete weak gradient operator for each $v \in W_{k,k}(K)$.

Definition 2.1. For any $v \in W_{k,k}(K)$, the discrete local weak gradient operator on $K \in \mathcal{T}_h$, denoted by $\nabla_{w,k-1}$, is defined as the unique polynomial $(\nabla_{w,k-1} v) \in [\mathbb{P}_{k-1}(K)]^d$ satisfying the following equation

$$(\nabla_{w,k} v, \mathbf{q})_K = -(v_0, \nabla \cdot \boldsymbol{q})_K + \langle v_b, \mathbf{q} \cdot \boldsymbol{n} \rangle_{\partial K}, \quad \forall\ \mathbf{q} \in [\mathbb{P}_{k-1}(K)]^d.$$

For a given $k \ge 1$, let V_h be the over-penalized weak Galerkin finite element space associated with $\mathcal{T}_h$ defined as follows

$$\begin{aligned} V_h = \{(v_0,\ v_b) : v_0|_K \in \mathbb{P}_k(K),\ K \in \mathcal{T}_h, \\ v_b|_e \in \mathbb{P}_j(e) \times \mathbb{P}_j(e), e \in \mathcal{E}_{\mathcal{I}};\ v_b|_e \in \mathbb{P}_j(e), e \in \partial\Omega\}, \end{aligned} \tag{3}$$

$$V_h^0 = \{(v_0,\ v_b) : v \in V_h,\ v_b = 0 \text{ on } \partial\Omega\}, \tag{4}$$

where we assume that $j = k$, and the case of $j = k - 1$ can be analyzed similarly. It is worth noting that v_b is a double-valued function on each interior edge.

For convenience, we introduce a set of normals of $\mathcal{E}$ as follows:

$$\mathcal{D} = \{\boldsymbol{n}_e : \boldsymbol{n}_e \text{ is unit and normal to } e, e \in \mathcal{E}\}. \tag{5}$$

If $e \in \partial\Omega$, we set $\boldsymbol{n}_e$ to be a unit normal vector exterior to the domain. If $e \in \mathcal{E}_{\mathcal{I}}$ and $e = K_1^e \cap K_2^e$, we always assume $\boldsymbol{n}_e$ being directed from K_1^e to K_2^e. Since there are two traces v_b along every interior edge e, we denote by $[\![v_b]\!]$ the weak jump for v_b:

$$[\![v_b]\!] = v_b|_{K_1^e} - v_b|_{K_2^e}, \quad \forall\ e = \partial K_1^e \cap \partial K_2^e.$$

We also extend the definition of jump to sides that belong to the domain boundary, i.e., $[\![v_b]\!] = v_b|_e,\ \forall\ e \in \partial\Omega$. In addition, we denote by $|e|$ length or area of e and by h_e the diameter of edge of flat face $e \in \mathcal{E}$. For a shape regular mesh, there exist constants κ, ρ_e such that $\kappa h_K \le h_e$ and $\rho_e h_K^{d-1} \le |e|,\ e \in \partial K$.

To investigate the approximation properties of the finite element space V_h and weak gradient space $[\mathbb{P}_{k-1}(K)]^d$, we introduce the following three projections:

$$\begin{aligned}
&Q_0 : L^2(K) \to \mathbb{P}_k(K), \quad \forall K \in \mathcal{T}_h,\\
&Q_b : L^2(e) \to \mathbb{P}_k(e), \quad \forall K \in \mathcal{E},\\
&\mathbb{Q}_h : [L^2(K)]^d \to [\mathbb{P}_{k-1}(K)]^d, \quad \forall K \in \mathcal{T}_h.
\end{aligned}$$

We shall combine Q_0 and Q_b by writing $Q_h = (Q_0, Q_b)$. It is easy to see the following identity (see [8]):

$$\nabla_w(Q_h\phi) = \mathbb{Q}_h(\nabla\phi), \quad \forall\phi \in H^1(K). \tag{6}$$

In fact, by using the definition of Q_h and $\mathbb{Q}_h$, for all $\boldsymbol{q} \in [\mathbb{P}_{k-1}(K)]^d$,

$$\begin{aligned}
(\nabla_w(Q_h\phi), \boldsymbol{q})_K &= -(Q_0\phi, \nabla\cdot\boldsymbol{q})_K + \langle Q_b\phi, \boldsymbol{q}\cdot\boldsymbol{n}\rangle_{\partial K}\\
&= -(\phi, \nabla\cdot\boldsymbol{q})_K + \langle \phi, \boldsymbol{q}\cdot\boldsymbol{n}\rangle_{\partial K}\\
&= (\nabla\phi, \boldsymbol{q})_K\\
&= (\mathbb{Q}_h\nabla\phi, \boldsymbol{q})_K.
\end{aligned}$$

It is worthy to note that the analysis of this paper is completely valid if we set the local weak function space to be $W_{k,k-1}(K), K \in \mathcal{T}_h$. In this case, the projection Q_b shall be defined as $Q_b : L^2(e) \to \mathbb{P}_{k-1}(e)$, $\forall K \in \mathcal{E}$.

3. The over-penalized weak Galerkin scheme.

For any $w, v \in V_h$, we define the following bilinear forms

$$\begin{aligned}
s(w,v) &:= \sum_{K\in\mathcal{T}_h} h_K^{-1}\langle Q_bw_0 - w_b, Q_bv_0 - v_b\rangle_{\partial K},\\
a_s(w,v) &:= (A\nabla_w w, \nabla_w v) + s(w,v) + \sum_{e\in\mathcal{E}_{\mathcal{I}}} |e|^{-\beta_0}\langle [\![w_b]\!], [\![v_b]\!]\rangle_e,
\end{aligned}$$

where $s(\cdot,\cdot)$ is the stabilizer, and $\beta_0 \ge 1$ is a penalty parameter. If $\beta_0 > 1$, it is referred to over-penalization or super-penalization.

Algorithm 1. *A numerical approximation for* (1)-(2) *can be obtained by seeking* $u_h = (u_0, u_b) \in V_h$ *satisfying* $u_b = Q_bg$ *on* $\partial\Omega$ *and such that*

$$a_s(u_h, v) = (f, v_0), \quad \forall v = (v_0, v_b) \in V_h^0. \tag{7}$$

Next, we justify the well-posedness of Algorithm 1. It's rational to define the following norm due to the fact that $a_s(v,v)$ is nonnegative for any $v \in V_h$:

$$|\!|\!| v |\!|\!| := \sqrt{a_s(v,v)}. \tag{8}$$

The functional $|||\cdot|||$ defines a semi-norm in V_h. It also defines a norm in V_h^0. To verify this, it suffices to check the positivity property for $|||\cdot|||$. Assume $v \in V_h^0$ and $|||\cdot||| = 0$, i.e.,

$$(A\nabla_w v, \nabla_w v) + \sum_{K\in\mathcal{T}_h} h_K^{-1}\langle Q_b v_0 - v_b, Q_b v_0 - v_b\rangle_{\partial K} + \sum_{e\in\mathcal{E}_h} |e|^{-\beta_0}\int_e [\![v_b]\!]^2 = 0,$$

which means $\nabla_w v = 0$ on each element $K \in \mathcal{T}_h$, $Q_b v_0 = v_b$ on ∂K and $[\![v_b]\!] = 0$ on each edge $e \in \mathcal{E}_{\mathcal{I}}$. Therefore, we have by the definition of weak gradient for any $\boldsymbol{q} \in [\mathbb{P}_{k-1}(K)]^d$

$$\begin{aligned}0 &= (\nabla_w v, \boldsymbol{q})_K\\ &= (\nabla v_0, \boldsymbol{q})_K - \langle v_0 - v_b, \boldsymbol{q}\cdot\boldsymbol{n}\rangle_{\partial K}\\ &= (\nabla v_0, \boldsymbol{q})_K - \langle Q_b v_0 - v_b, \boldsymbol{q}\cdot\boldsymbol{n}\rangle_{\partial K}\\ &= (\nabla v_0, \boldsymbol{q})_K.\end{aligned}$$

Let $\boldsymbol{q} = \nabla v_0$, then $\nabla v_0 = 0$ on each element, that is $v_0 = const$. This, together with $Q_b v_0 = v_b$ and $[\![v_b]\!] = 0$ on each edge, implies $v_0 = v_b = 0$.

Now, we shall give the following lemma without proof.

Lemma 3.1. *The over-penalized weak Galerkin scheme (7) has one and only one solution.*

Mass conversation. The model problem (1) can be rewritten in a conservative form as follows:

$$\nabla\cdot q = f, \quad q = -A\nabla u.$$

Let K be any control volume. Integrating the first equation over K yields the following integral form of mass conservation:

$$\int_{\partial K} q\cdot\boldsymbol{n} = \int_K f. \tag{9}$$

We claim that the numerical approximation from the over-penalized weak Galerkin finite element method for (1) retains the mass conservation property (9) with a numerical flux q_h. To this end, for any given $K \in \mathcal{T}_h$, we choose test function $v = (v_0, v_b = 0) \in V_h$ such that $v_0 = 1$ on K and $v_0 = 0$ elsewhere, which means the jump $[\![v_b]\!] = 0$ on each edge. Then Algorithm (7) becomes

$$\begin{aligned}\int_K f dK &= \int_K A\nabla_w u_h\cdot\nabla_w v + h_K^{-1}\int_{\partial K}(Q_b u_0 - u_b)\\ &= \int_K \mathbb{Q}_h(A\nabla_w u_h)\cdot\nabla_w v + h_K^{-1}\int_{\partial K}(Q_b u_0 - u_b)\\ &= -\int_K \nabla\cdot\mathbb{Q}_h(A\nabla_w u_h) + h_K^{-1}\int_{\partial K}(Q_b u_0 - u_b)\\ &= \int_{\partial K}\{-\mathbb{Q}_h(A\nabla_w u_h) + h_K^{-1}(Q_b u_0 - u_b)\cdot\boldsymbol{n}\}\cdot\boldsymbol{n},\end{aligned} \tag{10}$$

which means that the OPWG method conserves mass with a numerical flux given by

$$q_h = -\mathbb{Q}_h(A\nabla_w u_h) + h_K^{-1}(Q_b u_0 - u_b)\cdot\boldsymbol{n}.$$

Furthermore, we have another balance equation for u_h. For any element K, $\partial K \cap \partial\Omega = \varnothing$, if we choose test function $v = (v_0, v_b) \in V_h$ such that $v_0 = v_b = 1$ on

$\overline{K}$ and $v = 0$ elsewhere, then (7) becomes

$$\begin{aligned}\int_K f dK &= \sum_{e\in\partial K} \frac{1}{|e|^{\beta_0}} \int_e [\![u_b]\!] \\ &= \sum_{e\in\partial K} \frac{1}{|e|^{\beta_0}} \int_e (u_b|_K - u_b|_{\mathcal{N}(e;K)}),\end{aligned} \tag{11}$$

where we have used $\nabla_w v|_K = 0$ and $Q_b v_0 = v_b$, and $\mathcal{N}(e;K)$ denotes a neighbor element of K sharing the edge e.

4. **Preconditioning of the OPWG method.** In this section, we will present the conditioning of the OPWG method, and give a simple preconditioner and some theoretical analysis about its effectiveness. Note that for the conditioning of WG method with Raviart-Thomas elements, the readers are referred to an interesting work Ref. [3] by Li and Xie. Assume $\mathcal{T}_h$ be a family of shape regular triangulations of $\overline{\Omega}$ satisfying quasi-uniformity, i.e., there exists a constant τ such that

$$\min_{K\in\mathcal{T}_h} h_K \geq \tau h, \quad \forall\, h > 0. \tag{12}$$

First, we give trace and inverse inequalities for functions defined on shape regular partitions [18, Lemma A.3 and Lemma A.6].

Lemma 4.1. *Let $\mathcal{T}_h$ be a regular partition of the domain Ω into polygons in 2D or polyhedra in 3D. Then, there exists a constant C_t such that for any $K \in \mathcal{T}_h$ and edge/face $e \in \partial\mathcal{T}$,*

$$\|\theta\|_e^2 \leq C_t \left(h_K^{-1}\|\theta\|_K^2 + h_K\|\theta\|_K^2\right), \quad \forall\, \theta \in H^1(K). \tag{13}$$

Moreover, there exists a constant $C_i = C_i(k)$ such that

$$\|\nabla\phi\|_K \leq C_i h_K^{-1}\|\phi\|_K, \quad \forall\, \phi \in \mathbb{P}_k(K). \tag{14}$$

Define the following norms for any $v \in V_h$:

$$\text{Descrete } L^2 \text{ norm } \|v\|_h^2 := \sum_{K\in\mathcal{T}_h} \left(\|v_0\|_K^2 + h_K\|v_b\|_{\partial K}^2\right), \tag{15}$$

$$\text{Seminorm } \|v\|_{1,h}^2 := \sum_{K\in\mathcal{T}_h} \left(\|\nabla v_0\|_K^2 + h_K^{-1}\|Q_b v_0 - v_b\|_{\partial K}^2\right) + \sum_{e\in\mathcal{E}_\mathcal{I}} |e|^{-\beta_0}\|[\![v_b]\!]\|_e^2. \tag{16}$$

One can easily prove that $\|\cdot\|_{1,h}$ is equivalent for any $v \in V_h$ to the triple-bar norm $|\!|\!|\cdot|\!|\!|$ (see [8]).

Let $\{\phi_{0,i}\}_{i=1}^{N_0}$, $\{\phi_{b,i}\}_{i=1}^{N_b}$ be the Lagrange basis for V_h, and for any $v \in V_h$ there is a vector $\boldsymbol{\eta} = [\boldsymbol{\eta}_0;\ \boldsymbol{\eta}_b] = [\eta_{0,1}, \ldots, \eta_{0,N_0}, \eta_{b,1}, \ldots, \eta_{b,N_b}]^T$ such that

$$v = \{v_0, v_b\} = \left\{\sum_{i=1}^{N_0} \eta_{0,i}\phi_{0,i}, \sum_{i=1}^{N_b} \eta_{b,i}\phi_{b,i}\right\}.$$

We denote $|\boldsymbol{\eta}|$ to be the standard Euclidean norm of vector $\boldsymbol{\eta}$. Following the idea of [11, Proposition 6.3.1], we can get

Lemma 4.2. *Let $\mathcal{T}_h$ be a quasi-uniform family of triangulations of $\overline{\Omega}$. Then there exist positive constants C_1 and C_2 such that for each $v \in V_h$,*

$$C_1 h^d |\boldsymbol{\eta}|^2 \leq \|v\|_h^2 \leq C_2 h^d |\boldsymbol{\eta}|^2. \tag{17}$$

Next we shall present two estimates between norms $\|\cdot\|_h$ and $|||\cdot|||$. Since penalty parameter $\beta_0 \geq 1$, one can easily get

$$
\begin{aligned}
|||v|||^2 &\leq C\|v\|_{1,h}^2 \\
&\leq C \sum_{K\in\mathcal{T}_h} \left(h_K^{-2}\|v_0\|_K^2 + h_K^{-1}\|v_b\|_{\partial K}^2 + h_K^{-\beta_0(d-1)}\|v_b\|_{\partial K}^2 \right) \\
&\leq Ch^{-\beta_0(d-1)-1} \sum_{K\in\mathcal{T}_h} \left(h_K^{\beta_0(d-1)-1}\|v_0\|_K^2 + (h_K^{\beta_0(d-1)} + h_K)\|v_b\|_{\partial K}^2 \right) \\
&\leq Ch^{-\beta_0(d-1)-1}\|v\|_h^2.
\end{aligned} \tag{18}
$$

For the inverse estimate, we need the following lemma, a variation of [9, Lemma 7.1].

Lemma 4.3. *Assume that the finite element partition $\mathcal{T}_h$ is shape regular. Then, there exists a constant C independent of the mesh size h such that*

$$
\|v_0\| \leq C\, |||\, v|||, \qquad v \in V_h^0. \tag{19}
$$

Proof. For any $v = \{v_0, v_b\} \in V_h^0$, let $\boldsymbol{q} \in [H^1(\Omega)]^d$ satisfy $\nabla\cdot\boldsymbol{q} = v_0$ and $\|\boldsymbol{q}\|_{1,\Omega} \leq C\|v_0\|$. To see an existence of such a function $\boldsymbol{q}$, one may first extend v_0 by zero to a convex domain $\tilde{\Omega}$ which contains Ω, and then consider the Poisson equation $\Delta\Psi = v_0$ on the enlarged domain $\tilde{\Omega}$ and set $\boldsymbol{q} = \nabla\Psi$. The required properties of $\boldsymbol{q}$ follow immediately from the full regularity of Poisson's equation on convex domains.

Recall that $\mathbb{Q}_h$ is the L^2 projection to the space of discrete weak gradient space $[\mathbb{P}_{k-1}]^d$. Thus,

$$
\begin{aligned}
\|v_0\|^2 &= \sum_{K\in\mathcal{T}_h} (v_0, v_0)_K = \sum_{K\in\mathcal{T}_h} (v_0, \nabla\cdot\boldsymbol{q})_K \\
&= \sum_{K\in\mathcal{T}_h} \left(-(\nabla v_0, \boldsymbol{q})_K + \langle v_0, \boldsymbol{q}\cdot\boldsymbol{n}\rangle_{\partial K}\right) = \sum_{K\in\mathcal{T}_h} \left(-(\nabla v_0, \mathbb{Q}_h\boldsymbol{q})_K + \langle v_0, \boldsymbol{q}\cdot\boldsymbol{n}\rangle_{\partial K}\right) \\
&= \sum_{K\in\mathcal{T}_h} \left((v_0, \nabla\cdot\mathbb{Q}_h\boldsymbol{q})_K - \langle v_0, (\mathbb{Q}_h\boldsymbol{q})\cdot\boldsymbol{n}\rangle_{\partial K} + \langle v_0, \boldsymbol{q}\cdot\boldsymbol{n}\rangle_{\partial K}\right), \\
&= \sum_{K\in\mathcal{T}_h} \left(-(\nabla_w v_0, \mathbb{Q}_h\boldsymbol{q})_K + \langle v_0 - v_b, (\boldsymbol{q} - \mathbb{Q}_h\boldsymbol{q})\cdot\boldsymbol{n}\rangle_{\partial K}\right) + \sum_{e\in\mathcal{E}_{\mathcal{I}}} \langle [\![v_b]\!], \boldsymbol{q}\cdot\boldsymbol{n}_e\rangle_e,
\end{aligned}
$$

where we have used the continuity of $\boldsymbol{q}\cdot\boldsymbol{n}_e$ across each element edge/face and the fact that $v_b = 0$ on $\partial\Omega$, and the following two equations

$$
\sum_{K\in\mathcal{T}_h} \langle v_b, \boldsymbol{q}\cdot\boldsymbol{n}\rangle_{\partial K} = \sum_{e\in\mathcal{E}_{\mathcal{I}}} \langle [\![v_b]\!], \boldsymbol{q}\cdot\boldsymbol{n}_e\rangle_e
$$

$$
(v_0, \nabla\cdot\mathbb{Q}_h\boldsymbol{q})_K = -(\nabla_w v, \mathbb{Q}_h\boldsymbol{q})_K + \langle v_b, (\mathbb{Q}_h\boldsymbol{q})\cdot\boldsymbol{n}\rangle_{\partial K}.
$$

Since $h < 1$ and $\beta_0 \geq 1$ and estimate (28), we can get

$$
\begin{aligned}
\|v_0\|^2 &\leq \|\nabla_w v\|\|\mathbb{Q}_h\boldsymbol{q}\| + Ch\|\Psi\|_2\, |||\, v\, ||| + \left(\sum_{e\in\mathcal{E}_{\mathcal{I}}} \|\boldsymbol{q}\cdot\boldsymbol{n}_e\|_e^2\right)^{1/2} \left(\sum_{e\in\mathcal{E}_{\mathcal{I}}} \|[\![v]\!]\|_e^2\right)^{1/2} \\
&\leq C\, |||\, v\, |||\, \|\Psi\|_2 + C\left(\sum_{K\in\mathcal{T}_h} \left(h_K^{-1}\|\boldsymbol{q}\|_K^2 + h_K\|\Psi\|_K^2\right)\right)^{1/2} \left(\sum_{e\in\mathcal{E}_{\mathcal{I}}} \|[\![v]\!]\|_e^2\right)^{1/2} \\
&\leq C(1 + h^{\frac{\beta_0(d-1)-1}{2}})\, |||\, v\, |||\, \|\Psi\|_2 \leq C\, |||\, v\, |||\, \|v_0\|,
\end{aligned}
$$

where we have used the regularity of Poisson's equation. This completes the proof. □

By using (19), we can arrive at

$$\begin{aligned}\|v\|_h^2 &\le \sum_{K\in\mathcal{T}_h}\left(\|v_0\|_K^2 + h_K\|Q_bv_0 - v_b\|_{\partial K}^2 + h_K\|Q_bv_0\|_{\partial K}^2\right)\\ &\le C\sum_{K\in\mathcal{T}_h}\left(\|v_0\|_K^2 + h_K^{-1}\|Q_bv_0 - v_b\|_{\partial K}^2 + h_K^2\|\nabla v_0\|_K^2\right)\\ &\le C\left(\|v_0\|^2 + |||v|||^2\right)\\ &\le C\,|||\,v\,|||^2\,.\end{aligned} \tag{20}$$

Integrating inequalities (17)(18)(20) leads to

$$C_1h^d \le \frac{a_s(v,v)}{|\boldsymbol{\eta}|^2} = \frac{|||v|||^2}{|\boldsymbol{\eta}|^2} \le C_2h^{-\beta_0(d-1)-1+d}. \tag{21}$$

Consequently, the spectral condition number of the stiffness matrix A_h arising from OPWG Algorithm 1 can be estimated as follows

$$\chi_{sp}(A_h) := \frac{\lambda_{\max}(A_h)}{\lambda_{\min}(A_h)} \le \frac{C_2}{C_1}h^{-\beta_0(d-1)-1}. \tag{22}$$

A simple preconditioner. Conditioning result (22) indicates that the OPWG scheme may produce an ill-conditioned linear system with a large value of β_0 selected. Therefore, we propose the following symmetric positive-definite bilinear form

$$\begin{aligned}b(w,v) := &\sum_{K\in\mathcal{T}_h}\left(h_K^{-d}(w_0,v_0)_K + h_K^{-(d-1)}\langle w_b, v_b\rangle_{\partial K}\right)\\ &+\sum_{e\in\mathcal{E}_\mathcal{I}} h_e^{2-d}|e|^{-\beta_0}\int_e [\![w_b]\!][\![v_b]\!],\end{aligned} \tag{23}$$

which can be used to construct suitable *block diagonal* preconditioner as penalty parameter β_0 increases. The linear system B_h arising from $b(\cdot,\cdot)$ is symmetric positive-definite and is easily invertible. Moreover, B_h can be computed in advance, because it can be assembled simultaneously with the stiffness matrix (cf. Remark 1).

It is observed from the following theoretical results and numerical results in Section 7 that the preconditioner is highly efficient in reducing fast-increasing condition number. To this end, we give the main result of this section.

Theorem 4.4. *Assume $\mathcal{T}_h$ is a shape regular triangulation defined in [18] and satisfies quasi-uniformity (12). The preconditioning matrix B_h arising from bilinear form $b(\cdot,\cdot)$ gives*

$$\chi_{sp}(B_h^{-1}A_h) = \frac{\lambda_{max}(B_h^{-1}A_h)}{\lambda_{min}(B_h^{-1}A_h)} \le Ch^{-2}. \tag{24}$$

Proof. We shall limit a test function v in V_h^0. On the one hand, by using Lemma 4.1 and (19) we have

$$b(v,v) \le C\sum_{K\in\mathcal{T}_h} h_K^{-d}\left(\|v_0\|_K^2 + h_K\|v_0\|_{\partial K}^2 + h_K\|Q_bv_0 - v_b\|_{\partial K}^2\right) + \sum_{e\in\mathcal{E}_\mathcal{I}} h_e^{2-d}|e|^{-\beta_0}\|[\![v_b]\!]\|_e^2$$

$$
\begin{aligned}
&\leq C\sum_{K\in\mathcal{T}_h} h_K^{-d}\left(\|v_0\|_K^2 + h_K\|Q_b v_0 - v_b\|_{\partial K}^2\right) + \sum_{e\in\mathcal{E}_\mathcal{I}} h_e^{2-d}|e|^{-\beta_0}\|[\![v_b]\!]\|_e^2 \\
&\leq Ch^{-d}\|v_0\|^2 + Ch^{2-d}\|v\|_{1,h}^2 \\
&\leq Ch^{-d}\,|\!|\!|\, v|\!|\!|^2 \\
&\leq Ch^{-d}a_s(v,v).
\end{aligned}
$$

Hence, by the Rayleigh quotient formula, we have

$$
\lambda_{\min}(B_h^{-1}A_h) = \min_{v\in V_h^0\setminus\{0\}} \frac{a_s(v,v)}{b(v,v)} \geq Ch^d. \tag{25}
$$

On the other hand, we have for $d=2,3$

$$
\begin{aligned}
a_s(v,v) &\leq C\|v\|_{1,h}^2 \\
&\leq C\sum_{K\in\mathcal{T}_h}\left(\|\nabla v_0\|_K^2 + h_K^{-1}\|v_0\|_{\partial K}^2 + h_K^{-1}\|v_b\|_{\partial K}^2\right) + C\sum_{e\in\mathcal{E}_\mathcal{I}}|e|^{-\beta_0}\|[\![v_b]\!]\|_e^2 \\
&\leq C\sum_{K\in\mathcal{T}_h}\left(h_K^{-2}\|v_0\|_K^2 + h_K^{-1}\|v_0\|_{\partial K}^2 + h_K^{-1}\|v_b\|_{\partial K}^2\right) + C\sum_{e\in\mathcal{E}_\mathcal{I}}|e|^{-\beta_0}\|[\![v_b]\!]\|_e^2 \\
&\leq C\sum_{K\in\mathcal{T}_h} h_K^{d-2}\left(h_K^{-d}\|v_0\|_{\partial K}^2 + h_K^{-(d-1)}\|v_b\|_{\partial K}^2\right) + Ch^{d-2}\sum_{e\in\mathcal{E}_\mathcal{I}} h_e^{2-d}|e|^{-\beta_0}\|[\![v_b]\!]\|_e^2 \\
&\leq Ch^{d-2}b(v,v),
\end{aligned}
$$

and hence, by the Rayleigh quotient formula, we have

$$
\lambda_{\max}(B_h^{-1}A_h) = \max_{v\in V_h^0\setminus\{0\}} \frac{a_s(v,v)}{b(v,v)} \leq Ch^{d-2}. \tag{26}
$$

The estimate (24) follows from (25) and (26). □

Remark 1. For a shape regular triangulation, the following two norms are equivalent by the proof of Lemma 4.2

$$
\begin{aligned}
h_K^{-d}\|v_0\|_K^2 &\sim |\boldsymbol{\eta}_{0,K}|^2, \\
h_K^{-(d-1)}\|v_b\|_{\partial K}^2 &\sim |\boldsymbol{\eta}_{b,K}|^2,
\end{aligned}
$$

where $\boldsymbol{\eta}_{0,K}$ and $\boldsymbol{\eta}_{b,K}$ are restriction on $K\in\mathcal{T}_h$ of $\boldsymbol{\eta}_0$ and $\boldsymbol{\eta}_b$, respectively. Thus, by numbering the basis functions of V_h properly, we have

$$
b(v,v) = \boldsymbol{\eta}^t B_h\boldsymbol{\eta} \sim \boldsymbol{\eta}^t\tilde{B}_h\boldsymbol{\eta} \text{ with } \tilde{B}_h = I + blkdiag(0, M_1, M_2, \cdots, M_{N_e}),
$$

where I is the unit matrix, $blkdiag(\cdot)$ denotes a matrix with $0, M_i$ be its principal diagonal blocks, M_i denotes a matrix arising from the penalty term in (23) on each edge $e_i\in\mathcal{E}_I$, and N_e is the the number of basis functions defined on the interior edges. In practice, one can replace the precondition matrix B_h by $\tilde{B}_h$ to save computational cost.

5. **Error analysis.** The goal of this section is to establish some error estimates for the OPWG finite element solution u_h arising from Algorithm 1. The error will be measured in two natural norms: the triple-bar norm as defined in (8) and the standard L^2 norm.

We shall give some key estimates for error analysis.

Lemma 5.1. *Assume $\mathcal{T}_h$ to be a shape regular partitions. Then for any $u \in H^{k+1}(K)$ and $v = (v_0, v_b) \in V_h$, we have*

$$|s(Q_h u, v)| \le Ch^k \|u\|_{k+1} \, |\!|\!| v |\!|\!|, \tag{27}$$

$$|\ell_u(v)| \le Ch^k \|u\|_{k+1} \, |\!|\!| v |\!|\!|, \tag{28}$$

$$\left| \sum_{e\in\mathcal{E}_\mathcal{I}} \langle A\nabla u \cdot \boldsymbol{n}_e, [\![v_b]\!] \rangle_e \right| \le Ch^{\frac{\beta_0(d-1)-1}{2}} \|u\|_{k+1} \, |\!|\!| v |\!|\!|, \tag{29}$$

where $\ell_u(v) = \sum_{K\in\mathcal{T}_h} \langle A(\nabla u - \mathbb{Q}\nabla u)\cdot \boldsymbol{n}, v_0 - v_b\rangle_{\partial K}$.

Proof. The proof for the first two estimates can be found in [8], which could be used to show the equivalence between $\|\cdot\|_{1,h}$ and $|\!|\!|\cdot|\!|\!|$ in V_h, so we only prove the third inequality. Following the trace inequality (13), and $|e| \le C_e h_K^{d-1}, K \in \mathcal{T}_h$, we have

$$\begin{aligned}
\left| \sum_{e\in\mathcal{E}_\mathcal{I}} \langle A\nabla u \cdot \boldsymbol{n}_e, [\![v_b]\!] \rangle_e \right| &\le \sum_{e\in\mathcal{E}_\mathcal{I}} \|A\nabla u\cdot \boldsymbol{n}_e\|_e \|[\![v_b]\!]\|_e \\
&\le C \sum_{e\in\mathcal{E}_\mathcal{I}} \left(h_K^{-\frac{1}{2}} \|\nabla u\|_{K_1^e\cup K_2^e} + h_K^{\frac{1}{2}} \|\nabla^2 u\|_{K_1^e\cup K_2^e} \right) \|[\![v_b]\!]\|_e \\
&\le C h^{-\frac{1}{2}} \|u\|_{k+1} h^{\frac{\beta_0}{2}(d-1)} \, |\!|\!| v |\!|\!| \\
&\le C h^{\frac{\beta_0(d-1)-1}{2}} \|u\|_{k+1} \, |\!|\!| v |\!|\!| \, .
\end{aligned}$$

□

Now, we give an *error equation* for the over-penalized weak Galerkin scheme (7).

Lemma 5.2. *Let $e_h = (e_0, e_b) = (Q_0 u - u_0, Q_b u - u_b)$. Then, for any $v \in V_h^0$ we have*

$$a_s(e_h, v) = s(Q_h u, v) + \ell_u(v) + \sum_{e\in\mathcal{E}_\mathcal{I}} \langle A\nabla u\cdot \boldsymbol{n}_e, [\![v_b]\!]\rangle_e \tag{30}$$

Proof. Testing (1) with v_0 of $v = (v_0, v_b) \in V_h^0$ leads to

$$\begin{aligned}
(f, v_0) = -(\nabla\cdot A\nabla u, v_0) &= \sum_{K\in\mathcal{T}_h} \left((A\nabla u, \nabla v_0)_K - \langle A\nabla u\cdot \boldsymbol{n}, v_0\rangle_{\partial K} \right) \\
&= (A\nabla u, \nabla v_0) - \sum_{K\in\mathcal{T}_h} \langle A\nabla u \cdot \boldsymbol{n}, v_0 - v_b\rangle_{\partial K} - \sum_{e\in\mathcal{E}_\mathcal{I}} \langle A\nabla u\cdot \boldsymbol{n}_e, [\![v_b]\!]\rangle_e.
\end{aligned} \tag{31}$$

For $u \in H^1(K)$ and $v \in V_h$, we have

$$\begin{aligned}
(A\nabla_w Q_h u, \nabla_w v)_K &= (A\mathbb{Q}_h \nabla u, \nabla_w v)_K \\
&= (A\nabla u, \nabla v_0)_K - \langle (A\mathbb{Q}_h\nabla u)\cdot \boldsymbol{n}, v_0 - v_b\rangle_{\partial K}.
\end{aligned} \tag{32}$$

Combining above two equations, we have

$$\begin{aligned}
(A\nabla_w Q_h u, \nabla_w v) = &\sum_{K\in\mathcal{T}_h} \langle A(\nabla u - \mathbb{Q}_h \nabla u)\cdot \boldsymbol{n}, v_0 - v_b\rangle_{\partial K} \\
&+ (f, v_0) + \sum_{e\in\mathcal{E}_\mathcal{I}} \langle A\nabla u\cdot \boldsymbol{n}_e, [\![v_b]\!]\rangle_e.
\end{aligned}$$

Adding the stabilizer to above equation, and noting that $\sum_{e\in\mathcal{E}_\mathcal{I}} |e|^{-\beta_0}\langle [\![Q_b u]\!], [\![v_b]\!]\rangle_e = 0$, and subtracting (7) from it, we arrive at

$$\begin{aligned} a(e_h, v) = &\sum_{K\in\mathcal{T}_h} h_K^{-1}\langle Q_b(Q_0 u) - Q_b u, Q_b v_0 - v_b\rangle_{\partial K} \\ &+ \sum_{K\in\mathcal{T}_h} \langle A(\nabla u - \mathbb{Q}_h \nabla u)\cdot \boldsymbol{n}, v_0 - v_b\rangle_{\partial K} \\ &+ \sum_{e\in\mathcal{E}_\mathcal{I}} \langle A\nabla u\cdot \boldsymbol{n}_e, [\![v_b]\!]\rangle_e, \end{aligned}$$

which completes the proof of (30). □

Now, we are ready to show the first key error estimate by using Lemma 5.1.

Theorem 5.3. *Let $u_h \in V_h$ be the over-penalized weak Galerkin finite element solution of the problems (1)(2) arising from (7). Assume the exact solution $u \in H^{k+1}(\Omega)$ with $k \geq 1$. Then there exists a constant C independent of h such that*

$$|\!|\!| e_h |\!|\!| \leq C\left(h^k + h^{\frac{\beta_0(d-1)-1}{2}}\right)\|u\|_{k+1,\Omega}. \tag{33}$$

This estimate is optimal if $\beta_0(d-1) \geq 2k+1$.

In the rest of the section, we shall derive an optimal order error estimate for the over-penalized weak Galerkin finite element scheme (7) in the usual L^2-norm by using duality argument as was commonly employed in the standard Galerkin finite element methods. To this end, we consider a dual problem that seeks $w \in H_0^1(\Omega)$ satisfying

$$-\nabla\cdot(A\nabla w) = e_0, \quad \text{in } \Omega. \tag{34}$$

Assume that the above dual problem has the usual H^2-regularity, which means that there exists a constant C such that

$$\|w\|_2 \leq C\|e_0\|. \tag{35}$$

Theorem 5.4. *Let $u_h \in V_h$ be the over-penalized weak Galerkin finite element solution of the problems (1)(2) arising from (7). Assume the exact solution satisfies $u \in H^{k+1}(\Omega)$ with $k \geq 1$. In addition, assume that the dual problem (34) has the usual H^2-regularity. Then there exists a constant C independent of h such that*

$$\|e_0\| \leq Ch^{k+1}\|u\|_{k+1} + C\left(h + h^{\frac{\beta_0(d-1)-1}{2}}\right)|\!|\!| e_h |\!|\!| . \tag{36}$$

This estimate is optimal if $\beta_0(d-1) \geq 2k+1$.

Proof. By testing (34) with e_0 we obtain

$$\begin{aligned} \|e_0\|^2 &= \sum_{K\in\mathcal{T}_h} (-\nabla\cdot(A\nabla w), e_0)_K \\ &= \sum_{K\in\mathcal{T}_h} \left((A\nabla w, \nabla e_0)_K - \langle A\nabla w\cdot\boldsymbol{n}, e_0\rangle_{\partial K}\right) \\ &= (A\nabla w, \nabla v_0) - \sum_{K\in\mathcal{T}_h} \langle A\nabla w\cdot \boldsymbol{n}, v_0 - v_b\rangle_{\partial K} - \sum_{e\in\mathcal{E}_\mathcal{I}} \langle A\nabla w\cdot\boldsymbol{n}_e, [\![v_b]\!]\rangle_e. \end{aligned} \tag{37}$$

Setting $u = w$ and $v = e_h$ in (32) yields

$$\begin{aligned} (A\nabla_w Q_h w, \nabla_w e_h)_K &= (A\mathbb{Q}_h\nabla w, \nabla_w e_h)_K \\ &= (A\nabla w, \nabla e_0)_K - \langle (A\mathbb{Q}_h\nabla w)\cdot\boldsymbol{n}, e_0 - e_b\rangle_{\partial K}. \end{aligned} \tag{38}$$

Substituting (38) into (37) gives

$$\|e_0\|^2 = (A\nabla_w Q_h w, \nabla_w e_h) + \ell_w(e_h) - \sum_{e\in\mathcal{E}_\mathcal{I}} \langle A\nabla w \cdot \boldsymbol{n}_e, [\![e_b]\!]\rangle_e. \tag{39}$$

Following from error equation (30), we have

$$\begin{aligned}&(A\nabla_w e_h, \nabla_w Q_h w)\\ =& - s(e_h, Q_h w) + s(Q_h u, Q_h w) + \ell_u(Q_h w) + \sum_{e\in\mathcal{E}_\mathcal{I}} \langle A\nabla u \cdot \boldsymbol{n}_e, [\![Q_b w]\!]\rangle_e.\end{aligned} \tag{40}$$

Note that the fourth term of the right hand side is zero because the jump $[\![Q_b w]\!] = 0$ on each internal edge. Thus, we get from (39)(40)

$$\begin{aligned}\|e_0\|^2 =& - s(e_h, Q_h w) + s(Q_h u, Q_h w)\\ &+ \ell_u(Q_h w) + \ell_w(e_h) - \sum_{e\in\mathcal{E}_\mathcal{I}} \langle A\nabla w \cdot \boldsymbol{n}_e, [\![e_b]\!]\rangle_e.\end{aligned} \tag{41}$$

To this end, we need to estimate the five terms in (41). Analogous to the proof of Theorem 6.4 in [8], Lemma 5.1 and Theorem 5.3, we have

$$\begin{aligned}|s(e_h, Q_h w)| &\le Ch\|w\|_2 \, |\!|\!| e_h |\!|\!| \le Ch\|e_0\| \, |\!|\!| e_h |\!|\!|,\\ |s(Q_h u, Q_h w)| &\le Ch^{k+1}\|u\|_{k+1}\|e_0\|,\\ |\ell_u(Q_h w)| &\le Ch^{k+1}\|u\|_{k+1}\|e_0\|,\\ |\ell_w(e_h)| &\le Ch\|w\|_2 \, |\!|\!| e_h |\!|\!| \le Ch\|e_0\| \, |\!|\!| e_h |\!|\!| .\end{aligned}$$

Especially, by using (29), we have

$$\left|\sum_{e\in\mathcal{E}_\mathcal{I}} \langle A\nabla w \cdot \boldsymbol{n}_e, [\![e_b]\!]\rangle_e\right| \le Ch^{\frac{\beta_0(d-1)-1}{2}}\|w\|_2 \, |\!|\!| e_h |\!|\!| \le Ch^{\frac{\beta_0(d-1)-1}{2}}\|e_0\| \, |\!|\!| e_h |\!|\!| .$$

Thus, the proof is completed by combining above estimates with (41). □

6. **A Schur complement formulation.** One unique feature of weak Galerkin methods is that the globally coupled unknowns are those defined on element boundaries. In this section, we shall develop a practical method to eliminate degree of freedoms corresponding to u_0, and only involves freedoms corresponding to u_b defined on element edges.

Let $u_h \in V_h$ be the solution of the over-penalized weak Galerkin scheme. By restricting test functions, we can split (7) into two following equations:

$$a_s(u_h, v) = (f, v) \qquad \forall v = (v_0, 0) \in V_h^0, \tag{42}$$

$$a_s(u_h, v) = 0 \qquad \forall v = (0, v_b) \in V_h^0. \tag{43}$$

Thus, Algorithm 1 is equivalent to the following scheme: Find $u_h \in V_h$ satisfying $u_b = Q_b g$ on $\partial\Omega$, and the equation (42) and (43).

For given u_b, one can solve equation (42) for the u_0, we denote this dependence as

$$u_0 = u_0(u_b, f),$$

and note that this procedure can be done elementwisely. We further write

$$u_h = u_h(u_b, f) = (u_0(u_b, f), u_b) \in V_h,$$

Thus by the principle of superposition, we have

$$u_h = u_h(u_b, f) = u_h(u_b, 0) + u_h(0, f) \in V_h. \tag{44}$$

Then substituting (44) into (43), we have

$$a_s(u_h(u_b,0),v) = -a_s(u_h(0,f),v) \qquad \forall v=(0,v_b)\in V_h^0. \tag{45}$$

The above equation is nothing but the Schur complement formulation of the over-penalized weak Galerkin method, involving only unknown variables corresponding to u_b defined on the edges of each element. Due to this size reduction, for triangular or rectangular element, the OPWG scheme (with element $(\mathbb{P}_k,\mathbb{P}_{k-1},[\mathbb{P}_{k-1}]^2)$) has fewer degree of freedoms than Discontinuous Galerkin method (with element $\mathbb{P}_k$) as k is large enough. Take any triangular element K , for instance. The dimension of $\mathbb{P}_k$ for DG is $\frac{1}{2}(k+1)(k+2)$, and the dimension of $\mathbb{P}_{k-1}(\partial K)$ for the OPWG is $3k$. This technique makes the over-penalized WG method computationally efficient compared with DG methods.

Remark 2. If $\tilde{B}_h$ named as a preconditioner with the special structure in Remark 1, the original linear system arising from the OPWG scheme 1 becomes

$$\tilde{B}_h^{-1}A_hU = \tilde{B}_h^{-1}F = F,$$

equivalently,

$$\begin{bmatrix} I & \\ & D^{-1}\end{bmatrix}\begin{bmatrix} A_{00} & A_{0b}\\ A_{0b}^T & A_{bb}\end{bmatrix}\begin{bmatrix} U_0\\ U_b\end{bmatrix} = \begin{bmatrix} A_{00} & A_{0b}\\ D^{-1}A_{0b}^T & D^{-1}A_{bb}\end{bmatrix}\begin{bmatrix} U_0\\ U_b\end{bmatrix} = \begin{bmatrix} F_0\\ 0\end{bmatrix},$$

where D is a symmetry positive-definite block-diagonal matrix. Now, one can apply the Schur implementation of matrix $\tilde{B}_h^{-1}A_h$ to solve the above equations. Note that A_{00} is easy to inverse, since this process can be done elementwisely.

7. Numerical experiments.

In this section, we present some numerical results to confirm the theory in Section 5, and some preconditioning results associated with the preconditioner $b(\cdot,\cdot)$ introduced in Section 4. For simplicity, in all examples we take $\Omega=(0,1)\times(0,1)$, and successively-refined triangular meshes are employed.

Example 1. (A model problem with smooth solution.) Consider the problem (1)-(2) with the following analytical solution

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

and the diffusion matrix $A=\begin{bmatrix} x^2+y^2+1 & xy\\ xy & x^2+y^2+1\end{bmatrix}$, which was also investigated in [8].

The meshes are generated in the following ways. First, we partition the square domain into $N\times N$ subsquares uniformly; then we divide each subsquare into two triangles by the diagonal line with a negative slope, completing the construction of uniformly refined triangular meshes. The mesh size h is denoted by $h:=\frac{1}{N}$, where N is an integer number.

For comparison between the OPWG method and the original WG method, we list convergence results of the WG method with element $(\mathbb{P}_k,\mathbb{P}_k,[\mathbb{P}_{k-1}]^d)$ in Table 1. The over-penalized weak Galerkin errors e_h in triple-bar and L^2 norms are listed in the Table 2 with $\beta_0 = 1, 2, 3, 4$ for the lowest order element $(\mathbb{P}_1,\mathbb{P}_0,[\mathbb{P}_0]^2)$. It can be seen that the over-penalized WG solution of linear element is convergent with optimal rate $O(h)$ in triple-bar and $O(h^2)$ in L^2 norms when $\beta_0=3$. This is the same as the WG results in Table 1. Similarly, for quadratic element, the results are listed in the Tables 3, and the optimal rates of convergence are observed

TABLE 1. WG method with element $(\mathbb{P}_k, \mathbb{P}_k, \mathbb{P}_{k-1}^2)$ for Example 1.

h	$k=1$		$k=2$	
	$\|\|\|e_h\|\|\|$	$\|e_0\|$	$\|\|\|e_h\|\|\|$	$\|e_0\|$
1/8	9.9173e-01	5.3131e-02	7.8508e-02	3.2464e-03
1/16	4.9588e-01	1.3272e-02	1.9677e-02	4.0611e-04
1/32	2.4793e-01	3.3169e-03	4.9226e-03	5.0761e-05
1/64	1.2396e-01	8.2914e-04	1.2309e-03	6.3445e-06
Rate.	1.0001	2.0001	1.9997	3.0001

TABLE 2. OPWG with $(\mathbb{P}_1, \mathbb{P}_1, \mathbb{P}_0^2)$ and $\beta_0 = 1, 2, 3, 4$ for Example 1.

h	$\|\|\|e_h\|\|\|$	$\|e_0\|$	$\|\|\|e_h\|\|\|$	$\|e_0\|$
	$\beta_0 = 1$		$\beta_0 = 2$	
1/8	1.3258e+00	7.4931e-02	1.0550e+00	5.9839e-02
1/16	1.0306e+00	3.4203e-02	5.6097e-01	1.5700e-02
1/32	9.4663e-01	2.5074e-02	3.1098e-01	4.3112e-03
1/64	9.2663e-01	2.3053e-02	1.8215e-01	1.2815e-03
Rate.	0.0308	0.1212	0.7717	1.7503
	$\beta_0 = 3$		$\beta_0 = 4$	
1/8	1.0019e+00	5.7508e-02	9.9304e-01	5.7135e-02
1/16	5.0137e-01	1.4381e-02	4.9628e-01	1.4278e-02
1/32	2.5075e-01	3.5954e-03	2.4804e-01	3.5682e-03
1/64	1.2539e-01	8.9887e-04	1.2399e-01	8.9191e-04
Rate.	0.9998	2.	1.0003	2.0002

TABLE 3. OPWG with $(\mathbb{P}_2, \mathbb{P}_2, \mathbb{P}_1^2)$ and $\beta_0 = 2, 3, 4, 5$ for Example 1.

h	$\|\|\|e_h\|\|\|$	$\|e_0\|$	$\|\|\|e_h\|\|\|$	$\|e_0\|$
	$\beta_0 = 2$		$\beta_0 = 3$	
1/8	1.7917e+00	1.0256e-01	8.4595e-01	2.0487e-02
1/16	1.4060e+00	5.7784e-02	4.4264e-01	5.3334e-03
1/32	1.0580e+00	3.0997e-02	2.2464e-01	1.3530e-03
1/64	7.7468e-01	1.6118e-02	1.1291e-01	3.4023e-04
Rate.	0.4497	0.9435	0.9924	1.9916
	$\beta_0 = 4$		$\beta_0 = 5$	
1/8	3.6348e-01	4.6229e-03	1.6594e-01	3.2492e-03
1/16	1.2933e-01	5.9131e-04	4.1839e-02	4.0604e-04
1/32	4.5730e-02	7.4938e-05	1.0494e-02	5.0757e-05
1/64	1.6162e-02	9.4394e-06	2.6274e-03	6.3443e-06
Rate.	1.5005	2.9889	1.9979	3.0001

numerically for each norm when $\beta_0 = 5$. The numerical results are in reasonable agreement with the related theory.

It can be seen from the conditioning results in Section 4, the essential weakness of the OPWG is the relatively fast-increasing condition number compared to general WG methods. One remedy is to utilize an appropriate preconditioner along with a linear system generated by the over-penalized weak Galerkin method. At the end of Section 4, a preconditioner is proposed, and condition number has been analyzed. Next we will check its performance. For Example 1, in Table 4, we compare the condition numbers of discrete linear systems with/without preconditioning, by linear ($k = 1$) and quadratic ($k = 2$) elements with optimal penalty parameters, respectively. Numerical results indicate the preconditioner works effectively, coinciding with the theoretical result. Thus, the problem of fast-increasing condition number can be numerically overcome.

TABLE 4. Comparison of condition number with optimal penalty parameters.

h	Without preconditioning		Block-diagonal preconditioning	
	$k=1,\beta_0=3$	$k=2,\beta_0=5$	$k=1,\beta_0=3$	$k=2,\beta_0=5$
1/4	1.8272e+04	2.1075e+04	2.6265e+02	4.3079e+04
1/8	1.6957e+05	6.9533e+05	1.0135e+03	1.9257e+05
1/16	2.1585e+06	3.4759e+07	4.0481e+03	8.0866e+05
1/32	3.2471e+07	2.0653e+09	1.6228e+04	3.3427e+06
1/64	5.1218e+08	1.2974e+11	6.4995e+04	1.3570e+07
Order	-3.9794	-5.9731	-2.0018	-2.0213

In practical application, storage, consumed time and precision are three important issues deserving careful considerations. Mu etc. [8] proposed element $(\mathbb{P}_k, \mathbb{P}_{k-1}, \mathbb{P}_{k-1}^2)$ which is one of the best kind of elements with the least degree of freedoms (dof). By using such elements, the errors and condition numbers with the optimal penalty parameter $\beta_0 = 3, 5$ are reported in Tables 5 and 6, respectively. Fewer dof equally mean less storage, but does not means better precision under the same mesh, which can be seen easily by comparing with the results in Table 5. Besides, the preconditioner still performs really well. In practice, one can choose both kinds of elements to address the practical needs.

TABLE 5. Errors and condition numbers for Example 1 with $(\mathbb{P}_1, \mathbb{P}_0, \mathbb{P}_0^2)$ and $\beta_0 = 3$.

h	$\|\|\|e_h\|\|\|$	$\|e_0\|$	Cond.	Pre. Cond.
1/8	1.5226e+00	8.7459e-02	6.6551e+03	2.1055e+03
1/16	7.7405e-01	2.2035e-02	8.1739e+04	7.9132e+03
1/32	3.8921e-01	5.5253e-03	1.2109e+06	3.1176e+04
1/64	1.9499e-01	1.3832e-03	1.9007e+07	1.2425e+05
Rate.	0.9971	1.9980	-3.9724	-1.9947

TABLE 6. Errors and condition numbers for Example 1 with $(\mathbb{P}_2, \mathbb{P}_1, \mathbb{P}_1^2)$ and $\beta_0 = 5$.

h	$\|\|\|e_h\|\|\|$	$\|e_0\|$	Cond.	Pre. Cond.
1/8	1.6622e-01	3.3354e-03	3.5419e+05	1.0229e+05
1/16	4.1912e-02	4.1731e-04	1.9810e+07	4.2176e+05
1/32	1.0513e-02	5.2184e-05	1.2227e+09	1.7166e+06
1/64	2.6321e-03	6.5231e-06	7.7586e+10	6.9742e+06
Rate.	1.9979	3.0000	-5.9877	-2.0225

Example 2. (A model problem with a corner singularity.) We consider Laplace equation defined on the unit square with the following exact solution

$$u(x, y) = x(1-x)y(1-y)r^{-2+\alpha},$$

where $r = \sqrt{x^2 + y^2}$ and $\alpha \in (0, 1]$. As we known

$$u \in H_0^1(\Omega) \cap H^{1+\alpha-\epsilon}(\Omega) \text{ and } u \notin H^{1+\alpha}(\Omega),$$

where ϵ is any small, but a positive number.

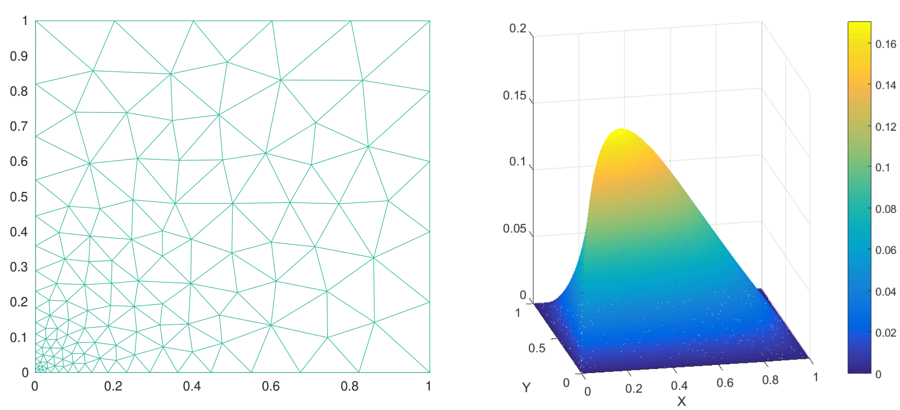

FIGURE 1. Initial mesh (*Left*) and the OPWG ($\beta_0 = 3$) solution on the finest mesh (*Right*) for Example 2.

The model problem has a corner singularity at the origin point. Thus we choose a locally-refined initial mesh as shown in Figure 1 (*Left*).

Taking linear element with $\beta_0 = 3$ (the optimal exponent), we test the model problem in both cases $\alpha = 0.5$ and $\alpha = 0.25$. Figure 1 (*Right*) is the surface plot of the OPWG solution with $\alpha = 0.5$ on the finest mesh. Numerical results are listed in Table 7, where the term 'Pre. Cond.' means the condition number after preconditioning, and the associated convergence rates are plotted in Figure 2. It can be observed that the OPWG solution converges with rates given by $O(N^{-\frac{\alpha}{2}})$ in energy norm; better convergence rates of error in L^2-norm than the theory predicted $O(N^{-\frac{1+\alpha}{2}})$. Moreover, the preconditioner for the singularity problem performs still well.

TABLE 7. OPWG with $(\mathbb{P}_1, \mathbb{P}_1, \mathbb{P}_0^2)$ and optimal penalty parameter for Example 2.

dof.	$\alpha = 0.5$		$\alpha = 0.25$		Condition Number	
	$\|\|\|e_h\|\|\|$	$\|e_0\|$	$\|\|\|e_h\|\|\|$	$\|e_0\|$	Cond.	Pre. Cond.
2.0880e+3	4.2303e-1	7.0060e-2	9.6614e-1	7.3906e-2	1.7335e+6	1.1357e+3
8.3520e+3	2.7461e-1	1.8365e-2	7.8292e-1	2.2116e-2	2.4882e+7	4.5578e+3
3.3408e+4	1.8503e-1	4.7471e-3	6.4655e-1	7.3419e-3	3.8905e+8	1.8270e+4
1.3363e+5	1.3025e-1	1.2371e-3	5.5838e-1	2.8274e-3	6.1854e+9	7.3140e+4
5.3453e+5	9.2125e-2	3.2954e-4	4.6994e-1	1.1394e-3	9.8828e+10	2.9267e+5

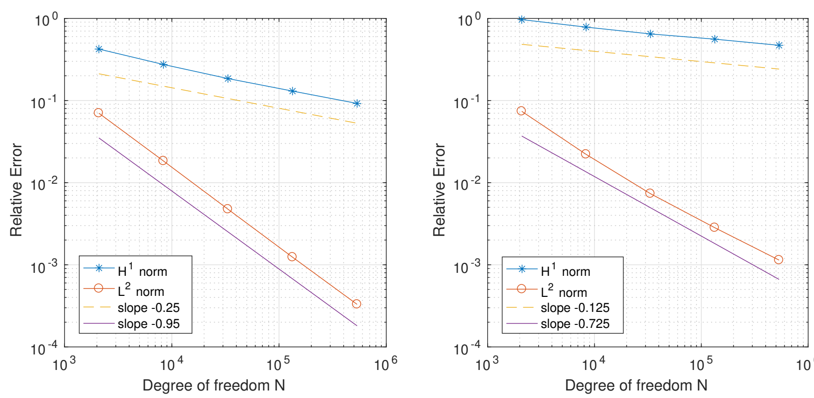

FIGURE 2. Convergence rates of the OPWG ($\beta_0 = 3$) solutions against degree of freedoms for different values of α in Example 2. (*Left*) $\alpha = 0.5$; (*Right*) $\alpha = 0.25$.

Acknowledgments. The research of the second author was supported in part by the Natural Science Foundation of Gansu Province, China (Grant 18JR3RA290). The research of Zhao was supported in part by the US National Science Foundation grant DMS-1812930.

REFERENCES

[1] D. N. Arnold, F. Brezzi, B. Cockburn and L. D. Marini, Unified analysis of discontinuous Galerkin methods for elliptic problems, *SIAM J. Numer. Anal.*, **39** (2002), 1749–1779.
[2] F. Brezzi, J. Douglas Jr. and L. D. Marini, Two families of mixed finite elements for second order elliptic problems, *Numer. Math.*, **47** (1985), 217–235.
[3] B. Li and X. Xie, A two-level algorithm for the weak Galerkin discretization of diffusion problems, *J. Comput. Appl. Math.*, **287** (2015), 179–195.
[4] K. Liu, L. Song and S. Zhou, An over-penalized weak Galerkin method for second-order elliptic problems, *J. Comput. Math.*, **36** (2018), 866–880.
[5] L. Mu, J. Wang, Y. Wang and X. Ye, A computational study of the weak Galerkin method for second-order elliptic equations, *Numer. Algor.*, **63** (2012), 753–777.
[6] L. Mu, J. Wang, G. Wei, X. Ye and S. Zhao, Weak Galerkin methods for second order elliptic interface problems, *J. Comput. Phys.*, **250** (2013), 106–125.
[7] L. Mu, J. Wang and X. Ye, A new weak Galerkin finite element method for the Helmholtz equation, *IMA J. Numer. Anal.*, **35** (2015), 1228–1255.

[8] L. Mu, J. Wang and X. Ye, A weak Galerkin finite element method with polynomial reduction, *J. Comput. Appl. Math.*, **285** (2015), 45–58.

[9] L. Mu, J. Wang and X. Ye, Weak Galerkin finite element methods on polytopal meshes, *Int. J. Numer. Anal. Model.*, **12** (2015), 31–53.

[10] L. Mu, J. Wang, X. Ye and S. Zhang, A weak Galerkin finite element method for the Maxwell equations, *J. Sci. Comput.*, **65** (2015), 363–386.

[11] A. Quarteroni and V. Alberto, *Numerical Approximation of Partial Differential Equations*, Springer Series in Computational Mathematics, 23, Springer, Berlin, Heidelberg, 1994.

[12] P.-A. Raviart and J.-M. Thomas, A mixed finite element method for 2-nd order elliptic problems, in Galligani I., Magenes E. (eds) Mathematical Aspects of Finite Element Methods. Lecture Notes in Mathematics, vol 606, Springer, Berlin, Heidelberg.

[13] B. Riviere, *Discontinuous Galerkin Methods for Solving Elliptic and Parabolic Equations: Theory and Implementation*, Society for Industrial and Applied Mathematics, Philadelphia, PA, 2008.

[14] L. Song, K. Liu and S. Zhao, A weak Galerkin method with an over-relaxed stabilization for low regularity elliptic problems, *J. Sci. Comput.*, **71** (2017), 195–218.

[15] L. Song, S. Zhao and K. Liu, A relaxed weak Galerkin method for elliptic interface problems with low regularity, *Appl. Numer. Math.*, **128** (2018), 65–80.

[16] C. Wang and J. Wang, An efficient numerical scheme for the biharmonic equation by weak Galerkin finite element methods on polygonal or polyhedral meshes, *Comput. Math. Appl.*, **68** (2014), 2314–2330.

[17] J. Wang and X. Ye, A weak Galerkin finite element method for second-order elliptic problems, *J. Comput. Appl. Math.*, **241** (2013), 103–115.

[18] J. Wang and X. Ye, A weak Galerkin mixed finite element method for second order elliptic problems, *Math. Comput.*, **83** (2014), 2101–2126.

[19] J. Wang and X. Ye, A weak Galerkin finite element method for the Stokes equations, *Adv. Comput. Math.*, **42** (2016), 155–174.

[20] Q. Zhai, X. Ye, R. Wang and R. Zhang, A weak Galerkin finite element scheme with boundary continuity for second-order elliptic problems, *Comput. Math. Appl.*, **74** (2017), 2243–2252.

Received November 2019; revised February 2020.

E-mail address: liukf17@lzu.edu.cn
E-mail address: song@lzu.edu.cn
E-mail address: szhao@ua.edu