INTERNATIONAL JOURNAL OF
NUMERICAL ANALYSIS AND MODELING, *SERIES B*
Volume 5, Number 4, Pages 299–316

BABUŠKA'S PENALTY METHOD FOR INHOMOGENEOUS DIRICHLET PROBLEM: ERROR ESTIMATES AND MULTIGRID ALGORITHMS

THIRUPATHI GUDI

Abstract. This article is two fold. Firstly, we derive optimal order *a priori* error estimates for Babuška's penalty method applied to inhomogeneous Dirichlet problem. Secondly, we derive convergence of W-cycle and V-cycle multigrid algorithms for the resulting system. To this end, a simple pre-conditioner is introduced to remedy the ill-condition due to over-penalty.

Key words. Finite element, multigrid, penalty, pre-conditioner.

1. Introduction

We consider the model problem of finding $u \in H^1(\Omega)$ such that $u|_{\partial\Omega} = g$ and

$$\int_\Omega \nabla u \cdot \nabla v \, dx = \int_\Omega f v \, dx \quad \forall v \in H_0^1(\Omega). \tag{1}$$

Here $f \in L_2(\Omega)$ and $g \in C^0(\partial\Omega)$ is such that there is a $\psi \in H^2(\Omega)$ with $\psi|_{\partial\Omega} = g$. We consider $\Omega \subset \mathbb{R}^2$ to be a bounded convex polygonal domain. Let the boundary $\partial\Omega = \cup_{i=1}^N \bar{\Gamma}_i$, where Γ_is are pairwise disjoint (open) line segments on $\partial\Omega$. From the regularity theory of elliptic problems [18], it holds that $u \in H^2(\Omega)$ and

$$\|u\|_{H^2(\Omega)} \le C \left(\|f\| + \sum_{i=1}^N \|g\|_{H^{3/2}(\Gamma_i)} \right), \tag{2}$$

hereafter $\|\cdot\|$ denotes the standard $L_2(\Omega)$ norm. Based on (1), it is well known that the finite element method with interpolated boundary condition is defined as follows: Find $u_h \in V_h$ with $u_h|_{\partial\Omega} = g_h$ such that

$$\int_\Omega \nabla u_h \cdot \nabla v \, dx = \int_\Omega f v \, dx \quad \forall v \in V_h^0, \tag{3}$$

where g_h is an approximation of g and $V_h^0 \subset H_0^1(\Omega)$ is a finite element subspace. In practice, g_h is considered to be a nodal interpolation of g in $V_h|_{\partial\Omega}$, where $V_h \subset H^1(\Omega)$ is a finite element subspace. However it is shown in [4] that the error estimates for (3) are better when g_h is chosen to be the L_2-projection of g onto $V_h|_{\partial\Omega}$. In particular when g_h is chosen to be a nodal interpolation, the L_2 error estimate requires g to be a piecewise H^2 function on $\partial\Omega$, see [4, Theorem 7.1]. Perhaps this estimate may not be improved. On the other-hand a mild disadvantage associated with the choice that g_h is L_2-projection, we need to solve a system before we solve for the solution u_h.

An alternative method due to Babuška [1] is to pose the boundary condition weakly by penalty. Namely, find $u_h \in V_h$ such that

$$a_h(u_h, v) = L_h(v) \quad \forall v \in V_h, \tag{4}$$

Received by the editors October 12, 2014.
1991 *Mathematics Subject Classification.* 65N30, 65N15.
The author is supported partly by the DST Fast Track Project and partly by the UGC Center for Advanced Study.

where

$$a_h(w,v) = \int_\Omega \nabla w \cdot \nabla v \, dx + \int_{\partial\Omega} \frac{1}{h^2} wv \, ds, \tag{5}$$

and

$$L_h(v) = \int_\Omega fv \, dx + \int_{\partial\Omega} \frac{1}{h^2} gv \, ds. \tag{6}$$

In [1], error estimates that are almost optimal order up to an arbitrary ϵ were derived for homogeneous Dirichlet boundary value problem. In this article, we improve upon the error estimates in [1] and derive optimal order error estimates for general inhomogeneous Dirichlet problem. To this end, we construct an interpolation which is partially based on a projection. We derive optimal order approximation properties for this interpolation and use as a key in our *a priori* error analysis. On the other hand since the method in (4) uses high penalty on posing the boundary condition weakly, the condition number results in the order h^{-3}. We construct a simple diagonal pre-conditioner to remedy this ill-condition and present W-cycle and V-cycle multigrid algorithms for the resulting preconditioned system. We prove optimal order convergence for the multigrid algorithms following the theory developed in [2, 6, 7, 20, 21, 11, 12] for W-cycle algorithm and the additive multigrid theory in [8] for V-cycle algorithm. The convergence analysis of V-cycle algorithm is established under a mild assumption on the initial triangulation of the domain. Namely, we require that each interior vertex in the initial mesh is connected by at most six edges. The results in this article will be useful in solving the feedback boundary control problems (see e.g. [17]) and also in the analysis of over penalized discontinuous Galerkin and non-conforming methods. There is a plenty of work on multigrid methods for finite element methods, we refer to a few articles and monographs for the references and related work [3, 5, 14, 15, 19, 23, 24, 25, 26].

Remark 1.1. It is well known that there is another alternative method due to Nitsche [22] posing the Dirichlet boundary conditions weakly. However this approach requires tuning of a suitable penalty parameter for obtaining a stable formulation. We therefore restrict ourself in this article to the penalty method by Babuška.

The rest of the article is organized as follows. In section 2, we present improved error estimates for the penalty method. In section 3, we construct a simple pre-conditioner and set up multigrid methods. In sections 4 and 5, we present convergence analysis of W-cycle and V-cycle algorithms, respectively. In section 6, we present some numerical experiments demonstrating the performance of the method and multigrid algorithms. Finally, we conclude the article in section 7.

2. A priori Analysis of Penalty Method

In this section, we derive optimal order *a priori* error estimates for the penalty method (4). Let $\mathcal{T}_h$ be a regular quasi-uniform simplicial triangulation of Ω. Let $h_T = \sqrt{|T|}$, where $|T|$ is the area of T. Set $h = \max\{h_T : T \in \mathcal{T}_h\}$. The P_1 finite element space V_h is defined as

$$V_h = \{v \in C^0(\bar{\Omega}) : v|_{T \in \mathcal{T}_h} \in P_1(T)\}.$$

We also need the following finite element space:

$$V_h^0 = \{v \in V_h : v|_{\partial\Omega} = 0\}.$$

We derive *a priori* error analysis in the following norms:

$$\|v\|_h^2 = \int_\Omega |\nabla v|^2\, dx + \int_{\partial\Omega} \frac{1}{h^2} v^2\, ds, \tag{7}$$

and

$$\|v\|_1^2 = \int_\Omega |\nabla v|^2\, dx + \int_{\partial\Omega} \frac{1}{h} v^2\, ds. \tag{8}$$

It is obvious that the method (4) has a unique solution.

Let $\Pi_h u \in V_h$ be the standard Lagrange interpolation of u, i.e., $\Pi_h u(p) = u(p)$ for all nodes in $\mathcal{T}_h$. The following approximation properties are well-known [10, 16]:

$$\|u - \Pi_h u\| + h\|u - \Pi_h u\|_{H^1(\Omega)} \le Ch^2 \|u\|_{H^2(\Omega)}, \tag{9}$$

$$\int_{\partial\Omega} \frac{1}{h}(u - \Pi_h u)^2\, ds \le Ch^2 \|u\|_{H^2(\Omega)}^2, \tag{10}$$

hereafter C denotes a generic positive constant which is independent of h and takes different values at different appearances. The estimate in (9) follows from Bramble-Hilbert lemma and the estimate in (10) is a consequence of the trace inequality [10]:

$$\|v\|_{L_2(\partial\Omega)} \le C\|v\|_{L_2(\Omega)}^{1/2} \|v\|_{H^1(\Omega)}^{1/2}. \tag{11}$$

Define a trace space W_h by $W_h = V_h|_{\partial\Omega}$, i.e.,

$$W_h = \{w : w = v|_{\partial\Omega} \text{ for some } v \in V_h\}.$$

Let $g_h \in W_h$ be the L_2 projection of g, i.e.,

$$\int_{\partial\Omega} (g - g_h) w_h\, ds = 0 \quad \forall w_h \in W_h. \tag{12}$$

For our error analysis, we construct an interpolation $\mathcal{P}_h u \in V_h$ of u in the following. Define $\mathcal{P}_h u$ by the following conditions: $\mathcal{P}_h u(p) = g_h(p)$ for all nodes $p \in \partial\Omega$ and $\mathcal{P}_h u(p) = u(p)$ for all nodes $p \in \Omega$. It is clear that $\mathcal{P}_h u = g_h$ on $\partial\Omega$. Note that on any triangle T such that any vertex of T do not fall on $\partial\Omega$, $\mathcal{P}_h u = \Pi_h u$. Then using an inverse inequality [10, 16] and (10), we obtain for $l = 0, 1$

$$\begin{aligned}\|\mathcal{P}_h u - \Pi_h u\|_{H^l(\Omega)}^2 &\le Ch^{2(1-l)} \sum_{p \in \partial\Omega} (\Pi_h u - g_h)^2(p) \le Ch^{1-2l} \int_{\partial\Omega} (\Pi_h u - g_h)^2\, ds \\ &\le Ch^{1-2l} \int_{\partial\Omega} (\Pi_h u - u)^2\, ds \le Ch^{4-2l} \|u\|_{H^2(\Omega)}^2.\end{aligned}$$

Therefore we have the following estimates:

$$\|u - \mathcal{P}_h u\| + h\|u - \mathcal{P}_h u\|_{H^1(\Omega)} \le Ch^2 \|u\|_{H^2(\Omega)}, \tag{13}$$

and

$$\int_{\partial\Omega} \frac{1}{h}(u - \mathcal{P}_h u)^2\, ds \le Ch^2 \|u\|_{H^2(\Omega)}^2. \tag{14}$$

We use the following trace inequality in our analysis. The proof is well known [18].

Lemma 2.1. *Let $v \in H^2(\Omega)$. Then there exists a constant C which depends on Ω such that*

$$\|\partial v / \partial n\|_{L_2(\partial\Omega)} \le C\|v\|_{H^2(\Omega)}.$$

Error estimates in [1] were almost optimal order up to an arbitrary $\epsilon > 0$. Below we present improved error estimates.

Theorem 2.2. *There exists some positive C such that*

$$\|\mathcal{P}_h u - u_h\|_h \le Ch\|u\|_{H^2(\Omega)}, \tag{15}$$

$$\|u - u_h\|_1 \le Ch\|u\|_{H^2(\Omega)}, \tag{16}$$

and

$$\|u - u_h\| \le Ch^2\|u\|_{H^2(\Omega)}. \tag{17}$$

Proof. Note that the solution u of (1) satisfies

$$\int_\Omega \nabla u \cdot \nabla v\, dx - \int_{\partial\Omega} \frac{\partial u}{\partial n} v\, ds = \int_\Omega fv\, dx \quad \forall v \in H^1(\Omega).$$

Let $w = \mathcal{P}_h u - u_h$. Then using (5), (4) and (12), we find

$$\begin{aligned}
\|w\|_h^2 = a_h(w,w) &= a_h(\mathcal{P}_h u, w) - (f,w) - \int_{\partial\Omega} \frac{1}{h^2} gw\, ds \\
&= \int_\Omega \nabla(\mathcal{P}_h u - u)\cdot\nabla w\, dx + \int_{\partial\Omega} \frac{\partial u}{\partial n} w\, ds + \int_{\partial\Omega} \frac{1}{h^2}(\mathcal{P}_h u - g) w\, ds \\
&= \int_\Omega \nabla(\mathcal{P}_h u - u)\cdot\nabla w\, dx + \int_{\partial\Omega} \frac{\partial u}{\partial n} w\, ds. \qquad (18)
\end{aligned}$$

Using the trace inequality in Lemma 2.1, we obtain

$$\begin{aligned}
\left|\int_{\partial\Omega} \frac{\partial u}{\partial n} w\, ds\right| &\le Ch\left(\int_{\partial\Omega}\left(\frac{\partial u}{\partial n}\right)^2 ds\right)^{1/2}\left(\int_{\partial\Omega}\frac{1}{h^2}w^2\, ds\right)^{1/2} \\
&\le Ch\|u\|_{H^2(\Omega)}\left(\int_{\partial\Omega}\frac{1}{h^2}w^2\, ds\right)^{1/2}. \qquad (19)
\end{aligned}$$

Now the estimate in (15) follows from (18), (19) and (13). The estimate (16) follows from (15), (13) and (14). In order to prove (17), we apply the duality argument.

Let $\phi \in H^2(\Omega)\cap H_0^1(\Omega)$ be the unique solution of

$$\begin{aligned}
-\Delta\phi &= u - u_h \quad \text{in} \quad \Omega, \\
\phi &= 0 \quad \text{on} \quad \partial\Omega.
\end{aligned}$$

From the elliptic regularity theory [18] on convex polygonal domains, there exists a positive C such that

$$\|\phi\|_{H^2(\Omega)} \le C\|u - u_h\|_{L_2(\Omega)}.$$

Let $e = u - u_h$. Then

$$\|u - u_h\|^2 = \int_\Omega \nabla\phi\cdot\nabla e\, dx - \int_{\partial\Omega}\frac{\partial\phi}{\partial n} e\, ds. \tag{20}$$

Let $\phi_h \in V_h^0$ be the Lagrange interpolation of ϕ. Then,

$$\begin{aligned}
\|u - u_h\|^2 &= \int_\Omega \nabla(\phi - \phi_h)\cdot\nabla e\, dx + \int_\Omega \nabla\phi_h\cdot\nabla e\, dx - \int_{\partial\Omega}\frac{\partial\phi}{\partial n} e\, ds \\
&= \int_\Omega \nabla(\phi - \phi_h)\cdot\nabla e\, dx - \int_{\partial\Omega}\frac{\partial\phi}{\partial n}(u - \mathcal{P}_h u)\, ds - \int_{\partial\Omega}\frac{\partial\phi}{\partial n}(\mathcal{P}_h u - u_h)\, ds \\
&= \int_\Omega \nabla(\phi - \phi_h)\cdot\nabla e\, dx - \int_{\partial\Omega}\left(\frac{\partial\phi}{\partial n} - w_h\right)(u - \mathcal{P}_h u)\, ds \\
&\quad - \int_{\partial\Omega}\frac{\partial\phi}{\partial n}(\mathcal{P}_h u - u_h)\, ds, \qquad (21)
\end{aligned}$$

for any $w_h \in W_h$. The first term on the right-hand side of (21) is estimated as

$$\left|\int_{\Omega} \nabla(\phi - \phi_h) \cdot \nabla e \, dx\right| \le Ch\|\phi\|_{H^2(\Omega)}\|\nabla e\|_{L_2(\Omega)}.$$

For the third term, we use Cauchy-Schwarz and trace inequalities,

$$\left|\int_{\partial\Omega} \frac{\partial\phi}{\partial n}(\mathcal{P}_h u - u_h) \, ds\right| \le Ch\|\phi\|_{H^2(\Omega)} \left(\int_{\partial\Omega} \frac{1}{h^2}(\mathcal{P}_h u - u_h)^2 \, ds\right)^{1/2}.$$

To estimate the second term on the right-hand side of (21), let $\mathbf{p} = \nabla\phi$. Then from [4, Theorem 5.1], there exists a $\mathbf{p}_h \in V_h \times V_h$ such that $\mathbf{p}_h$ is zero at all the corners of Ω and

$$\|h^{-1}(\mathbf{p} - \mathbf{p}_h)\|_{L_2(\Omega)} + \|\nabla(\mathbf{p} - \mathbf{p}_h)\|_{L_2(\Omega)} \le C\|\nabla\mathbf{p}\|_{L_2(\Omega)}. \tag{22}$$

Let $w_h = \mathbf{p}_h \cdot n$. Then it is clear that $w_h|_{\partial\Omega} \in W_h$. Using Green's theorem, we obtain

$$\begin{aligned}
&\int_{\partial\Omega} \left(\frac{\partial\phi}{\partial n} - w_h\right)(\mathcal{P}_h u - u) \, ds \\
&= \int_{\partial\Omega} (\mathbf{p} - \mathbf{p}_h) \cdot n(\mathcal{P}_h u - u) \, ds \\
&= \int_{\Omega} (\mathbf{p} - \mathbf{p}_h) \cdot \nabla(\mathcal{P}_h u - u) \, dx + \int_{\Omega} (\mathcal{P}_h u - u)\nabla \cdot (\mathbf{p} - \mathbf{p}_h) \, dx.
\end{aligned}$$

Cauchy-Schwarz inequality and (22) complete the proof. □

Remark 2.3. For a general Dirichlet data g, it is difficult to obtain optimal error estimate for $\|u - u_h\|_h$. However it is possible to derive optimal error estimate for $\|u - u_h\|_h$ when g is trace of a finite element function. In particular, when g is zero. This can be seen from the estimate (15) and the fact that $\mathcal{P}_h u = u$ on $\partial\Omega$ when g is trace of a finite element function. For this reason, we have obtained optimal order L_2 norm error estimate since the adjoint problem has homogeneous boundary condition. This fact is used as a key in the convergence analysis of multigrid algorithms below.

3. Multigrid Algorithm

In this section, we present W-cycle and V-cycle multigrid algorithms for the penalty method (4). Let $\mathcal{T}_0$ be an initial triangulation of Ω. The triangulation $\mathcal{T}_k (k \ge 1)$ is then created recursively by subdividing each $T \in \mathcal{T}_{k-1}$ into four triangles (red refinement). Denote the set of all boundary edges of $\mathcal{T}_k$ by $\mathcal{E}_k^b$. We define $h_k = \max_{T \in \mathcal{T}_k} h_T$, where $h_T = \sqrt{|T|}$ and $|T|$ is the area of T. The mesh parameters satisfy $h_k = h_{k-1}/2$.

Let $V_k \subset H^1(\Omega)$ and $V_k^0 \subset H_0^1(\Omega)$ be the P_1 finite element spaces associated with $\mathcal{T}_k$. The k-th level penalty method for (1) is: Find $u_k \in V_k$ such that

$$a_k(u_k, v) = L_k(v) \quad \forall v \in V_k, \tag{23}$$

where

$$a_k(w, v) = \int_{\Omega} \nabla w \cdot \nabla v \, dx + \int_{\partial\Omega} \frac{1}{h_k^2} wv \, ds, \tag{24}$$

and

$$L_k(v) = \int_{\Omega} fv \, dx + \int_{\partial\Omega} \frac{1}{h_k^2} gv \, ds. \tag{25}$$

We can write (23) as

$$A_k u_k = f_k, \tag{26}$$

where $A_k : V_k \to V_k'$ and $f_k \in V_k'$ are defined by

$$\langle A_k w, v\rangle = a_k(w,v) \qquad \forall w, v \in V_k, \tag{27}$$

and

$$\langle f_k, v\rangle = L_k(v) \qquad \forall v \in V_k. \tag{28}$$

Here $\langle\cdot,\cdot\rangle$ is the canonical bilinear form on $V_k' \times V_k$.

Let $\mathcal{N}_k^i$(resp. $\mathcal{N}_k^b$) be the set of all vertices that are in Ω (resp. $\partial\Omega$). Denote $\mathcal{N}_k = \mathcal{N}_k^i \cup \mathcal{N}_k^b$. Let

$$b_k(w,v) = h_k^2 \sum_{\nu\in\mathcal{N}_k^i} w(\nu)v(\nu) + \sum_{\nu\in\mathcal{N}_k^b} h_k w(\nu)v(\nu) \qquad \forall w, v \in V_k, \tag{29}$$

and define $B_k : V_k \to V_k'$ by

$$\langle B_k w, v\rangle = b_k(w,v) \qquad \forall w, v \in V_k. \tag{30}$$

Note that the matrix B_k is diagonal with respect to the canonical nodal basis of V_k.

We prove the following estimate on condition number.

Theorem 3.1. *It holds for the condition number $\kappa(B_k^{-1}A_k)$ of $B_k^{-1}A_k$ that*

$$\kappa(B_k^{-1}A_k) \le Ch_k^{-2}.$$

Proof. By Rayleigh quotient formula, we note that

$$\lambda_{max}(B_k^{-1}A_k) = \max_{v\in V_k\setminus\{0\}} \frac{\langle A_k v, v\rangle}{\langle B_k v, v\rangle}$$

and

$$\lambda_{min}(B_k^{-1}A_k) = \min_{v\in V_k\setminus\{0\}} \frac{\langle A_k v, v\rangle}{\langle B_k v, v\rangle}$$

Let $v \in V_k$. Then using a standard inverse inequality,

$$\begin{aligned}\langle A_k v, v\rangle = a_k(v,v) &= \int_\Omega |\nabla v|^2\,dx + \int_{\partial\Omega} \frac{1}{h_k^2} v^2\,ds \le C \sum_{\nu\in\mathcal{N}_k} v(\nu)^2 + \sum_{e\in\mathcal{E}_k^b} \frac{1}{h_k^2}\int_e v^2\,ds \\ &\le Ch_k^{-2}\left(h_k^2 \sum_{\nu\in\mathcal{N}_k} v(\nu)^2 + \sum_{\nu\in N_k^b} h_k v(\nu)^2\right) \le Ch_k^{-2}\langle B_k v, v\rangle,\end{aligned}$$

which proves

$$\lambda_{max}(B_k^{-1}A_k) \le Ch_k^{-2}. \tag{31}$$

Using Poincaré-Friedrichs inequality, we obtain

$$\begin{aligned}\langle B_k v, v\rangle = b_k(v,v) = h_k^2 \sum_{\nu\in\mathcal{N}_k^i} v(\nu)^2 + \sum_{\nu\in N_k^b} h_k v(\nu)^2 &\le C\left(\|v\|_{L_2(\Omega)}^2 + \int_{\partial\Omega} v^2\,ds\right) \\ &\le C\langle A_k v, v\rangle,\end{aligned}$$

and

$$\lambda_{min}(B_k^{-1}A_k) \ge 1/C. \tag{32}$$

Hence the proof is completed by using $\kappa(B_k^{-1}A_k) = \lambda_{max}(B_k^{-1}A_k)/\lambda_{min}(B_k^{-1}A_k)$. □

In order to define multigrid algorithms, we need inter-grid transfer operators to move functions between grids. The coarse-to-fine operator $I_{k-1}^k : V_{k-1} \to V_k$ is the natural injection from $V_{k-1} \to V_k$, i.e., $I_{k-1}^k v = v,\ \forall v \in V_{k-1}$. The fine-to-coarse operator I_k^{k-1} is defined to be the transpose of I_{k-1}^k with respect to the canonical bilinear forms, i.e.,

$$\langle I_k^{k-1} w, v\rangle = \langle w, I_{k-1}^k v\rangle \quad \forall w \in V_k',\ v \in V_{k-1}.$$

We are now ready to define W-cycle and V-cycle algorithms for the equation

$$A_k z = b, \quad \text{where} \quad b \in V_k'.$$

Multigrid Algorithms:
Symmetric V-cycle Algorithm. Let $b \in V_k'$ and $z_0 \in V_k$ be an initial guess. The output of the V-cycle algorithm will be denoted by $MG_V(k, b, z_0, m)$, where m is the number of pre-smoothing and post-smoothing steps.
For $k = 1$, the output is taken to be $A_1^{-1} b$.
For $k > 1$, we proceed in three steps as follows:
Pre-smoothing. Compute $z_l \in V_k$ for $1 \le l \le m$, recursively by

$$z_l = z_{l-1} + \omega_k B_k^{-1}(b - A_k z_{l-1}), \tag{33}$$

where ω_k is a damping factor to be chosen later in Remark 3.2.
Coarse-grid correction. Compute $q \in V_{k-1}$ by

$$q = MG_V(k-1, r, 0, m), \quad \text{where} \quad r = I_k^{k-1}(b - A_k z_m), \tag{34}$$

and take $z_{m+1} = z_m + I_{k-1}^k q$.
Post-smoothing. Compute $z_l \in V_k$ for $m + 2 \le l \le 2m + 1$ recursively by

$$z_l = z_{l-1} + \omega_k B_k^{-1}(b - A_k z_{l-1}). \tag{35}$$

The final output is

$$MG_V(k, b, z_0, m) = z_{2m+1}.$$

Symmetric W-cycle Algorithm. The output $MG_W(k, b, z_0, m)$ of the W-cycle algorithm is obtained by replacing the coarse grid correction in (34) with
Coarse-grid correction. Compute $q \in V_{k-1}$ by

$$\begin{aligned} r &= I_k^{k-1}(b - A_k z_m), \\ q_* &= MG_W(k-1, r, 0, m), \\ q &= MG_W(k-1, r, q_*, m), \end{aligned} \tag{36}$$

and take $z_{m+1} = z_m + I_{k-1}^k q$.

Remark 3.2. In view of (31), we will take

$$\omega_k = c_* h_k^2 \tag{37}$$

in the smoothing steps (33) and (35), where the constant c_* is chosen according to $\omega_k \lambda_{max}(B_k^{-1} A_k) < 1$. For simplicity, we have used Richardson relaxation as our smoother. However the analysis developed here can be extended to other smoothers.

4. Convergence of W-Cycle Algorithm

Over the past few decades, there has been a huge attention on the convergence analysis of multigrid methods. The analysis can systematically be derived from the error propagation operator $E_k : V_k \to V_k$ for the k-th level W-cycle algorithm, i.e.,

$$E_k(z - z_0) = z - MG_W(k, b, z_0, m).$$

The following recursive relation is well known [10, 20, 21]:

$$E_k = R_k^m (Id_k - I_k^{k-1} P_k^{k-1} + I_{k-1}^k E_{k-1}^2 P_k^{k-1}) R_k^m, \tag{38}$$

where Id_k is the identity operator on V_k, the operator $R_k : V_k \to V_k$ which measures the effect of one smoothing step is defined by

$$R_k = (Id_k - \omega_k B_k^{-1} A_k),$$

and the operator $P_k^{k-1} : V_k \to V_{k-1}$ is the transpose of I_{k-1}^k with respect to the variational forms, i.e.,

$$a_{k-1}(P_k^{k-1} v, w) = a_k(v, I_{k-1}^k w) \qquad \forall w \in V_{k-1},\ v \in V_k. \tag{39}$$

For the convergence analysis, we define the mesh dependent norms $\|v\|_{j,k}$ for $j = 0, 1, 2$ by

$$\|v\|_{j,k} = \sqrt{\langle B_k (B_k^{-1} A_k)^j v, v \rangle} \qquad \forall v \in V_k,\ k \geq 1. \tag{40}$$

When $j = 0$, we obtain

$$\|v\|_{0,k} = \sqrt{\langle B_k v, v \rangle} \qquad \forall v \in V_k,\ k \geq 1, \tag{41}$$

and when $j = 1$, we get

$$\|v\|_{1,k} = \sqrt{\langle A_k v, v \rangle} \qquad \forall v \in V_k,\ k \geq 1. \tag{42}$$

From the Cauchy-Schwarz inequality and theory in [2], there hold

$$|a_k(v, w)| \leq \|v\|_{1,k} \|w\|_{1,k}, \tag{43}$$

and

$$|a_k(v, w)| \leq \|v\|_{2,k} \|w\|_{0,k}. \tag{44}$$

For completeness, we sketch the proof of the above two estimates in the appendix at the end of the article. In the rest of the article, C with or without subscript denotes a generic positive constant that is independent of h_k, k and m.

To prove the convergence of W-cycle algorithm, we only need to establish smoothing property and approximation property [6, 7, 10, 20, 21].

Following estimates on smoothing effect follow easily from the calculus techniques as in [10, Theorem 6.5.7]:

Lemma 4.1. *Let $v \in V_k$ and $k \geq 1$. Then*

$$\|R_k^m v\|_{1,k} \leq \|v\|_{1,k}, \tag{45}$$

$$\|R_k^m v\|_{1,k} \leq C h_k^{-1} (1 + m)^{-1/2} \|v\|_{0,k}, \tag{46}$$

and

$$\|R_k^m v\|_{2,k} \leq C h_k^{-1} (1 + m)^{-1/2} \|v\|_{1,k}. \tag{47}$$

Below, we establish an approximation property:

Lemma 4.2. *Let $v \in V_k$ and $k \geq 1$. It holds that*

$$\|(Id_k - I_{k-1}^k P_k^{k-1}) v\|_{0,k} \leq C h_k \|(Id_k - I_{k-1}^k P_k^{k-1}) v\|_{1,k}.$$

Proof. We prove this by duality argument. Let $v \in V_k$ and $\phi = (Id_k - I_{k-1}^k P_k^{k-1})v$. Let $\psi \in H_0^1(\Omega) \cap H^2(\Omega)$ be the solution of

$$\begin{aligned} -\Delta\psi &= \phi \quad \text{in} \quad \Omega, \\ \psi &= 0 \quad \text{on} \quad \partial\Omega. \end{aligned} \tag{48}$$

From the elliptic regularity theory, it holds that

$$\|\psi\|_{H^2(\Omega)} \le C\|\phi\|_{L_2(\Omega)}.$$

Let $\Pi_{k-1}\psi \in V_{k-1}^0$ be the Lagrange interpolation of ψ. Then from (39), for $v \in V_k$

$$a_k(v, I_{k-1}^k \Pi_{k-1}\psi) = a_{k-1}(P_k^{k-1}v, \Pi_{k-1}\psi) = a_k(I_{k-1}^k P_k^{k-1}v, I_{k-1}^k \Pi_{k-1}\psi).$$

From (48) and (5),

$$\begin{aligned} \|\phi\|_{L_2(\Omega)}^2 &= \int_\Omega \nabla\psi\cdot\nabla\phi\,dx - \int_{\partial\Omega}\frac{\partial\psi}{\partial n}\phi\,ds = a_k(\phi,\psi) - \int_{\partial\Omega}\frac{\partial\psi}{\partial n}\phi\,ds \\ &= a_k((Id_k - I_{k-1}^k P_k^{k-1})v, \psi) - \int_{\partial\Omega}\frac{\partial\psi}{\partial n}\phi\,ds \\ &= a_k((Id_k - I_{k-1}^k P_k^{k-1})v, \psi - I_{k-1}^k\Pi_{k-1}\psi) - \int_{\partial\Omega}\frac{\partial\psi}{\partial n}\phi\,ds \\ &\le Ch_k\|\nabla\phi\|_{L_2(\Omega)}\|\psi\|_{H^2(\Omega)} + Ch_k\|\psi\|_{H^2(\Omega)}\left(\int_{\partial\Omega}\frac{1}{h_k^2}\phi^2\,ds\right)^{1/2}, \end{aligned}$$

which implies

$$\|\phi\|_{L_2(\Omega)} \lesssim h_k\left(\|\nabla\phi\|_{L_2(\Omega)}^2 + \int_{\partial\Omega}\frac{1}{h_k^2}\phi^2\,ds\right)^{1/2}.$$

Since $\|\phi\|_{0,k}^2 = \langle B_k\phi,\phi\rangle \le C\left(\|\phi\|_{L_2(\Omega)}^2 + \int_{\partial\Omega}\phi^2\,ds\right)$, we find

$$\begin{aligned} \|\phi\|_{0,k}^2 &\le C\left(h_k^2\left(\|\nabla\phi\|_{L_2(\Omega)}^2 + \int_{\partial\Omega}\frac{1}{h_k^2}\phi^2\,ds\right) + \int_{\partial\Omega}\phi^2\,ds\right) \\ &= Ch_k^2\left(\|\nabla\phi\|_{L_2(\Omega)}^2 + \int_{\partial\Omega}\frac{1}{h_k^2}\phi^2\,ds\right) = Ch_k^2 a_k(\phi,\phi) = Ch_k^2\|\phi\|_{1,k}^2. \end{aligned}$$

This completes the proof. □

Below, we prove stability estimates for the inter-grid operators.

Lemma 4.3. *Let $k \ge 1$. Then, it holds that*

$$\|I_{k-1}^k v\|_{1,k} \le C\|v\|_{1,k-1} \quad \forall v \in V_{k-1}, \tag{49}$$

$$\|P_k^{k-1}v\|_{1,k-1} \le C\|v\|_{1,k} \quad \forall v \in V_k. \tag{50}$$

Proof. First we prove (49). Let $v \in V_{k-1}$ and $w = I_{k-1}^k v$. Then

$$\begin{aligned} \|w\|_{1,k}^2 = a_k(w,w) &= \int_\Omega |\nabla w|^2\,dx + \int_{\partial\Omega}\frac{1}{h_k^2}w^2\,ds \\ &= \int_\Omega |\nabla v|^2\,dx + 4\int_{\partial\Omega}\frac{1}{h_{k-1}^2}v^2\,ds \le 4\|v\|_{1,k-1}^2. \end{aligned}$$

Next to prove (50), let $v \in V_k$. Then using (39), (43) and (49),

$$\begin{aligned} \|P_k^{k-1}v\|_{1,k-1}^2 &= a_{k-1}(P_k^{k-1}v, P_k^{k-1}v) = a_k(v, I_{k-1}^k P_k^{k-1}v) \\ &\le C\|v\|_{1,k}\|I_{k-1}^k P_k^{k-1}v\|_{1,k} \le C\|v\|_{1,k}\|P_k^{k-1}v\|_{1,k-1}, \end{aligned}$$

which completes the proof. □

In the next lemma, we derive another approximation property.

Lemma 4.4. *Let $v \in V_k$ and $k \geq 1$. Then it holds that*

$$\|(Id_k - I_{k-1}^k P_k^{k-1})v\|_{0,k} \leq C h_k^2 \|v\|_{2,k}.$$

Proof. First of all, we note that for $v, w \in V_k$,

$$\begin{aligned}
&a_k((Id_k - I_{k-1}^k P_k^{k-1})v, w) \\
=&a_k(v,w) - a_k(I_{k-1}^k P_k^{k-1} v, w) = a_k(v,w) - a_k(w, I_{k-1}^k P_k^{k-1} v) \\
=&a_k(v,w) - a_{k-1}(P_k^{k-1} w, P_k^{k-1} v) = a_k(v,w) - a_{k-1}(P_k^{k-1} v, P_k^{k-1} w) \\
=&a_k(v,w) - a_k(v, I_{k-1}^k P_k^{k-1} w) = a_k(v, (Id_k - I_{k-1}^k P_k^{k-1})w).
\end{aligned}$$

For $w \in V_k$, we note using (49) and (50) that

$$\|(Id_k - I_{k-1}^k P_k^{k-1})w\|_{1,k} \leq \|w\|_{1,k} + \|I_{k-1}^k P_k^{k-1} w\|_{1,k} \leq C\|w\|_{1,k}. \tag{51}$$

Let $w = (Id_k - I_{k-1}^k P_k^{k-1})v$. Using (44), Lemma 4.2 and (51),

$$\begin{aligned}
\|(Id_k - I_{k-1}^k P_k^{k-1})v\|_{1,k}^2 &= a_k((Id_k - I_{k-1}^k P_k^{k-1})v, w) = a_k(v, (Id_k - I_{k-1}^k P_k^{k-1})w) \\
&\leq C\|v\|_{2,k} \|(Id_k - I_{k-1}^k P_k^{k-1})w\|_{0,k} \\
&\leq C h_k \|v\|_{2,k} \|(Id_k - I_{k-1}^k P_k^{k-1})w\|_{1,k} \leq C h_k \|v\|_{2,k} \|w\|_{1,k}.
\end{aligned}$$

Therefore

$$\|(Id_k - I_{k-1}^k P_k^{k-1})v\|_{1,k} \leq C h_k \|v\|_{2,k},$$

which together with Lemma 4.2 completes the proof. □

The following result on two-grid method will be useful in the convergence of W-cycle algorithm.

Theorem 4.5. *There exists a positive constant C_{TG} independent of h and m such that for $v \in V_k$ and $k \geq 1$,*

$$\|R_k^m (Id_k - I_{k-1}^k P_k^{k-1}) R_k^m v\|_{1,k} \leq C_{TG}(1+m)^{-1} \|v\|_{1,k}. \tag{52}$$

Proof. Using (46), Lemma 4.4 and (47), we find

$$\begin{aligned}
\|R_k^m (Id_k - I_{k-1}^k P_k^{k-1}) R_k^m v\|_{1,k} &\leq C h_k^{-1/2} (1+m)^{-1/2} \|(Id_k - I_{k-1}^k P_k^{k-1}) R_k^m v\|_{0,k} \\
&\leq C h_k^{1/2} (1+m)^{-1/2} \|R_k^m v\|_{2,k} \\
&\leq C(1+m)^{-1} \|v\|_{1,k}.
\end{aligned}$$

This completes the proof. □

We are now ready to prove the convergence of W-cycle algorithm. The prove will be similar to the ones in [2, 6, 7, 10].

Theorem 4.6. *There exists a positive integer m_* and $C_* > 0$ independent of k such that*

$$\|E_k v\|_{1,k} \leq C_*(1+m)^{-1} \|v\|_{1,k}, \tag{53}$$

provided $m \geq m_*$.

Proof. We prove this by mathematical induction. The case $k = 0$ holds for any m_* since $A_0 z = g$ is solved exactly. Assume $k \geq 1$ and (53) holds for $k-1$. Recall that

$$E_k v = R_k^m (Id_k - I_k^{k-1} P_k^{k-1}) R_k^m v + R_k^m (I_{k-1}^k E_{k-1}^2 P_k^{k-1}) R_k^m v.$$

Using (45), (49), induction hypothesis and (50), we find

$$
\begin{aligned}
\|R_k^m(I_{k-1}^k E_{k-1}^2 P_k^{k-1}) R_k^m v\|_{1,k} &\leq \|(I_{k-1}^k E_{k-1}^2 P_k^{k-1}) R_k^m v\|_{1,k} \\
&\leq C\|E_{k-1}^2 P_k^{k-1} R_k^m v\|_{1,k-1} \\
&\leq CC_*^2(1+m)^{-2}\|P_k^{k-1} R_k^m v\|_{1,k-1} \\
&\leq CC_*^2(1+m)^{-2}\|R_k^m v\|_{1,k} \\
&\leq CC_*^2(1+m)^{-2}\|v\|_{1,k}. \qquad (54)
\end{aligned}
$$

Theorem 4.5 and (54) imply

$$\|E_k v\|_{1,k} \leq \left(C_{TG} + CC_*^2(1+m)^{-1}\right)(1+m)^{-1}\|v\|_{1,k}.$$

Therefore, choose C_* large and then m_* large enough so that

$$C_{TG} + CC_*^2(1+m)^{-1} \leq C_* \quad \text{for} \quad m \geq m_*.$$

This completes the proof of the theorem. □

5. Convergence of V-Cycle Algorithm

Since the variational form in the finite element penalty method (4) is mesh dependent, the convergence of V-cycle algorithm needs to be treated as in the case of nonconforming finite element methods. We use the additive multigrid theory developed in [8]. In the recent years, additive multigrid theory is successfully used in analyzing the nonconforming [8] and the discontinuous Galerkin V-cycle algorithms [9, 11, 13].

In order to use additive multigrid theory, we need to establish the following bounds besides the ones in Section 4:

$$\|I_{k-1}^k v\|_{1,k}^2 \leq (1+\theta^2)\|v\|_{1,k-1}^2 + C\theta^{-2}h_k^2\|v\|_{2,k-1}^2 \qquad \forall v \in V_{k-1},\ \ \theta \in (0,1), \qquad (55)$$

$$\|I_{k-1}^k v\|_{0,k}^2 \leq (1+\theta^2)\|v\|_{0,k-1}^2 + C\theta^{-2}h_k^2\|v\|_{1,k-1}^2 \qquad \forall v \in V_{k-1},\ \ \theta \in (0,1), \qquad (56)$$

(57)

$$\|P_k^{k-1} v\|_{0,k-1}^2 \leq (1+\theta^2)\|v\|_{0,k}^2 + C\theta^{-2}h_k^2\|v\|_{1,k}^2 \qquad \forall v \in V_k,\ \ \theta \in (0,1),$$

and

$$\|(Id_{k-1} - P_k^{k-1} I_{k-1}^k)v\|_{0,k-1} \leq Ch_k\|v\|_{1,k-1} \qquad \forall v \in V_{k-1}. \qquad (58)$$

Comparing with the abstract assumptions in [8, Section 3], the estimates (55)-(57) are related to the assumptions on I_{k-1}^k and P_k^{k-1}, the estimate in Lemma 4.2 and the estimate (58) are related to the assumptions on $I_{k-1}^k P_k^{k-1}$ and $P_k^{k-1} I_{k-1}^k$. The assumptions on V_k in [8, Section 3] are related to the condition number estimate, see (3.13) of [8]. In our case, the condition number estimate of the preconditioned system is already established in Theorem 3.1. Therefore, we can apply the additive multigrid theory once we establish (55)-(58).

Lemma 5.1. *Estimate* (55) *is valid.*

Proof. Since $I_{k-1}^k : V_{k-1} \to V_k$ is a natural injection, it is enough to prove

$$\|w\|_{1,k}^2 \leq (1+\theta^2)\|w\|_{1,k-1}^2 + C\theta^{-2}h_k^2\|w\|_{2,k-1}^2 \qquad \forall w \in V_{k-1},\ \ \theta \in (0,1).$$

Let $w \in V_{k-1}$ be arbitrary. Then, it is trivial that there is a unique $\psi \in V_{k-1}$ such that

$$\int_\Omega \psi v\,dx + \int_{\partial\Omega} \psi v\,ds = a_{k-1}(w,v) \quad \forall v \in V_{k-1}.$$

Using (44), (29) and (41), we find that

$$\|\psi\|_{0,k-1} \le \|w\|_{2,k-1}. \tag{59}$$

By the definition of the norms and since $h_k = h_{k-1}/2$, we note that

$$\begin{aligned}\|w\|_{1,k}^2 = a_k(w,w) = \|\nabla w\|^2 + \int_{\partial\Omega} \frac{1}{h_k^2} w^2\, ds = \|\nabla w\|^2 + \int_{\partial\Omega} \frac{4}{h_{k-1}^2} w^2\, ds \\ = \|w\|_{1,k-1}^2 + \int_{\partial\Omega} \frac{3}{h_{k-1}^2} w^2\, ds.\end{aligned}$$

Therefore it is enough to establish

$$\int_{\partial\Omega} \frac{1}{h_{k-1}^2} w^2\, ds \le C h_k^2 \|w\|_{2,k-1}^2.$$

Let $\phi \in H^2(\Omega) \cap H_0^1(\Omega)$ be the solution of

$$\begin{aligned} -\Delta\phi &= \psi \quad \text{in} \quad \Omega, \\ \phi &= 0 \quad \text{on} \quad \partial\Omega. \end{aligned}$$

From the elliptic regularity theory [18], it holds that

$$\|\phi\|_{H^2(\Omega)} \le C\|\psi\|.$$

Let $\phi_{k-1} = \Pi_{k-1}\phi \in V_{k-1}^0$ be the Lagrange interpolation of ϕ and let $\eta = \phi_{k-1} - w$. Then, using Lemma 2.1,

$$\begin{aligned} a_{k-1}(\eta, v) &= a_{k-1}(\phi_{k-1}, v) - a_{k-1}(w, v) = a_{k-1}(\phi_{k-1}, v) - \left(\int_\Omega \psi v\, dx + \int_{\partial\Omega} \psi v\, ds \right) \\ &= \int_\Omega \nabla(\phi_{k-1} - \phi)\cdot\nabla v\, dx + \int_{\partial\Omega} \frac{\partial\phi}{\partial n} v\, ds - \int_{\partial\Omega} \psi v\, ds \\ &\le C h_{k-1} \left(\|\phi\|_{H^2(\Omega)}^2 + \|\psi\|_{L_2(\partial\Omega)}^2 \right)^{1/2} \left(\|\nabla v\|^2 + \int_{\partial\Omega} \frac{1}{h_{k-1}^2} v^2\, ds \right)^{1/2} \\ &\le C h_{k-1} \left(\|\psi\|_{L_2(\Omega)}^2 + \|\psi\|_{L_2(\partial\Omega)}^2 \right)^{1/2} \left(\|\nabla v\|^2 + \int_{\partial\Omega} \frac{1}{h_{k-1}^2} v^2\, ds \right)^{1/2} \\ &\le C h_{k-1} \|\psi\|_{0,k-1} \left(\|\nabla v\|^2 + \int_{\partial\Omega} \frac{1}{h_{k-1}^2} v^2\, ds \right)^{1/2}. \end{aligned}$$

Choosing $v = \eta$ and using (59), we get

$$\|\eta\|_{1,k-1} = \left(\|\nabla\eta\|^2 + \int_{\partial\Omega} \frac{1}{h_{k-1}^2} \eta^2\, ds \right)^{1/2} \le C h_{k-1} \|w\|_{2,k-1}.$$

Now using the fact that $\phi_{k-1} = 0$ on $\partial\Omega$, we complete the proof. □

Lemma 5.2. *Assume that each interior node in $\mathcal{T}_0$ is connected by at most six edges. Then the estimate* (56) *is valid.*

Proof. Let $v \in V_{k-1}$. Let $\mathcal{M}_k^i$(resp. $\mathcal{M}_k^b$) be the set of all midpoints on the interior (resp. boundary) edges of $\mathcal{T}_k$. Then by the definition of the norm

$$\begin{aligned} \|v\|_{0,k}^2 &= h_k^2 \sum_{p\in\mathcal{N}_k^i} v^2(p) + h_k \sum_{p\in\mathcal{N}_k^b} v^2(p) \\ &= h_k^2 \sum_{p\in\mathcal{N}_{k-1}^i} v^2(p) + h_k \sum_{p\in\mathcal{N}_{k-1}^b} v^2(p) + h_k^2 \sum_{p\in\mathcal{M}_{k-1}^i} v^2(p) + h_k \sum_{p\in\mathcal{M}_{k-1}^b} v^2(p) \end{aligned}$$

$$= \frac{h_{k-1}^2}{4} \sum_{p\in\mathcal{N}_{k-1}^i} v^2(p) + \frac{h_{k-1}}{2} \sum_{p\in\mathcal{N}_{k-1}^b} v^2(p)$$
$$+ h_k^2 \sum_{p\in\mathcal{M}_{k-1}^i} v^2(p) + h_k \sum_{p\in\mathcal{M}_{k-1}^b} v^2(p)$$
$$= \|v\|_{0,k-1}^2 + \frac{h_{k-1}^2}{4} \left(\sum_{p\in\mathcal{M}_{k-1}^i} v^2(p) - 3 \sum_{p\in\mathcal{N}_{k-1}^i} v^2(p) \right)$$
$$+ \frac{h_{k-1}}{2} \left(\sum_{p\in\mathcal{M}_{k-1}^b} v^2(p) - \sum_{p\in\mathcal{N}_{k-1}^b} v^2(p) \right).$$

Let p be the midpoint of an edge e in $\mathcal{E}_{k-1}$ with end points p_1 and p_2. Then

$$v^2(p) \le \frac{1}{2}\left(v^2(p_1) + v^2(p_2)\right),$$

and using the assumption in the hypothesis, we obtain

$$\sum_{p\in\mathcal{M}_{k-1}^i} v^2(p) \le 3 \sum_{p\in\mathcal{N}_{k-1}^i} v^2(p) + 3 \sum_{p\in\mathcal{N}_{k-1}^b} v^2(p).$$

Therefore

$$\left| \|v\|_{0,k}^2 - \|v\|_{0,k-1}^2 \right| \le \frac{3h_{k-1}^2}{4} \sum_{p\in\mathcal{N}_{k-1}^b} v^2(p) + \frac{h_{k-1}}{2} \left(\sum_{p\in\mathcal{M}_{k-1}^b} v^2(p) - \sum_{p\in\mathcal{N}_{k-1}^b} v^2(p) \right).$$

A use of an inverse inequality implies

$$h_{k-1} \sum_{p\in\mathcal{N}_{k-1}^b} v^2(p) \le C h_k^2 \left(\int_{\partial\Omega} \frac{1}{h_{k-1}^2} v^2\, ds \right)$$

and

$$\frac{h_{k-1}}{2} \left(\sum_{p\in\mathcal{M}_{k-1}^b} v^2(p) + \sum_{p\in\mathcal{N}_{k-1}^b} v^2(p) \right) \le C h_k^2 \left(\int_{\partial\Omega} \frac{1}{h_{k-1}^2} v^2\, ds \right).$$

Thus

$$\left| \|v\|_{0,k}^2 - \|v\|_{0,k-1}^2 \right| \le C h_k^2 \left(\int_{\partial\Omega} \frac{1}{h_{k-1}^2} v^2\, ds \right) \le C h_k^2 \|v\|_{1,k-1}^2. \tag{60}$$

This completes the proof. □

Lemma 5.3. *Assume that each interior node in $\mathcal{T}_0$ is connected by at most six edges. Then the estimate* (57) *is valid.*

Proof. Let $w \in V_k$ be arbitrary and let $w_{k-1} = P_k^{k-1} w$. Then using (60) and (50), we find

$$\|w_{k-1}\|_{0,k-1}^2 \le \|w_{k-1}\|_{0,k}^2 + C h_k^2 \|w_{k-1}\|_{1,k-1}^2$$
$$\le \|w_{k-1}\|_{0,k}^2 + C h_k^2 \|w\|_{1,k}^2.$$

Using the triangle inequality, we find

$$\|w_{k-1}\|_{0,k}^2 \le \|w_{k-1} - w\|_{0,k}^2 + \|w\|_{0,k}^2.$$

Lemma 4.2 implies that

$$\|w_{k-1} - w\|_{0,k}^2 \leq Ch_k^2 \|w\|_{1,k}^2.$$

This completes the proof. □

The proof of the following lemma is similar to that of Lemma 4.2 and uses the estimate (50). Therefore we skip the proof.

Lemma 5.4. *Estimate* (58) *is valid.*

We have validated all the assumptions for the additive multigrid theory in [8]. Therefore, we can conclude the convergence of V-cycle algorithm from the additive multigrid theory.

Theorem 5.5. *Assume that each interior node in the initial mesh $\mathcal{T}_0$ is connected by at most six edges and the number of smoothing steps m is sufficiently large. Then the output of V-cycle algorithm $MG_V(k, b, z_0, m)$ satisfies the estimate*

$$\|z - MG_V(k, b, z_0, m)\|_{1,k} \leq \frac{C}{m} \|z - z_0\|_{1,k},$$

where C is a positive constant independent of k and m.

Remark 5.6. The assumption in Theorem 5.5 on the initial mesh $\mathcal{T}_0$ is not too restrictive since in practice such meshes can be generated for polygonal domains.

Remark 5.7. The results in Section 5 also imply the convergence of F-cycle algorithm.

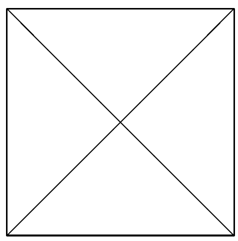

FIGURE 1. Initial mesh satisfying the assumption in Theorem 5.5.

TABLE 1. Order of Convergence.

h	$\|u - u_h\|_1$	order	$\|u - u_h\|$	order
$1/2^0$	3.8703e+000	–	7.7738e-001	–
$1/2^1$	1.7592e+000	1.1375	2.2581e-001	1.7835
$1/2^2$	8.4926e-001	1.0506	5.4616e-002	2.0477
$1/2^3$	4.1977e-001	1.0166	1.3222e-002	2.0464
$1/2^4$	2.1047e-001	0.9960	3.2658e-003	2.0174
$1/2^5$	1.0575e-001	0.9929	8.1583e-004	2.0011
$1/2^6$	5.3062e-002	0.9949	2.0450e-004	1.9961
$1/2^7$	2.6586e-002	0.9970	5.1273e-005	1.9959

6. Numerical Experiments

In this section, we report some numerical results demonstrating the performance of the penalty method (4) and multigrid algorithms. We consider the computational domain to be the unit square $(0,1)\times(0,1)$ and take the data f and g be such that the exact solution is $u(x,y)=e^{(x+y)}$. We take the initial mesh $\mathcal{T}_0$ as in the Figure 1 satisfying the assumption in Theorem 5.5.

TABLE 2. Contraction numbers of the W-cycle algorithm.

	$k=1$	$k=2$	$k=3$	$k=4$	$k=5$	$k=6$	$k=7$
$m=1$	1.1681	1.1083	1.0138	0.8960	1.0702	0.6723	0.4326
$m=2$	0.5413	0.3877	0.3930	0.3258	0.3439	0.3322	0.3340
$m=3$	0.2706	0.2550	0.2805	0.1619	0.1496	0.1432	0.1399
$m=4$	0.1393	0.1586	0.2580	0.2833	0.0648	0.0574	0.0538
$m=5$	0.0732	0.1003	0.0860	0.1842	0.0321	0.0247	0.0212

TABLE 3. Contraction numbers of the V-cycle algorithm.

	$k=1$	$k=2$	$k=3$	$k=4$	$k=5$	$k=6$	$k=7$
$m=1$	1.1681	2.9577	5.5919	9.3608	14.6869	22.2081	32.8730
$m=2$	0.5413	0.9145	1.1373	1.2238	1.2130	1.1507	1.0720
$m=3$	0.2706	0.3732	0.4055	0.3841	0.3425	0.3046	0.2772
$m=4$	0.1393	0.1863	0.1989	0.1841	0.1665	0.1560	0.1515
$m=5$	0.0732	0.1091	0.1176	0.1115	0.1052	0.1022	0.1012

We generate the meshes $\mathcal{T}_k$ (for $k\geq 1$) recursively by dividing each triangle in $\mathcal{T}_{k-1}$ into four sub-triangles. On these nested meshes, we have computed the order of convergence in both the energy and the L_2 norms and depicted them in the Table 1. The numerical convergence rates illustrate the theoretical convergence rates in Theorem 2.2. We next verify the convergence of multigrid algorithms. The damping parameter c_* in (37) is taken to be 1/5 that is required in the multigrid smoothing steps (33) and (35). The contraction numbers for various smoothing steps at various mesh levels are listed in Table 2 and Table 3 for W-cycle and V-cycle algorithms, respectively. These computations demonstrate that the W-cycle algorithm is contraction for smoothing steps two or more and the V-cycle algorithm is contraction for smoothing steps three or more.

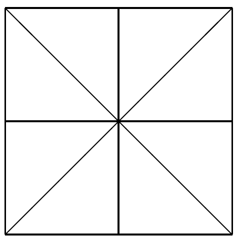

FIGURE 2. Initial mesh not satisfying the condition in Theorem 5.5.

In the second experiment, we consider the same model problem but consider the initial mesh that violates the condition in Theorem 5.5. In this case also we have observed the V-cycle algorithm is contraction for smoothing steps five or more. Perhaps, a proof of Lemma 5.2 without the assumption on the initial mesh may be derived. We leave this subject to our future research.

TABLE 4. Contraction numbers of the V-cycle algorithm.

	$k=1$	$k=2$	$k=3$	$k=4$	$k=5$	$k=6$
$m=3$	0.8936	1.9576	3.1365	4.3559	5.5692	6.7638
$m=4$	0.6268	1.1659	1.5706	1.8150	1.9210	1.9330
$m=5$	0.4460	0.7402	0.8928	0.9254	0.8853	0.8160
$m=6$	0.3206	0.4962	0.5604	0.5462	0.4970	0.4425
$m=7$	0.2322	0.3481	0.3794	0.3583	0.3207	0.2878

7. Conclusions

In this article, the penalty method by Babuška is considered for inhomogeneous Dirichlet problem and optimal order *a priori* error estimates are recovered by the help of an interpolation. A simple pre-conditioner is introduced to remedy the ill-condition of the stiffness matrix due to over-penalty. Convergence analysis of W-cycle and V-cycle algorithms is derived where the analysis of V-cycle algorithm is established under a mild assumption on the initial mesh, see Lemma 5.2. Numerical experiments in the article illustrate the performance of multigrid algorithms.

Acknowledgements

The author would like to acknowledge the fruitful discussions with Prof. J. P. Raymond on numerical approximation of boundary control problems during the seminars and discussions on the IFCAM Project “PDE Control”.

References

[1] I. Babuska, The finite element method with penalty, Math. Comp., 27 (1973) 221–228.
[2] R. E. Bank and T. Dupont, An optimal order process for solving finite element equations, Math. Comp., 36 (1981) 35–51.
[3] R.E. Bank and C.C. Douglas, Sharp estimates for multigrid rates of convergence with general smoothing and acceleration, SIAM J. Numer. Anal., 22 (1985) 617–633.
[4] S. Bartels, C. Carstensen and G. Dolzmann, Inhomogeneous Dirichlet conditions in a priori and a posteriori finite element error analysis, Numer. Math., 99 (2004) 1–24.
[5] J.H. Bramble and J.E. Pasciak, New convergence estimates for multigrid algorithms, Math. Comp., 49 (1987) 311–329.
[6] J. H. Bramble, J. E. Pasciak and J. Xu, The analysis of multigrd algorithms with nonnested spaces or noninherited quadratic forms, Math. Comp., 56 (1991) 1–34.
[7] J. H. Bramble, Multigrid Methods, Pitman Research Notes in Mathematics, v. 294, John Wiley and Sons, 1993.
[8] S.C. Brenner, Convergence of nonconforming V-cycle and F-cycle multigrid algorithms for second order elliptic boundary value problem, Math. Comp., 73 (2004) 1041–1066.
[9] S.C. Brenner and L.-Y. Sung, Multigrid algorithms for C^0 interior penalty methods, SIAM J. Numer. Anal., 44 (2006) 199–223.
[10] S.C. Brenner and L.R. Scott, The Mathematical Theory of Finite Element Methods (Third Edition), Springer-Verlag, New York, 2008.
[11] S.C. Brenner and J. Zhao, Convergence of multigrid algortihms for interior penalty methods, Appl. Num. Anal. Comp. Math., 2 (2005) 3–18.
[12] S. C. Brenner, J. Cui and L. Y. Sung, Multigrid methods for the symmetric interior penalty method on graded meshes, Numer. Linear Algebra Appl., 16 (2009) 481–501.
[13] S.C. Brenner, J. Cui, T. Gudi and L. Y. Sung, Multigrid algortihms for symmetric discontinuous Galerkin methods on graded meshes, Numer. Math., 119 (2011) 21–47.

[14] D. Braess and W. Hackbusch, A new convergence proof for the multigrid method including the V-cycle, SIAM J. Numer. Anal., 20 (1983) 967–975.
[15] D. Braess and R. Verfürth, Multigrid methods for nonconforming finite element methods, SIAM J. Numer. Anal., 27 (1990) 979–986.
[16] P.G. Ciarlet. The Finite Element Method for Elliptic Problems, North-Holland, Amsterdam, 1978.
[17] C. Eduardo, M. Mariano and J. P. Raymond, Penalization of Dirichlet optimal control problems, ESAIM Control Optim. Calc. Var., 15 (2009) 782–809.
[18] P. Grisvard, Elliptic Problems in Nonsmooth Domains, Pitman, Boston, 1985.
[19] J. Gopalakrishnan and G. Kanschat, A multilevel discontinuous Galerkin method, Numer. Math., 95 (2003) 527–550.
[20] W. Hackbusch, Multi-gird Methods and Applications, Springer, Berlin, 1985.
[21] J. Mandel, S. McCormick and R. Bank, *Variational multigrid theory*, in:S. McCormik (Ed.), Multigrid Methods, Frontiers in Applied Mathematics, vol. 3. SIAM Philadelphia, PA, 1987.
[22] J. A. Nitsche, Ü ber ein Variationsprinzip zur Lösung Dirichlet-Problemen bei Verwendung von Teilrä umen, die keinen Randbedingungen unteworfen sind, Abh. Math. Sem. Univ. Hamburg, 36 (1971) 9–15.
[23] U. Trottenberg, C. Oosterlee and A. Schüller, Multigrid, Academic Press, San Diego, 2001
[24] P. Wesseling, An Introduction to Multigrid Methods, John Wiley & Sons, Chichester, 1992.
[25] H. Yserentant, Old and new convergence proofs for multigrid methods, Acta Numerica, pp: 285–326, 1993.
[26] X. Zhang, Multilevel Schwarz methods, Numer. Math., 63 (1992) 521–539.

Appendix A. Proof of (43)-(44)

We prove the estimates by similar arguments in [2]. Let N_k be the dimension of V_k. Consider the eigenvalue problem:

$$A_k\psi = \lambda B_k\psi.$$

Since A_k is a positive definite matrix, there exist eigenvalues $\lambda_i > 0$ for all $1 \leq i \leq N_k$. Let $\{\psi_i\}_{i=1}^{N_k}$ be the corresponding eigenvectors such that $\langle B_k\psi_i, \psi_j\rangle = \delta_{ij}$ and $\langle A_k\psi_i, \psi_j\rangle = \lambda_i\delta_{ij}$ for $1 \leq i, j \leq N_k$, where δ_{ij} is the Kronecker delta function. Take any $v \in V_k$ and write it as

$$v = \sum_{i=1}^{N_k} \alpha_i\psi_i. \tag{A.1}$$

Then, we have

$$\|v\|_{0,k}^2 = \langle B_k v, v\rangle = \sum_{i=1}^{N_k} \alpha_i^2,$$

and

$$\|v\|_{1,k}^2 = \langle A_k v, v\rangle = \sum_{i=1}^{N_k} \alpha_i^2\lambda_i.$$

Consider any $w \in V_k$ and write it as

$$w = \sum_{i=1}^{N_k} \beta_i\psi_i. \tag{A.2}$$

Then, we find

$$\langle A_k v, w\rangle = \sum_{i=1}^{N_k} \alpha_i\beta_i\lambda_i \leq \|v\|_{1,k}\|w\|_{1,k}.$$

This proves (43). For given v as in (A.1), let $w = B_k^{-1} A_k v$ and write w in the form (A.2). Then $\beta_i = \alpha_i \lambda_i$ and

$$\|v\|_{2,k}^2 = \langle B_k (B_k^{-1} A_k)^2 v, v\rangle = \langle A_k B_k^{-1} A_k v, v\rangle = \langle A_k w, v\rangle = \sum_{i=1}^{N_k} \alpha_i \beta_i \lambda_i = \sum_{i=1}^{N_k} \alpha_i^2 \lambda_i^2.$$

Now take any $v \in V_k$ and $w \in V_k$ of the form (A.1) and (A.2), respectively. Then using the Cauchy-Schwarz inequality, we find

$$\langle A_k v, w\rangle = \sum_{i=1}^{N_k} \alpha_i \beta_i \lambda_i \leq \left(\sum_{i=1}^{N_k} \alpha_i^2 \lambda_i^2 \right)^{1/2} \left(\sum_{i=1}^{N_k} \beta^2 \right)^{1/2} = \|v\|_{2,k} \|w\|_{0,k}.$$

This proves (44).

Department of Mathematics, Indian Institute of Science, Bangalore - 560012
E-mail: gudi@math.iisc.ernet.in