A GENERALIZED BPX MULTIGRID FRAMEWORK
COVERING NONNESTED V-CYCLE METHODS

HUO-YUAN DUAN, SHAO-QIN GAO, ROGER C. E. TAN, AND SHANGYOU ZHANG

Abstract. More than a decade ago, Bramble, Pasciak and Xu developed a
framework in analyzing the multigrid methods with nonnested spaces or non-
inherited quadratic forms. It was subsequently known as the BPX multigrid
framework, which was widely used in the analysis of multigrid and domain
decomposition methods. However, the framework has an apparent limit in the
analysis of nonnested V-cycle methods, and it produces a variable V-cycle, or
nonuniform convergence rate V-cycle methods, or other nonoptimal results in
analysis thus far.

This paper completes a long-time effort in extending the BPX multigrid
framework so that it truly covers the nonnested V-cycle. We will apply the
extended BPX framework to the analysis of many V-cycle nonnested multigrid
methods. Some of them were proven previously only for two-level and W-cycle
iterations. Some numerical results are presented to support the theoretical
analysis of this paper.

1. Introduction

The multigrid method, consisting of the fine-level smoothing and the coarse-level
correction, is an effective iterative method for solving the linear system arising from,
e.g., the finite element discretization of boundary-value problems. The multigrid
method provides the optimal-order computation in such a case, in the sense that
the number of arithmetic operations is proportional to the number of unknowns
in the system of linear equations; cf. [1], [5], [28], [31], [33]. The constant rate of
W-cycle multigrid iterations was proved in several early papers, one of them is [1],
which is generalized to many nonnested cases, for example, [13], [15], [42], [43].

The multigrid method is often nonnested because the multilevel discrete spaces
may not be nested, or discrete bilinear forms may be different on different levels.
For example, the nonnesteness may be caused by bubble elements [18], composite
elements [18], nonconforming elements [1], [13], nonnested meshes [42], the mortar
method [2], [25], numerical integrations [24], or other situations (cf. [11]), such
as finite difference equations. The multigrid methods with noninherited forms but
nested spaces, other than the cases in [11], are studied in [26], [27], [30] for the
discontinuous Galerkin method and the edge element. Many earlier two-level and
W-cycle nonnested multigrid iterations were analyzed by extending the method of
[1]. However, a generalized framework [11], referred as the BPX multigrid frame-
work, is widely used in the analysis of multigrid iterations; e.g., [5], [6], [8], [9], [10],

Received by the editor July 21, 2001 and, in revised form, November 27, 2005.
2000 Mathematics Subject Classification. Primary 65N55, 65N30, 65F10.
Key words and phrases. V-cycle nonnested multigrid method.

©2006 American Mathematical Society
The framework is rooted in [7] and [38]. Although the convergence theory for the W-cycle was established [1], [11], [15], [4], the problem of how to establish the convergence rate for the V-cycle nonnested multigrid method is subtle, and is still an active research subject; see [3], [11], [10], [9], [7], [16], [28], [20], [21], [29], [6], [12], etc. The BPX framework [11] was generalized to allow nonsymmetric smoothings and can be applied to some nonnested multigrid methods. In particular, it provides a constant convergence rate for the nonnested V-cycle under the assumption that

$$A_k(I_k u, I_k u) \leq A_{k-1}(u, u) \quad \forall u \in U_{k-1}, \forall k,$$

where $I_k : U_{k-1} \to U_k$ is the coarse-to-fine intergrid transfer operator and $A_k$ is the bilinear form on $U_k$. However, (1) does not hold for most nonnested multigrid methods. Thus the BPX framework produces some nonoptimal mathematics results, such as the variable V-cycle, nonuniform convergence rate, and multigrid preconditioners, (cf. [11], [37]) though most of these methods provide the optimal order of computation. The question has remained open for a long time whether one can lift this obvious limit, the inequality (1), from the BPX framework.

This question will be answered in this paper. We will extend the BPX framework so that the number of smoothings can play its important role in the V-cycle analysis so that the BPX framework can provide a uniform convergence rate without the nearly nested bound (1). We will then apply the extended BPX framework to show the uniform convergence rates of several common nonnested multigrid methods. Some of them were proven previously for two-level and W-cycle iterations only.

So far, we still require the full elliptic regularity assumption in our applications of the BPX framework. Brenner recently gave a proof in [17] for the nonconforming V-cycle multigrid method applied to the second-order elliptic problem, under a lower regularity requirement. It is a better result. In addition, the analysis [17] can be extended to some other nonnested multigrid methods; cf. [44]. However, it is not straightforward to apply Brenner’s analysis to different cases in general, due to its lengthy analysis and its long list of approximation properties and inverse estimates. For example, the standard inverse estimate fails to hold on the combined space of finite element functions on two nonnested grids such as the ones in Figure 1. In contrast to [17], our extended BPX framework is simple in analysis and can be applied to all common nonnested cases.

The outline of this paper is as follows. In Section 2 we recall the V-cycle multigrid method. The convergence analysis is given in Section 3. In Section 4 we provide a proof for the regularity-approximation assumption for various nonnested methods. Some numerical results will be provided in Section 5 to support the theoretical analysis in this paper.

2. THE V-CYCLE MULTIGRID METHOD

For $k \geq 0$, let $U_k$ be a sequence of finite-dimensional vector spaces, along with coarse-to-fine intergrid transfer operators $I_k : U_{k-1} \to U_k$. Let $A_k(\cdot, \cdot)$ and $(\cdot, \cdot)_k$ be symmetric positive definite discrete bilinear forms on $U_k \times U_k$. We solve the following linear system of equations. Given $f \in U_k$, find $v \in U_k$ satisfying

$$A_k(v, \phi) = (f, \phi)_k \quad \forall \phi \in U_k.$$
To define a V-cycle multigrid method for (1), following the notations in [11], we introduce operators $A_k : U_k \to U_k$, $P_{k-1} : U_k \to U_{k-1}$ and $P_{k-1}^0 : U_k \to U_{k-1}$ as:

$$(A_k w, \phi)_k = A_k(w, \phi) \quad \forall \phi \in U_k,$$

$$(A_{k-1}(P_{k-1} w, \phi) = A_k(w, I_k \phi) \quad \forall \phi \in U_{k-1},$$

$$(P_{k-1}^0 w, \phi)_{k-1} = (w, I_k \phi)_{k-1} \quad \forall \phi \in U_{k-1}.$$  

We also introduce linear smoothing operators $R_k : U_k \to U_k$, along with the adjoint operators $R_k^*$ with respect to the inner product $(\cdot, \cdot)_k$. We define

$$R_k^{(l)} = \begin{cases} R_k & \text{if } l \text{ is odd}, \\ R_k^* & \text{if } l \text{ is even}. \end{cases}$$

Now we define the standard (symmetric) V-cycle multigrid method [11]. Let $m$ be a positive integer, the number of fine-level smoothings. The multigrid operator $B_k : U_k \to U_k$ is defined by induction as follows. Set $B_0 = A_0^{-1}$. Assume that $B_{k-1}$ has been defined, and define $B_k g \in U_k$ for $g \in U_k$ as follows.

(i) Set $x^0 = 0$.

(ii) Define $x^l$ for $l = 1, 2, \ldots, m$ by

$$x^l = x^{l-1} + R_k^{(l+m)}(g - A_k x^{l-1}).$$

(iii) Define $y^m = x^m + I_k q^1$, where $q^1$ is defined by

$$q^1 = B_{k-1} P_{k-1}^0 (g - A_k x^m).$$

(iv) Define $y^l$ for $l = m + 1, m + 2, \ldots, 2m$ by

$$y^l = y^{l-1} + R_k^{(l+m)}(g - A_k y^{l-1}).$$

(v) Set $B_k g = y^{2m}$.

3. THE CONVERGENCE ANALYSIS

To analyze the convergence, we set $J_k = I - R_k A_k$ and $J_k^* = I - R_k^* A_k$, where $J_k^*$ denotes the adjoint of $J_k$ with respect to $A_k(\cdot, \cdot)$ and $I$ is the identity operator. Set

$$\tilde{J}_k^{(m)} = \begin{cases} (J_k^* J_k)^{m/2} & \text{if } m \text{ is even,} \\ (J_k^* J_k)^{(m-1)/2} J_k^* & \text{if } m \text{ is odd.} \end{cases}$$

We then have the following recursive relation among the multigrid operators (cf. [11])

$$I - B_k A_k = (\tilde{J}_k^{(m)})^* [(I - I_k P_{k-1}) + I_k (I - B_{k-1} A_{k-1}) P_{k-1}] \tilde{J}_k^{(m)}.$$  

We make two standard hypotheses (cf. [11]) as follows:

(C1) Regularity-approximation assumption

$$|A_k((I - I_k P_{k-1})u, u)| \leq C_1 \frac{|A_k u|^2}{\lambda_k} \quad \forall u \in U_k,$$

where $\lambda_k$ is the largest eigenvalue of $A_k$, $C_1$ is independent of $k$, and $\| \cdot \|_k$ is the norm corresponding to $(\cdot, \cdot)_k$. In addition, we require that (see remarks below)

$$A_k((I - I_k P_{k-1})u, (I - I_k P_{k-1})u))^{1/2} \leq C_Q (A_k(u, u))^{1/2} \quad \forall u \in U_k,$$

where $C_Q$ is independent of $k$.  

(C2) \[ \frac{||u||^2}{\lambda_k} \leq C_R(\bar{R}_k u, u) \quad \forall u \in U_k, \]

where \( \bar{R}_k = (I - J_k^* J_k) A_k^{-1} \) and \( C_R \) is independent of \( k \).

Remark 3.1. The smoothing hypothesis (C2) can be easily verified for point, line, and block versions of the Jacobi and Gauss-Seidel iterations (cf. [8], for example). The verification of the regularity-approximation hypothesis (C1) will be carried out in the next section for many examples. The requirement (C3) can be verified easily for all practical cases. Inequality (3) is a simple corollary (cf. [43] for example) of the stability estimate (see (11)).

\[ A_k(I_k u, I_k u) \leq C A_{k-1}(u, u) \quad \forall u \in U_{k-1}. \]

Theorem 3.1. Assume that (C1) and (C2) hold. Then, for all \( k \geq 0 \),

\[ |A_k((I - B_k A_k)u, u)| \leq \delta A_k(u, u) \quad \forall u \in U_k, \]

where

\[ \delta = \frac{C_1 C_R}{m - C_1 C_R} \]

with \( m > 2 C_1 C_R \).

Proof. The method here is motivated by [11],[7], reasoning by mathematical induction. For \( k = 0 \), we have a zero on the left-hand side of (5), and (5) holds. It is assumed that (5) and (6) hold for \( k - 1 \). In what follows, we show that (5) and (6) hold for \( k \) too.

In view of (C1), we have

\[ |A_k((I - I_k P_{k-1}) J^{(m)}_k u, J^{(m)}_k u)| \leq C_1 \frac{|A_k J^{(m)}_k u|^2}{\lambda_k}. \]

Define

\[ \bar{J}_k = \begin{cases} J^*_k J_k & \text{if } m \text{ is even}, \\ J_k J^*_k & \text{if } m \text{ is odd}. \end{cases} \]

By (C2) we have

\[ \frac{|A_k J^{(m)}_k u|^2}{\lambda_k} \leq C_R A_k((I - \bar{J}_k) J^{m}_k u, u). \]

Since the spectrum of \( \bar{J}_k \) is in \([0, 1]\), as shown in [11],[7], we have

\[ A_k((I - \bar{J}_k) J^{m}_k u, u) \leq \frac{1}{m} \sum_{i=0}^{m-1} A_k((I - \bar{J}_k) J^{m}_k u, u) = \frac{1}{m} \{ A_k(u, u) - A_k(\bar{J}^{m}_k u, u) \}. \]

Note that \( A_k(\bar{J}^{m}_k u, u) = A_k(\bar{J}^{(m)}_k u, \bar{J}^{(m)}_k u) \). We then get, by (7)–(9), that

\[ |A_k((I - I_k P_{k-1}) J^{(m)}_k u, J^{(m)}_k u)| \leq \frac{C_1 C_R}{m} \{ A_k(u, u) - A_k(\bar{J}^{(m)}_k u, J^{(m)}_k u) \}. \]

Set

\[ t := \frac{A_k(\bar{J}^{(m)}_k u, J^{(m)}_k u)}{A_k(u, u)} \quad \forall u \neq 0, u \in U_k, \]
or \( t := 0 \) for \( u = 0 \). Clearly, \( t \in [0, 1] \). We now rewrite (10) as

\[
|A_k((I - I_k P_{k-1}) \tilde{J}_k^{(m)} u, \tilde{J}_k^{(m)} u)| \leq \frac{C_1 C_R (1 - t)}{m} A_k(u, u).
\]

On the other hand, from the Cauchy–Schwarz inequality and (3) we have

\[
|A_k((I - I_k P_{k-1}) \tilde{J}_k^{(m)} u, \tilde{J}_k^{(m)} u)| \leq \left\{ A_k((I - I_k P_{k-1}) \tilde{J}_k^{(m)} u, (I - I_k P_{k-1}) \tilde{J}_k^{(m)} u) \right\}^\frac{1}{2} \left\{ A_k(\tilde{J}_k^{(m)} u, \tilde{J}_k^{(m)} u) \right\}^\frac{1}{2}
\]

\[
\leq C_Q A_k(\tilde{J}_k^{(m)} u, \tilde{J}_k^{(m)} u) = C_Q t A_k(u, u).
\]

Combining (12) and (13), we get

\[
|A_k((I - I_k P_{k-1}) \tilde{J}_k^{(m)} u, \tilde{J}_k^{(m)} u)| \leq \min\{C_Q t, \frac{C_1 C_R}{m} (1 - t) \} A_k(u, u).
\]

By the relation

\[
A_k-1(P_{k-1}\tilde{J}_k^{(m)} u, P_{k-1}\tilde{J}_k^{(m)} u) = A_k(\tilde{J}_k^{(m)} u, \tilde{J}_k^{(m)} u) - A_k(\tilde{J}_k^{(m)} u, (I - I_k P_{k-1}) \tilde{J}_k^{(m)} u),
\]

the induction hypothesis and the symmetry of \( A_k \), we get

\[
|A_k((I - B_k A_k) u, u)|
\]

\[
\leq |A_k((I - I_k P_{k-1}) \tilde{J}_k^{(m)} u, \tilde{J}_k^{(m)} u)| + |A_k-1((I - B_k-1 A_k-1) P_{k-1} \tilde{J}_k^{(m)} u, P_{k-1} \tilde{J}_k^{(m)} u)|
\]

\[
\leq (1 + \delta) |A_k((I - I_k P_{k-1}) \tilde{J}_k^{(m)} u, \tilde{J}_k^{(m)} u)| + \delta A_k(\tilde{J}_k^{(m)} u, \tilde{J}_k^{(m)} u)
\]

\[
\leq (1 + \delta) \min\{C_Q t, \frac{C_1 C_R}{m} (1 - t) \} A_k(u, u) + \delta t A_k(u, u).
\]

Now, to show that (5) and (6) for \( k \), we only need to verify

\[
(1 + \delta) \min\{C_Q t, \frac{C_1 C_R}{m} (1 - t) \} + \delta t \leq \frac{C_1 C_R}{m - C_1 C_R} \quad \forall t \in [0, 1].
\]

When \( t = 0 \), the left-hand side of (15) is zero. When \( t = 1 \), (15) is the induction hypothesis. Next, we consider the case of \( t \in (0, 1) \). To show (15), by the hypothesis (6) on level \( k - 1 \), it suffices to show that

\[
(1 + \delta) C_Q \min\left( \frac{t}{1 - t}, \frac{C_1 C_R}{C_Q m} \right) \leq \delta.
\]

We consider two cases. First,

\[
\frac{C_1 C_R}{C_Q m} + \frac{C_1 C_R}{C_R} \leq t < 1,
\]

i.e.,

\[
\frac{t}{1 - t} \geq \frac{C_1 C_R}{C_Q m}.
\]

Thus

\[
\min\left( \frac{t}{1 - t}, \frac{C_1 C_R}{C_Q m} \right) = \frac{C_1 C_R}{C_Q m},
\]

and

\[
(1 + \delta) C_Q \min\left( \frac{t}{1 - t}, \frac{C_1 C_R}{C_Q m} \right) \leq \frac{C_1 C_R}{m - C_1 C_R}.
\]
For the second case,
\[ 0 < t \leq \frac{C_1 C_R}{C_Q \, m + C_1 C_R}, \]
i.e.,
\[ \frac{t}{1 - t} \leq \frac{C_1 C_R}{C_Q \, m}, \]
we have
\[ \min\left( \frac{t}{1 - t}, \frac{C_1 C_R}{C_Q \, m} \right) = \frac{t}{1 - t}, \]
and
\[ (1 + \delta) C_Q \, \min\left( \frac{t}{1 - t}, \frac{C_1 C_R}{C_Q \, m} \right) = \frac{m C_Q}{m - C_1 C_R} \, \frac{t}{1 - t} \leq \frac{C_1 C_R}{m - C_1 C_R}. \]
Thus, equation (10) holds for both cases (17) and (18).

**Remark 3.2.** Note that from (14) we can get
\[ |A_k((I - I_k P_{k-1}) (j_k^{(m)} u, j_k^{(m)} v))| \leq \frac{C_1 C_R}{m + \frac{C_1 C_R}{C_Q}} A_k(u, u), \]
which indicates that the number of smoothings, \( m \), has to be large enough for the convergence rate in the interval \((0, 1)\) in general, even for the two-level method.

On the other hand, if \( C_Q = 1 \), we get a convergence rate in \((0, 1)\) by (11), for the two-level method, for any \( m \geq 1 \). Note that if
\[ A_k(I_k P_{k-1} v, I_k P_{k-1} v) \leq 2 A_{k-1}(P_{k-1} v, P_{k-1} v) \quad \forall v \in U_k, \quad \forall k, \]
then \( C_Q = 1 \). In some cases, (20) holds (see [11] and [22]). Note that (20) is a generalization of (11).

**Remark 3.3.** The key step in our proof is the introduction of a variable \( t \) in (11). By it, we extend the BPX framework from the very limited case (11) to the general case (3), or just (1).

### 4. Verification of (C1)

In this section, we provide a proof for the regularity-approximation assumption (C1) in solving the symmetric and positive definite second-order elliptic problems by various nonnested methods.

Let \( \Omega \) be a bounded, connected domain in \( \mathbb{R}^n \), \( n = 2 \) or 3, with Lipschitz continuous boundary \( \partial \Omega \). We will use the Sobolev space \( H^l(\Omega) \), \( l \geq 0 \), with the norm and seminorm \( || \cdot ||_{H^l(\Omega)} \) and \( | \cdot |_{H^l(\Omega)} \). The \( L^2(\Omega) (= H^0(\Omega)) \) inner product is denoted by \( (\cdot, \cdot)_{L^2(\Omega)} \).

We set \( U = H^1_0(\Omega) \) and \( A(u, v) = \sum_{|\alpha|, |\beta| \leq 1} \int_\Omega a_{\alpha \beta}(x) \partial^{\alpha} u \partial^{\beta} v, \) where \( \alpha, \beta \) are \( n \)-indexes and \( a_{\alpha \beta}(x) \in L^\infty(\Omega) \). Let \( \mathcal{J}_k, k \geq 0 \) denote a sequence of shape-regular triangulations of \( \Omega \), with the mesh-size \( h_k \); cf. [19]. On \( \mathcal{J}_k \), let \( U_k \) be a finite-dimensional space and \( A_k(\cdot, \cdot) \) a discrete form on \( U_k \times U_k \).

We first list some general hypotheses.

**H1** We require that \( A(\cdot, \cdot) \) is symmetric and positive definite, and that for any given \( f \in L^2(\Omega) \) there is a unique solution \( u \in U \) such that
\[ A(u, v) = (f, v)_{L^2(\Omega)} \quad \forall v \in U, \]
and that \( u \in H^2(\Omega) \) satisfies

\[ \|u\|_{H^2(\Omega)} \leq C \|f\|_{L^2(\Omega)}. \]

**H2**) For all \( k \), we require that \( A_k(\cdot, \cdot) \) is a symmetric, positive definite and bounded bilinear form. We set

\[ \|v\|_{1,k} := \sqrt{A_k(v,v)} \quad \forall v \in U_k, \forall k. \]

**H3**) Let \( \Pi_j u \in U_j \) denote the standard interpolant to \( u \in H^2(\Omega) \). For all \( k \), we require that

\[ \|u - \Pi_j u\|_{L^2(\Omega)} + h_k \|u - \Pi_j u\|_{1,j} \leq C \, h_k^2 \|u\|_{H^2(\Omega)}, \quad j = k - 1, k. \]

**H4**) Let \( I_k : U_{k-1} \to U_k \) denote the coarse-to-fine intergrid transfer operator. For all \( k \), we require that

\[ \|I_k v\|_{L^2(\Omega)} \leq C \|v\|_{L^2(\Omega)} \quad \forall v \in U_{k-1}. \]

**H5**) For all \( k \), we require that

\[ C^{-1} \|v\|_{L^2(\Omega)} \leq \|v\|_k \leq C \|v\|_{L^2(\Omega)} \quad \forall v \in U_k. \]

**H6**) For all \( k \), we require that the following inverse inequality holds,

\[ \|v\|_{1,k} \leq C \, h_k^{-1} \|v\|_{L^2(\Omega)} \quad \forall v \in U_k. \]

**H7**) Let \( u_j \in U_j \) be a finite-element approximation to \( u \), the exact solution for a given \( f \in L^2(\Omega) \), i.e.

\[ A(u, v) = (f, v)_{L^2(\Omega)} \quad \forall v \in U, \quad A_j(u_j, v) = (f, v)_{L^2(\Omega)} \quad \forall v \in U_j. \]

For all \( k \), we require that

\[ \|u - u_j\|_{L^2(\Omega)} + h_k \|u - u_j\|_{1,j} \leq C \, h_k^2 \|f\|_{L^2(\Omega)}, \quad j = k - 1, k. \]

**Remark 4.1.** For \( C^0 \) conforming elements and nonconforming elements such as the Crouzeix–Raviart element, usually **H7** results from **H1**-**H3** (cf. [19], [35] and [13]).

**Theorem 4.1.** Assume **H2**, **H5** and **H6**. If the following assumption,

\[ (I - I_k P_{k-1}) v |v|_k \leq C \, h_k^2 \|A_k v\|_k \quad \forall v \in U_k \]

holds, then (C1) holds.

**Proof.** By **H2**), the eigenvalues \( \lambda_{k,i} \) and eigenvectors \( \psi_{k,i}, 1 \leq i \leq N_k \), satisfy

\[ A_k(\psi_{k,i}, v) = \lambda_{k,i}(\psi_{k,i}, v)_k \quad \forall v \in U_k, \]

\[ 0 < \lambda_{k,1} \leq \lambda_{k,2} \leq \cdots \leq \lambda_{k,N_k}, \]

\[ (\psi_{k,i}, \psi_{k,j})_{0,k} = \delta_{ij}, \quad A_k(\psi_{k,i}, \psi_{k,j}) = \lambda_{k,i} \delta_{ij}, \]

where \( \delta_{ij} \) is the Kronecker symbol. It can be easily seen that

\[ \|A_k^{1/2} w\|_k^2 = (A_k w, w)_k \leq \lambda_{k,N_k} (w, w)_k \quad \forall w \in U_k. \]

Set \( \lambda_k := \lambda_{k,N_k} \). From **H5** and **H6** we see that

\[ \lambda_k \leq C \, h_k^{-2}. \]

By (C1’) we can conclude that (C1) holds, since

\[ |A_k((I - I_k P_{k-1}) v, v)| \leq \|(I - I_k P_{k-1}) v\|_k \|A_k v\|_k \leq C \, h_k^2 \|A_k v\|_k^2 \leq C_1 \|A_k v\|_k^2 \lambda_k \]

License or copyright restrictions may apply to redistribution; see http://www.ams.org/journal-terms-of-use
and
\[
(A_k((I - I_k P_{k-1})v, (I - I_k P_{k-1})v))^{1/2} = \|(I - I_k P_{k-1})v\|_{1,k} \leq C h_k^{-1} \|(I - I_k P_{k-1})v\|_k \leq C h_k \|A_k v\| = C h_k \|A_k^{1/2} A_k^{1/2} v\|_k \leq C h_k \lambda_k^{1/2} (A_k(v, v))^{1/2} \leq C_M (A_k(v, v))^{1/2}.
\]

**Theorem 4.2.** Let hypotheses H1)–H5) and H7) hold. If
\[
||w_k - I_k w_{k-1}||_{L^2(\Omega)} \leq C h_k^2 ||g||_{L^2(\Omega)}
\]
holds, then (C1)' holds. Here \(w_j \in U_j\) denotes the finite-element solution on the \(j\)-th level for a given \(g \in L^2(\Omega)\); i.e.
\[
A_j(w_j, q) = (g, q)_{L^2(\Omega)} \quad \forall q \in U_j.
\]

**Proof.** The proof is divided into two steps. In the first step, we show that (21) implies
\[
||w_k - P_{k-1} w_k||_{L^2(\Omega)} \leq C h_k^2 ||g||_{L^2(\Omega)}.
\]
To do so, we consider a dual problem: Find \(z \in U\) such that
\[
A(z, q) = (w_k - P_{k-1} w_k, q)_{L^2(\Omega)} \quad \forall q \in U.
\]
Denote by \(z_j \in U_j\) the finite element solution to (23); i.e.
\[
A_j(z_j, q) = (w_k - P_{k-1} w_k, q)_{L^2(\Omega)} \quad \forall q \in U_j.
\]
Applying (21) with a right-hand side function \(w_k - P_{k-1} w_k \in L^2(\Omega)\), by H7) and the triangle inequality, we have
\[
||z_{k-1} - I_k z_{k-1}||_{L^2(\Omega)} \leq ||z_{k-1} - z_k||_{L^2(\Omega)} + ||z_k - I_k z_{k-1}||_{L^2(\Omega)} \leq ||z_{k-1} - z||_{L^2(\Omega)} + ||z - z_k||_{L^2(\Omega)} + ||z_k - I_k z_{k-1}||_{L^2(\Omega)} \leq C h_k^2 ||w_k - P_{k-1} w_k||_{L^2(\Omega)}.
\]
Therefore,
\[
||w_k - P_{k-1} w_k||^2_{L^2(\Omega)} = A_k(z_{k-1}, w_{k-1} - P_{k-1} w_k) = A_k(z_{k-1}, w_{k-1} - A_{k-1}(z_{k-1}, I_k z_{k-1}, w_k)) = A_k(z_{k-1}, w_{k-1}) - A_{k-1}(z_{k-1}, P_{k-1} w_k) = (g, z_{k-1} - I_k z_{k-1})_{L^2(\Omega)} \leq C h_k^2 ||g||_{L^2(\Omega)} ||w_k - P_{k-1} w_k||_{L^2(\Omega)}.
\]
It follows that (22) holds.

Now we take the second step, showing (C1)'. To do so, set \(E_k := I - I_k P_{k-1}\). Again, we consider a dual problem: Find \(z \in U\) such that
\[
A(z, q) = (E_k v, q)_{L^2(\Omega)} \quad \forall q \in U.
\]
Let \(z_j \in U_j\) be the finite-element solution, approximating \(z\); i.e.,
\[
A_j(z_j, q) = (E_k v, q)_{L^2(\Omega)} \quad \forall q \in U_j.
\]
From (21) and (22) we have
\[
||z_k - I_k z_{k-1}||_{L^2(\Omega)} + ||z_{k-1} - P_{k-1} z_k||_{L^2(\Omega)} \leq C h_k^2 ||E_k v||_{L^2(\Omega)}.
\]
Therefore, in view of H4) and H5),
\[ \|E_kv\|^2_{L^2(\Omega)} = A_k(E_kv, z_k) \]
\[ = A_k(v, z_k) - A_{k-1}(P_{k-1}v, P_{k-1}z_k) \]
\[ = A_k(v, z_k - I_kz_k) + A_k(v, I_k(z_k - P_{k-1}z_k)) \]
\[ \leq \|A_kv\|_k \|z_k - I_kz_k\|_k + \|A_kv\|_k \|I_k(z_k - P_{k-1}z_k]\|_k \]
\[ \leq C \|A_kv\|_k \{\|z_k - I_kz_k\|_{L^2(\Omega)} + \|I_k(z_k - P_{k-1}z_k)\|_{L^2(\Omega)}\} \]
\[ \leq C \|A_kv\|_k \{|\|z_k - I_kz_k\|_{L^2(\Omega)} + |\|z_k - P_{k-1}z_k\|_{L^2(\Omega)}\} \]
\[ \leq C h_k^2 \|A_kv\|_k \|E_kv\|_{L^2(\Omega)}, \]
The proof is completed. □

Proposition 4.1. Assume H1)–H4) and H7). If the following estimate holds,
\[ \|\Pi_kw - I_k \Pi_{k-1}w\|_{L^2(\Omega)} \leq C h_k^2 \|w\|_{H^2(\Omega)} \quad \forall w \in H^2(\Omega), \]
then (21) holds.

Proof. With a right-hand side function \( g \in L^2(\Omega) \), let \( w \in H^2(\Omega) \) be the solution to
\[ A(w, q) = (g, q)_{L^2(\Omega)} \quad \forall q \in U. \]
Let \( w_j \in U_j \) be the finite-element solution to \( w \). From H7) we know that
\[ \|w - w_j\|_{L^2(\Omega)} \leq C h_k^2 \|g\|_{L^2(\Omega)}, \quad j = k - 1, k. \]
Rewriting
\[ w_k - I_k w_{k-1} = w_k - w + w - \Pi_kw + \Pi_kw - I_k \Pi_{k-1}w + I_k(\Pi_{k-1}w - w + w - w_{k-1}), \]
by (20), (24), H3) and H4), we get (21). □

Proposition 4.2. Assume H1)–H4) and H7). If there exists a finite-dimensional space \( \Sigma_{k-1} \subseteq U_k \cap U_{k-1} \), which has the same order of approximation as that of \( U_k \) and \( U_{k-1} \), such that
\[ I_k w \equiv w \quad \forall w \in \Sigma_{k-1}, \]
or, if the following estimate holds,
\[ \|z - I_k q\|_{L^2(\Omega)} \leq C \|z - q\|_{L^2(\Omega)} \quad \forall z \in U_k, \forall q \in U_{k-1}, \]
then (21) holds.

Proof. Inequality (21) trivially results from (28), the triangle inequality and H7).
Let us assume (27). Let \( w_j \in U_j, \quad j = k, k - 1 \) and \( q_{k-1} \in \Sigma_{k-1} \) be the finite-element solutions to \( w \), for \( g \in L^2(\Omega) \); cf. (25) and (26).
\[ w_k - I_k w_{k-1} = w_k - q_{k-1} + q_{k-1} - I_k w_{k-1} \]
\[ = w_k - w + w - q_{k-1} + I_k(q_{k-1} - w + w - w_{k-1}). \]
Inequality (21) follows. □

Remark 4.2. For \( P_1 \) and Wilson’s nonconforming elements, (27) is obviously true, with \( \Sigma_{k-1} \) being the conforming \( P_1 \) and \( Q_1 \) elements, respectively; see [13] and [11].
For \( C^0 \) elements with nonnested triangulations, (28) was shown in [12]. For other nonnested \( C^0 \) elements such as a bubble-enriching element and a composed element, (28) was shown in [23].
For nonconforming elements such as $Q_{rot}^{rot}$ element and a discretely divergence-free $P_1$ element, (24) was shown in [36], [4] and [14].

For the Mortar element, (C1)’ was shown in [25].

Note that for all these nonnested cases there are different assumptions on the triangulations.

To avoid proliferation, here we leave out more detailed description and verification of assumptions (23), (26) and (27), associated with the assumption (20) for (C1)’, for various nonnested V-cycle methods. Readers can refer to the cited references for details.

Remark 4.3. To our best knowledge, all the existing intergrid transfer operators satisfy $H_4$; cf. [43], [11], [4], [2], [14], [13], [41] and [23]. Other hypotheses $H_1$–$H_3$ and $H_5$–$H_7$ are often trivial and standard. We therefore do not insist on details here.

5. Numerical results

In our numerical tests, we studied $P_1$ linear triangles and $P_1$ nonconforming linear elements, where the nodal values are defined at the vertices or the midedge points, respectively. We tested both nested and nonnested, but uniform grids, (see Figure 1), on the unit square domain $\Omega = (0, 1)^2$.

The bilinear form is the semi-$H^1$ product $A_k(u, v) = \int_\Omega u_x v_x + u_y v_y$. The discrete $L^2$ inner product is $(u, v)_k = h^2 \sum u_i v_i$ where the summation is over all nodal points and $h$ is the grid size.

In Table 1 below, we listed the constants computed numerically for the $P_1$ conforming elements on nonnested grids, where $C_1$ and $C_Q$ are used in the regularity-approximation assumption (C.1), the constant $C_H$ is from the smoothing hypothesis (C.2), $\delta_k$ (less than the theoretic constant $\delta$ in (5)) is the error reduction factor of the V-cycle nonnested multigrid method, and $\delta'_k$ is the two-level error reduction factor. Here we solve the coarse-level correction problem exactly in the two-level multigrid method and $\delta'_k$ is the the spectral radius of such a multigrid operator:

$$\delta'_k = \rho \left( (J_k^{(m)})^* (I - I_k P_{k-1}) J_k^{(m)} \right).$$

We note in particular that all constants are computed by Matlab as they are the maximum or the minimum of certain eigenvalues. Here the number of the smoothing parameter $m$ is set to 8 and the Richardson iteration is used for the presmoothing and the postsmoothing. The same constants for the $P_1$ nonconforming elements on nonnested grids are listed in Table 2. Because of the nonnested grids, the intergrid transfer operator $I_k$ is simply the nodal value interpolation operator $\Pi_k$ as all fine-level nodal points are in the interior of some coarse-level triangles. This avoids

Figure 1. Nonnested grids. ($h = 1/3$ and $h = 1/5$)
Table 1. Constants for $P_1$ conforming elements on nonnested grids.

<table>
<thead>
<tr>
<th>level $k$</th>
<th>grid</th>
<th>$C_1$</th>
<th>$C_Q$</th>
<th>$C_R$</th>
<th>$\delta_k$</th>
<th>$\delta_k'$</th>
</tr>
</thead>
<tbody>
<tr>
<td>2</td>
<td>3 x 3</td>
<td>1.7676</td>
<td>5.3645</td>
<td>1.0000</td>
<td>0.0008</td>
<td>0.0000</td>
</tr>
<tr>
<td>3</td>
<td>5 x 5</td>
<td>3.3794</td>
<td>7.0620</td>
<td>1.0000</td>
<td>0.0470</td>
<td>0.0468</td>
</tr>
<tr>
<td>4</td>
<td>9 x 9</td>
<td>4.8324</td>
<td>7.7069</td>
<td>1.0000</td>
<td>0.0923</td>
<td>0.0787</td>
</tr>
<tr>
<td>5</td>
<td>17 x 17</td>
<td>6.0642</td>
<td>7.9184</td>
<td>1.0000</td>
<td>0.1289</td>
<td>0.1132</td>
</tr>
<tr>
<td>6</td>
<td>33 x 33</td>
<td>6.9498</td>
<td>7.9785</td>
<td>1.0000</td>
<td>0.1554</td>
<td>0.1358</td>
</tr>
<tr>
<td>7</td>
<td>65 x 65</td>
<td>7.5868</td>
<td>7.9945</td>
<td>1.0000</td>
<td>0.1705</td>
<td>0.1485</td>
</tr>
</tbody>
</table>

Table 2. Constants for $P_1$ nonconforming elements on nonnested grids.

<table>
<thead>
<tr>
<th>level $k$</th>
<th>grid</th>
<th>$C_1$</th>
<th>$C_Q$</th>
<th>$C_R$</th>
<th>$\delta_k$</th>
<th>$\delta_k'$</th>
</tr>
</thead>
<tbody>
<tr>
<td>2</td>
<td>3 x 3</td>
<td>10.9296</td>
<td>31.6111</td>
<td>1.0000</td>
<td>0.0965</td>
<td>0.0261</td>
</tr>
<tr>
<td>3</td>
<td>5 x 5</td>
<td>10.5700</td>
<td>26.3688</td>
<td>1.0000</td>
<td>0.0751</td>
<td>0.0741</td>
</tr>
<tr>
<td>4</td>
<td>9 x 9</td>
<td>17.0298</td>
<td>38.4648</td>
<td>1.0000</td>
<td>0.1539</td>
<td>0.1521</td>
</tr>
<tr>
<td>5</td>
<td>17 x 17</td>
<td>20.8358</td>
<td>44.6080</td>
<td>1.0000</td>
<td>0.2421</td>
<td>0.2051</td>
</tr>
<tr>
<td>6</td>
<td>33 x 33</td>
<td>23.3766</td>
<td>52.3772</td>
<td>1.0000</td>
<td>0.2862</td>
<td>0.2422</td>
</tr>
</tbody>
</table>

the trouble of defining nodal (midedge) values of $I_k v_{k-1}$, which is done usually by averaging the values of $v_{k-1}$ at nearby nodes; cf. [4], [13], [14], [20], and [22].

Comparing the data in Tables 1 and 2 we can see the constants $C_1$ and $C_Q$ are much worse for the nonconforming $P_1$ elements. However, the V-cycle and the two-level convergence rates $\delta_k$ and $\delta_k'$ do not differ much between the conforming and nonconforming elements. We note that here the grid size ratio of the fine-to-coarse levels is more than $1/2$—better than that in the nested multigrid method. So the nonnested multigrid convergence rate is better than that of the standard nested multigrid (shown in Table 3).

For Table 3 we have nested grids. We note that because we used one-sided value interpolation operator (since the fine-level midedge points are no longer inside coarse-level triangles) as the intergrid transfer operator, instead of some averaging operators (cf. [13]), the rates of the nonconforming multigrid method (listed in the last two columns) are much worse than that of the conforming method (listed in the middle two columns in Table 3). Otherwise, the difference in rates should be small, as shown in the nonnested cases (listed in Tables 1 and 2). We further remark that, due to the perturbation to the subspace $A_k$-projection, the two-level nonconforming multigrid method is worse than its V-cycle multigrid version in terms of the rate of convergence. In other words, a more accurate coarse-level correction would produce a bigger error to the high-frequency components of the iterative solution on the finer grid. Therefore, if the fine-level smoothing number is not high enough, the multigrid iteration may even diverge. This phenomenon shows up clearly in the last numerical example in this paper.

In Figure 2 we plot the estimated rate $\delta_k$ of the V-cycle multigrid method in (6) and the estimated rate in (19) for the two-level multigrid method, against the actual (computed) rates. Here the estimates $\frac{C_1 C_R}{m C_1 C_R}$ and $\frac{C_1 C_R}{m C_1 C_R / C_Q}$ are both for $P_1$ conforming elements, and computed by numerical data $C_1$ and $C_Q$. The grid level is $k = 5$ in Figure 2.
Table 3. The convergence rate for nested-grid $P_1$ conforming and nonconforming elements.

<table>
<thead>
<tr>
<th>level $k$</th>
<th>grid</th>
<th>$\delta_k(c)$</th>
<th>$\delta'_k(c)$</th>
<th>$\delta_k(nc)$</th>
<th>$\delta'_k(nc)$</th>
</tr>
</thead>
<tbody>
<tr>
<td>2</td>
<td>$4 \times 4$</td>
<td>0.0844</td>
<td>0.0186</td>
<td>0.3211</td>
<td>0.1088</td>
</tr>
<tr>
<td>3</td>
<td>$8 \times 8$</td>
<td>0.1619</td>
<td>0.1332</td>
<td>0.3989</td>
<td>0.3775</td>
</tr>
<tr>
<td>4</td>
<td>$16 \times 16$</td>
<td>0.2143</td>
<td>0.1663</td>
<td>0.4091</td>
<td>0.4469</td>
</tr>
<tr>
<td>5</td>
<td>$32 \times 32$</td>
<td>0.2404</td>
<td>0.1791</td>
<td>0.4359</td>
<td>0.4678</td>
</tr>
</tbody>
</table>

Figure 2. The convergence rate of the V-cycle multigrid

Finally, we would show a counterexample where the number of smoothings $m$ must be sufficiently large, larger than one, depending on a parameter $\sigma$ in the mesh perturbation, in order for the V-cycle nonnested multigrid methods to converge. In the example, we use the cubic Lagrange element to solve the Poisson equation with a homogeneous boundary condition. The domain is the unit square. On the first level, we have only two triangles. We then use the multigrid refinement to generate the higher level meshes. On the fifth level, we perturb the mesh by moving all internal nodes by the mapping $(x, y)/(r/1.5)^\sigma$. In Figure 3, the fine grid is plotted by solid lines and all the coarse grids are plotted (overlapped) by dash lines.

In Table 4 we list the number of V-cycle iterations needed for the $P_3$-finite element iterative solution to reach its approximation accuracy. For the nested case, i.e., $\sigma = 0$ on the finest level, one smoothing is enough to make V-cycle iteration converge, as predicted by the standard multigrid theory. However, when the meshes are perturbed as shown in Figure 3, one smoothing is not enough for the nonnested V-cycle iteration to converge. As shown in Table 4, the number of smoothings $m$
must be larger than 4 (when $\sigma = 0.36$) or 7 (when $\sigma = 0.4$), respectively. Otherwise the V-cycle iterations would diverge. From Table 3, it seems that the when $m$ is large, the converge rates of nested and nonnested V-cycles differ very little. We remark that the reduction rate $\delta_5$ listed in Table 3 does not decrease monotonically when $m$ increases. This is caused by the way we average the error reduction factors by the number of V-cycles:

$$\delta_5 = \frac{1}{n_v} \sum_{i=1}^{n_v} \frac{|u_h - u_i|_{H^1}}{|u_h - u_{i-1}|_{H^1}},$$

where $n_v$ is the number of V-cycles.

In Figure 4, we plot the iterative error before doing a nonnested coarse-level correction and after doing such a correction. In the nonnested coarse-level correction, the low-frequency components of the iterative error are usually reduced well. However, due to nonnestedness, some high-frequency errors would be amplified. This can be seen by comparing the two graphs in Figure 4. Therefore, the fine-level smoothing has to be performed enough times in order for the nonnested multigrid method to converge.

We make a final remark on selecting the counterexample. The difficulty here arises when we use the multigrid refinement to generate meaningful, or likely practical, grids. With reasonable perturbations of the grids, we could not find a case where the $P1$ multigrid V-cycle diverges. After numerous successful tries, we turned to $P2$, $P3$ and high-order elements where one fine-level smoothing is not powerful enough to smooth out the non-$a(\cdot, \cdot)$-projection component of the coarse-level correction.
Table 4. The number of V-cycles and the error reduction rate for nonnested $P_3$ elements.

<table>
<thead>
<tr>
<th>$m$</th>
<th>$# V$-cycle, $\delta_2(\sigma = 0)$</th>
<th>$# V$-cycle, $\delta_2(\sigma = .36)$</th>
<th>$# V$-cycle, $\delta_2(\sigma = .4)$</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>72</td>
<td>$\infty$</td>
<td>$\infty$</td>
</tr>
<tr>
<td>2</td>
<td>36</td>
<td>$\infty$</td>
<td>$\infty$</td>
</tr>
<tr>
<td>3</td>
<td>24</td>
<td>$\infty$</td>
<td>$\infty$</td>
</tr>
<tr>
<td>4</td>
<td>18</td>
<td>$\infty$</td>
<td>$\infty$</td>
</tr>
<tr>
<td>5</td>
<td>14</td>
<td>$&gt;40$</td>
<td>$0.9304$</td>
</tr>
<tr>
<td>6</td>
<td>12</td>
<td>$&gt;40$</td>
<td>$0.8708$</td>
</tr>
<tr>
<td>7</td>
<td>10</td>
<td>13</td>
<td>$0.6203$</td>
</tr>
<tr>
<td>8</td>
<td>8</td>
<td>11</td>
<td>$0.5767$</td>
</tr>
<tr>
<td>9</td>
<td>7</td>
<td>9</td>
<td>$0.5107$</td>
</tr>
<tr>
<td>10</td>
<td>6</td>
<td>8</td>
<td>$0.4733$</td>
</tr>
<tr>
<td>11</td>
<td>6</td>
<td>7</td>
<td>$0.4251$</td>
</tr>
<tr>
<td>12</td>
<td>5</td>
<td>7</td>
<td>$0.4206$</td>
</tr>
<tr>
<td>13</td>
<td>5</td>
<td>6</td>
<td>$0.3673$</td>
</tr>
<tr>
<td>14</td>
<td>5</td>
<td>6</td>
<td>$0.3839$</td>
</tr>
<tr>
<td>15</td>
<td>4</td>
<td>6</td>
<td>$0.4006$</td>
</tr>
<tr>
<td>16</td>
<td>4</td>
<td>5</td>
<td>$0.2917$</td>
</tr>
<tr>
<td>17</td>
<td>4</td>
<td>5</td>
<td>$0.3013$</td>
</tr>
<tr>
<td>18</td>
<td>4</td>
<td>5</td>
<td>$0.3160$</td>
</tr>
<tr>
<td>19</td>
<td>4</td>
<td>5</td>
<td>$0.3341$</td>
</tr>
<tr>
<td>20</td>
<td>4</td>
<td>5</td>
<td>$0.3386$</td>
</tr>
<tr>
<td>21</td>
<td>4</td>
<td>5</td>
<td>$0.3608$</td>
</tr>
<tr>
<td>22</td>
<td>4</td>
<td>5</td>
<td>$0.3729$</td>
</tr>
<tr>
<td>23</td>
<td>3</td>
<td>4</td>
<td>$0.2117$</td>
</tr>
<tr>
<td>24</td>
<td>3</td>
<td>4</td>
<td>$0.2120$</td>
</tr>
<tr>
<td>25</td>
<td>3</td>
<td>4</td>
<td>$0.2159$</td>
</tr>
</tbody>
</table>

Figure 4. The iterative error before and after doing a coarse-level correction.
References


DEPARTMENT OF MATHEMATICS, NATIONAL UNIVERSITY OF SINGAPORE, 2 SCIENCE DRIVE 2, SINGAPORE 117543
E-mail address: scidhy@nus.edu.sg

COLLEGE OF MATHEMATICS AND COMPUTERS, HEBEI UNIVERSITY, 071002, 1 HEZUO ROAD, BAODING, HEBEI, CHINA
E-mail address: gaoshq@amss.ac.cn.

DEPARTMENT OF MATHEMATICS, NATIONAL UNIVERSITY OF SINGAPORE, 2 SCIENCE DRIVE 2, SINGAPORE 117543
E-mail address: scitance@nus.edu.sg.

DEPARTMENT OF MATHEMATICAL SCIENCES, UNIVERSITY OF DELAWARE, NEWARK, DELAWARE 19716
E-mail address: szhang@udel.edu.