Abstract Cascading Multigrid Methods in Besov Spaces

Jingtang Ma and Hermann Brunner

Math. & Stats., Memorial University of Newfoundland, St. John's, NL, Canada A1C 5S7

E-mail: jingtang@math.mun.ca, hermann@math.mun.ca

November 24, 2003

Abstract. In this paper, the general framework of cascading multigrid methods is studied. As an application, the cascading multigrid methods are applied to the heat equation with mild regularity in Besov spaces. Furthermore we analyze a cascading multigrid method for the interior penalty discontinuous Galerkin method.

Mathematics subject classification 1991. 65F10, 65N55, 65N30, 65R20

Keywords. Cascading multigrid method, finite element method, Besov space

1 Introduction

At present there exist three types of multigrid algorithms: the V-cycle algorithm, the W-cycle algorithm and the cascading multigrid algorithm. You may consult Hackbusch [12] and Bramble [6] for the background materials and Brandt [7] and Trottenberg *et al.* [19] for references. The cascading algorithm is the new member of the family of multigrid methods. As a distinctive feature, the algorithm does not need nested correction at all and performs more iterations on coarser levels so as to obtain less iterations on finer levels. The first publication of this algorithm (Deuflhard [10] in 1994) contained rather convincing numerical results, but no theoretical justification. In 1996, Bornemann and Deuflhard [5] provided a theoretical analysis of this algorithm. Later in 1998 and 1999, Shi and Xu [15, 16] generalized the idea and applied it to nonconforming FEM.

It is worthwhile to mention that Bramble [6] contributed the general analysis of the V-cycle and the W-cycle in an abstract setting. To solve the FEM equation (2.1.1):

$$A_k y_k = Q_k g,$$

where A_k is symmetric positive definite. Let λ_{max} and λ_{min} denote the largest and smallest eigenvalue of A_k, respectively. Condition number of A_k is defined by

$$K(A_k) := \frac{\lambda_{max}}{\lambda_{min}}.$$

In view of the process of iteration, we know that the smaller the condition number $K(A_k)$ the more effective the iteration method is. Therefore, we define the precondition operator

$$B_k : V_k \to V_k,$$

such that the condition number $K(B_k A_k)$ is as small as possible. Instead we solve the equation

$$B_k A_k y_k = B_k Q_k g,$$

by the iteration method. Indeed, this approach is more effective. However the new problem arising is how to construct such a precondition operator B_k with as less computational cost as possible. Bramble [6] and the references mentioned there showed the multigrid processes: the V-cycle and the W-cycle are the ideal methods. We shall prove that the cascading multigrid process is also a simple and robust method. Bornemann and Deuflhard [5] provided the numerical examples to compare the cascading multigrid algorithm with the V-cycle algorithm. We analyze this by using the general framework with the assumptions abstracted from the FEM discretization of a given problem. In fact, the abstract multigrid framework provided us a good way to describe the method for more complicated problems. To see this, we shall apply the cascading multigrid algorithms to heat equation with mild regularity in Besov spaces [2] and the equation discretized by the interior penalty discontinuous Galerkin method (see [11] and references therein). As comparison with Shi and Xu [17], a distinctive feature of our method is to use block Jacobi and symmetric Gauss-Seidel iteration as smoothers.

In Section 2, we study the abstract cascading multigrid method with nested spaces and unvarying forms and method in Besov space; in Section 3, we analyze the abstract multigrid method with non-nested forms and the method for the interior penalty discontinuous Galerkin method; in Section 4, we point to some future work, including the extension of our approach to mesh-free methods and to parabolic problems with memory terms.

2 Abstract cascading multigrid method

2.1 Cascading multigrid algorithm

We define a finite-dimensional space V equipped with an inner product $(\cdot,\cdot)$ and a bilinear form $A(\cdot,\cdot)$ with corresponding norms $||\cdot||$ and $|||\cdot|||$. We further assume that $A(\cdot,\cdot)$ is elliptic (i.e., continuous and coercive) and symmetric. For given subspaces V_k satisfying

$$V_0 \subset V_1 \subset \cdots \subset V_J \equiv V.$$

We define the linear operator $A_k : V_k \to V_k$ by

$$(A_k u, v) := A(u, v), \quad \text{for all } u,\ v \in V_k.$$

Obviously A_k is symmetric and positive definite. Furthermore, we define two projectors $P_k : V_J \to V_k$ and $Q_k : V_J \to V_k$ by

$$A(P_k u, v) := A(u, v),$$

and

$$(Q_k u, v) := (u, v),$$

for all $u \in V_J$ and $v \in V_k$. We also need to introduce a generic linear smoothing operator $R_k : V_k \to V_k$ with the assumption $R_0 := A_0^{-1}$. For ease of analysis, we assume also that R_k is symmetric. To solve the equations for each level $k := 0, 1, \ldots, J$:

$$A_k y_k = Q_k g, \tag{2.1.1}$$

we shall now define the preconditioning operator $B \equiv B_J$ by the following cascading multigrid algorithm.

CMG Algorithm I:

0) $B_0 := A_0^{-1}$;

Define B_k implicitly in terms of B_{k-1}, for $k = 1, \ldots, J$:

1) For $l = 1, \ldots, m_k$, we set

$$y_k^l := y_k^{l-1} + R_k(Q_k g - A_k y_k^{l-1}).$$

Here $y_k^0 := B_{k-1} Q_{k-1} g$.

2) Solve $B_k Q_k g := y_k^{m_k}$ for B_k.

In order to analyze CMG Algorithm I, we need to make the following four natural assumptions.

Assumption 2.1.1. *Let λ_k denote the maximum eigenvalue of A_k, i.e.,*

$$\lambda_k := \sup_{v \in V_k} \frac{(A_k v, v)}{(v, v)}.$$

For some $\alpha \in [0, 1]$, there exists C_α independent of k such that

$$(A_k^{1-\alpha}(I - P_{k-1})u, (I - P_{k-1})u) \le C_\alpha \lambda_k^{-\alpha} A(u, u), \quad \textit{for all } u \in V_k.$$

Assumption 2.1.2. *Let $K_k := I - R_k A_k$, $R_{k,\omega} := \omega \lambda_k^{-1} I$ and $K_{k,\omega} := I - R_{k,\omega} A_k$. There exists $\omega \in (0, 1]$ such that*

$$A(K_k v, K_k v) \le A(K_{k,\omega/2} v, K_{k,\omega/2} v), \quad \textit{for all } v \in V_k.$$

Hence it holds that

$$A(K_k^m v, K_k^m v) \le A(K_{k,\omega/2}^m v, K_{k,\omega/2}^m v), \quad \textit{for nonnegative number } m \textit{ and for } v \in V_k.$$

Assumption 2.1.3. *There is a constant $b > 0$ so that*

$$c\frac{\lambda_J}{b^{J-k}} \leq \lambda_k \leq C\frac{\lambda_J}{b^{J-k}}.$$

Here and throughout the paper, c and C denote generic positive constants which are independent of k.

Assumption 2.1.4. *For $\beta > 0$, let*

$$m_k := \lfloor \beta^{J-k} m_J \rfloor,$$

where $\lfloor \cdot \rfloor$ means the greatest integer function.

Remark 2.1.1. *We see that Assumption 2.1.1 is reasonable, if we refer to the elliptic boundary value problems with less than full regularity (cf. [5] [16]).*

Remark 2.1.2. *Bramble [6] tells us that block Jacobi and (symmetric) Gauss-Seidel are the examples of R_k satisfying Assumption 2.1.2.*

We need also the following lemma about a property of $K_{k,\omega}$, which is necessary for the analysis of the cascading multigrid method.

Lemma 2.1.1. *The iterates of $K_{k,\omega}$ possess the following two properties:*

$$A(K_{k,\omega}^{m_k} v, K_{k,\omega}^{m_k} v) \leq A(v, v), \tag{2.1.2}$$

and

$$A(K_{k,\omega}^{m_k} v, K_{k,\omega}^{m_k} v) \leq C\frac{\lambda_k^{-\alpha}}{m_k^{\alpha}}(A^{1-\alpha} v, v), \tag{2.1.3}$$

for $\omega \in (0, 1]$ and $\alpha \in (0, 1]$.

Proof. See (pp. 164, Brenner and Scott [8]).

Theorem 2.1.1. *Let B_J be defined by CMG Algorithm I and Assumptions {2.1.1, 2.1.2, 2.1.3, 2.1.4} hold. Then we have*

$$A((I - B_J A_J)u, u) \leq \begin{cases} C \cdot \dfrac{1}{1-(b/\beta)^{\alpha/2}} \cdot \dfrac{\lambda_J^{-\alpha/2}}{m_J^{\alpha/2}} A(u, u), & \textit{for } \beta > b, \\ C \cdot J \cdot \dfrac{\lambda_J^{-\alpha/2}}{m_J^{\alpha/2}} A(u, u), & \textit{for } \beta = b, \end{cases}$$

for all $u \in V_J$.

Proof. We begin with the estimate:

$$
\begin{aligned}
y_J - B_J A_J y_J &= y_J - B_J Q_J g = y_J - y^{m_J} \\
&= K_J^{m_J}(y_J - y^0) = K_J^{m_J}(y_J - B_{J-1}Q_{J-1}g) \\
&= K_J^{m_J}(y_J - y_{J-1}) + K_J^{m_J}(y_{J-1} - B_{J-1}Q_{J-1}g) \\
&\quad \ldots\ldots \\
&= K_J^{m_J}(y_J - y_{J-1}) + K_J^{m_J}K_{J-1}^{m_{J-1}}(y_{J-1} - y_{J-2}) \\
&+ \ldots + K_J^{m_J}K_{J-1}^{m_{J-1}}\ldots K_1^{m_1}(y_1 - y_0).
\end{aligned}
\tag{2.1.4}
$$

In view of (2.1.1), we obtain

$$
y_k = A_k Q_k Q_J^{-1} A_J y_J = A_k Q_k A_J y_J, \quad \text{for } k = 0, \ldots, J-1. \tag{2.1.5}
$$

Since Bramble [6] proved

$$
Q_k A_J = A_k P_k, \quad \text{on } V_k,
$$

we have

$$
y_k = P_k y_J, \quad \text{for } k = 0, \ldots, J-1.
$$

Bringing these equations into (2.1.4) leads to

$$
I - B_J A_J = \sum_{k=1}^{J} \prod_{i=1}^{J-k+1} K_{J+1-i}^{m_{J+1-i}}(P_k - P_{k-1}).
$$

Consequently by using the Cauchy-Schwartz inequality, we obtain

$$
\begin{aligned}
& A((I - B_J A_J)u, u) \qquad (2.1.6) \\
= & \sum_{k=1}^{J} A\left(\prod_{i=1}^{J-k+1} K_{J+1-i}^{m_{J+1-i}}(P_k - P_{k-1})u, u\right) \\
\le & \sum_{k=1}^{J} [A\left(\prod_{i=1}^{J-k+1} K_{J+1-i}^{m_{J+1-i}}(P_k - P_{k-1})u, \prod_{i=1}^{J-k+1} K_{J+1-i}^{m_{J+1-i}}(P_k - P_{k-1})u\right)]^{1/2} \cdot [A(u,u)]^{1/2}.
\end{aligned}
$$

In view of {(3.1.2), Assumption 2.1.2, (3.1.3), Assumption 2.1.1}, we estimate

$$
\begin{aligned}
& A\left(\prod_{i=1}^{J-k+1} K_{J+1-i}^{m_{J+1-i}}(P_k - P_{k-1})u, \prod_{i=1}^{J-k+1} K_{J+1-i}^{m_{J+1-i}}(P_k - P_{k-1})u\right) \\
\le & A\left(K_{k,\omega/2}^{m_k}(P_k - P_{k-1})u, K_{k,\omega/2}^{m_k}(P_k - P_{k-1})u\right) \\
\le & C\frac{\lambda_k^{-\alpha}}{m_k^{\alpha}}(A^{1-\alpha}(P_k - P_{k-1})u, (P_k - P_{k-1})u) \\
\le & C\frac{\lambda_k^{-\alpha}}{m_k^{\alpha}}\lambda_k^{-\alpha}A(u,u).
\end{aligned}
\tag{2.1.7}
$$

(2.1.6) and (2.1.7) lead to

$$A((I - B_J A_J)u, u) \leq C \left(\sum_{k=1}^{J} \frac{\lambda_k^{-\alpha}}{m_k^{\alpha/2}} \right) A(u, u). \tag{2.1.8}$$

Combining (2.1.8), Assumption 2.1.3 and Assumption 2.1.4, we obtain our desired result.

Theorem 2.1.2. *Let $N_k := \dim(V_k)$ and assume $N_{k+1} \geq a N_k$. Then the computational cost of the CMG Algorithm I is proportional to*

$$\sum_{k=1}^{J} m_k N_k \leq \begin{cases} C \cdot \dfrac{1}{1 - \beta/a} \cdot m_J N_J, & \text{for } \beta < a, \\ C \cdot J \cdot m_J N_J, & \text{for } \beta = a. \end{cases} \tag{2.1.9}$$

Proof. $N_{k+1} \geq a N_k$ implies $N_k \leq a^{-(J-k)} N_J$. Recalling Assumption 2.1.4, we estimate

$$\sum_{k=1}^{J} m_k N_k \leq C \sum_{k=1}^{J} \left(\frac{\beta}{a} \right)^{J-k} \cdot m_J N_J.$$

Hence (2.1.9) is true.

2.2 Application

Shi and Xu [17] analyzed the cascading multigrid method for the heat equation with Richardson and Conjugate Gradient iteration as smoothers. We shall solve the heat equation with mild regularity by the abstract cascading multigrid method with Block Jacobi and symmetric Gauss-Seidel iteration as smoothers.

Let $\Omega \subset \mathbb{R}^d$ $(d = 1, 2, 3)$ be a polygonal domain with the largest corner of measure ϖ . We consider the Heat equation:

$$\begin{cases} \dfrac{\partial u(x,t)}{\partial t} - \Delta u(x,t) = f(x,t), & \text{on } \Omega \times [0, T], \\ u(x,t) = 0, & \text{on } \partial\Omega \times [0, T], \\ u(x, 0) = u_0(x), & \text{on } \Omega, \end{cases} \tag{2.2.1}$$

where $f \in H^{1-\alpha}(\Omega)$. The weak form of (2.2.1) is to find $u \in H_0^1(\Omega)$, $u(x, 0) = u_0(x) \in H^{1+\alpha}(\Omega) \cap H_0^1(\Omega)$ such that

$$(\frac{\partial u}{\partial t}, v) + B(u, v) = (f, v), \quad \forall v \in H_0^1(\Omega), \ t \in [0, T], \tag{2.2.2}$$

where the bilinear form B is

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

and

$$(f, v) = \int_\Omega f v dx.$$

We use the backward Euler scheme for the time-stepping. Let

$$I_\tau := \{t_n : 0 =: t_0 < t_1 < \cdots < t_N := T\}$$

be the mesh on I, with

$$\tau_n := t_n - t_{n-1},\ \tau := \max_{(n)} \tau_n.$$

Moreover, we set

$$w := u^n - u^{n-1},\ (g, v) := (f, v) - B(u^{n-1}, v).$$

Then (2.2.2) leads to the weak form with time-stepping: To find $w \in H_0^1(\Omega)$ such that

$$A_\tau(w, v) = (g, v), \quad \forall v \in H_0^1(\Omega), \tag{2.2.3}$$

where

$$A_\tau(w, v) := \tau^{-1}(w, v) + B(w, v).$$

We conclude that (2.2.3) has a unique solution $w \in H^{1+\alpha}(\Omega) \cap H_0^1(\Omega)$ which satisfies

$$||w||_{H^{1+\alpha}} \leq C||g||_{H^{1-\alpha}},$$

where $0 < \alpha < \alpha_0 = \frac{\pi}{\varpi}$ (see Ciarlet [9] and Johnson [13]).

Let $\mathcal{T}_k$ $(k := 0, 1, \ldots, J)$ be a quasi-uniform triangular partition of Ω with mesh size $h_k = h_0 2^{-k}$. Let V_k denote the P1 conforming finite element space on $\mathcal{T}_k$. It is well known that

$$V_0 \subset V_1 \subset \ldots \subset V_J.$$

We derive the discrete form of (2.2.3): Find $w_k \in V_k$ such that

$$A_\tau(w_k, v) = (g, v), \quad \forall v \in V_k. \tag{2.2.4}$$

Define

$$(A_k w_k, v) := A_\tau(w_k, v), \quad \forall w_k,\ v \in V_k.$$

Then (2.2.4) can be expressed by

$$A_k w_k = g_k, \tag{2.2.5}$$

where $g_k \in V_k,\ (g_k, v) := (g, v),\ \forall v \in V_k$. It is well-known that

$$|||w - w_k|||_k \leq C h_k^\alpha (1 + \tau^{-1} h_k)^{1/2} ||g||_{H^{1-\alpha}}. \tag{2.2.6}$$

We can see that

$$\lambda_k = \mathcal{O}(h_k^{-2}) + \tau^{-1}.$$

Since [3] observed that some usual iterative methods, like the Richardson iteration, can already guarantee good convergence, we therefore only consider the case

$$\lambda_J^{-1} \leq \tau \leq \frac{\lambda_J^{-1}}{\gamma_0}, \quad \text{for some } \gamma_0 \in (0,1). \tag{2.2.7}$$

Now we shall verify Assumptions {(2.1.1), (2.1.2), (2.1.3), (2.1.4)} hold true in this case. Let

$$A := A_\tau.$$

From [6] and [18], we conclude that Assumption 2.1.1 is satisfied. Assumption 2.1.3 can be easily written as

$$c\left(\frac{\lambda_J}{2^{2(J-k)}} + \frac{2^{2(J-k)}-1}{2^{2(J-k)}}\tau^{-1}\right) \leq \lambda_k \leq C\left(\frac{\lambda_J}{2^{2(J-k)}} + \frac{2^{2(J-k)}-1}{2^{2(J-k)}}\tau^{-1}\right). \tag{2.2.8}$$

Assumption 2.1.2 and {(3.1.2), (3.1.3)} concern only R_k. Therefore they remain true as discussed in Section 2.1. Thus, recalling (2.1.8), we obtain

$$\begin{aligned} & A((I - B_J A_J)w, w) \\ \leq\ & C\left(\sum_{k=1}^{J} \frac{\lambda_k^{-\alpha}}{m_k^{\alpha/2}}\right) A(w,w) \\ \leq\ & C\left(\frac{1}{1+\gamma_0}\right)^{\alpha} \frac{h_J^\alpha}{m_J^{\alpha/2}} \sum_{k=1}^{J} \left(\frac{2^\alpha}{\beta^{\alpha/2}}\right)^k A(w,w). \end{aligned} \tag{2.2.9}$$

So (2.2.9) yields the following theorem. The last estimate in (2.2.9) reveals that the attainable (optimal or nearly optimal) order depends on the value of β; for $\beta = 4$ the estimate will contain the factor J, i.e. the number of the grid points. Compare also Bornemann and Deuflhard [5].

Theorem 2.2.1. *Under Assumption 2.1.4, we have the following convergence estimate of CMG Algorithm I for (2.2.5):*

$$A((I - B_J A_J)w, w) \leq \begin{cases} C(\alpha) \cdot \dfrac{1}{1 - \left(\frac{2}{\beta^{1/2}}\right)^{\alpha}} \cdot \dfrac{h_J^\alpha}{m_J^{\alpha/2}} A(w,w), & \textit{when } \beta > 4, \\ C(\alpha) \cdot J \cdot \dfrac{h_J^\alpha}{m_J^{\alpha/2}} A(w,w), & \textit{when } \beta = 4, \end{cases}$$

with $C(\alpha) := C(1/(1+\gamma_0))^\alpha$. These estimates hold for all $w \in V_J$.

In view of

$$N_{k+1} \geq 2^d N_k,$$

we conclude that Theorem 2.1.2 remains true with $a = 2^d$.

For a discussion on how to choose β so that the cascading multigrid method has optimal accuracy and complexity, we refer to shi and Xu [17] or Bornemann and Deuflhard [5].

2.3 The cascading multigrid method in Besov spaces

For $s \in (0,1)$, define the interpolation space

$$[X, Y]_{s,\infty} := \{w \in Y : \sup_{t>0} t^{-2s} K(t,w)^2 < \infty\},$$

and the K-functional by

$$K(t,w) := \inf_{w_0 \in X} (||w_0||_X^2 + t^2 ||w - w_0||_Y^2)^{1/2}.$$

Recall the equation (2.2.3): To find $w \in H_0^1(\Omega)$ such that

$$A_\tau(w,v) = (g,v), \quad \forall v \in H_0^1(\Omega), \tag{2.3.1}$$

where

$$A_\tau(w,v) := \tau^{-1}(w,v) + B(w,v).$$

For $0 < \alpha < \alpha_0 = \frac{\pi}{\varpi}$, we have the regularity

$$||w||_{H^{1+\alpha}} \leq C||g||_{H^{1-\alpha}}, \quad \forall g \in H^{1-\alpha}(\Omega).$$

For the critical case $\alpha = \alpha_0$, Bacuta, Bramble and Xu [2] proved the estimate

$$||w||_{B^{1+\alpha_0}(\Omega)} \leq C||g||_{B_{01}^{-1+\alpha_0}(\Omega)}, \quad \forall g \in B_{01}^{-1+\alpha_0}(\Omega), \tag{2.3.2}$$

where $B^{1+\alpha_0}(\Omega)$ is a standard Besov space and $B_{01}^{-1+\alpha_0}(\Omega)$ was defined in [2]. Correspondingly, the convergence estimate of the P1 finite element method for (2.3.1) is

$$\begin{aligned} |||w - w_k|||_k &\leq Ch_k^{\alpha_0}(1 + \tau^{-1}h_k)^{1/2} ||w||_{[H^2(\Omega)\cap H_0^1(\Omega), H_0^1(\Omega)]_{1-\alpha_0,\infty}} \\ &\leq Ch_k^{\alpha_0}(1 + \tau^{-1}h_k)^{1/2} ||g||_{B_{01}^{-1+\alpha_0}(\Omega)}. \end{aligned} \tag{2.3.3}$$

We begin our analysis with equation (2.1.7):

$$\begin{aligned} & A\left(K_{k,w/2}^{m_k}(P_k - P_{k-1})w, K_{k,w/2}^{m_k}(P_k - P_{k-1})w\right) \\ = \ & \sup_{t>0} t^{-2(1-\alpha_0)} \inf_{w_0 \in V_J} [A\left(K_{k,w/2}^{m_k}(P_k - P_{k-1})w_0, K_{k,w/2}^{m_k}(P_k - P_{k-1})w_0\right) \\ + \ & t^2 A\left(K_{k,w/2}^{m_k}(P_k - P_{k-1})(w - w_0), K_{k,w/2}^{m_k}(P_k - P_{k-1})(w - w_0)\right)]. \end{aligned} \tag{2.3.4}$$

We use Lemma 2.1.1 and Assumption 2.1.1 with $\alpha = 0$ and $\alpha = 1$ to get the equations

$$\begin{aligned}
& A\left(K_{k,w/2}^{m_k}(P_k - P_{k-1})w_0, K_{k,w/2}^{m_k}(P_k - P_{k-1})w_0\right) \\
\leq \quad & C\frac{\lambda_k^{-1}}{m_k}((P_k - P_{k-1})w_0, (P_k - P_{k-1})w_0) \leq C\frac{\lambda_k^{-2}}{m_k}A(w_0, w_0), \qquad (2.3.5)
\end{aligned}$$

and

$$A\left(K_{k,w/2}^{m_k}(P_k - P_{k-1})(w - w_0), K_{k,w/2}^{m_k}(P_k - P_{k-1})(w - w_0)\right) \leq CA(w - w_0, w - w_0). \qquad (2.3.6)$$

Hence {(2.3.4), (2.3.5), (2.3.6)} lead to

$$\begin{aligned}
& A\left(K_{k,w/2}^{m_k}(P_k - P_{k-1})w, K_{k,w/2}^{m_k}(P_k - P_{k-1})w\right) \\
\leq \quad & C\frac{\lambda_k^{-2\alpha_0}}{m_k^{\alpha_0}}\sup_{t>0}(m_k^{1/2}\lambda_k t)^{-2(1-\alpha_0)}\inf_{w_0\in V_J}[A(w_0, w_0) + (m_k^{1/2}\lambda_k t)^2 A(w - w_0, w - w_0)] \\
= \quad & C\frac{\lambda_k^{-2\alpha_0}}{m_k^{\alpha_0/2}}A(w, w). \qquad (2.3.7)
\end{aligned}$$

Therefore we obtain

$$A((I - B_J A_J)w, w) \leq C\left(\sum_{k=1}^{J}\frac{\lambda_k^{-\alpha_0}}{m_k^{\alpha_0/2}}\right)A(w, w).$$

This indicates the following theorem:

Theorem 2.3.1. *Under Assumption 2.1.4, we have the convergence estimate of CMG Algorithm I for (2.3.1)*

$$A((I - B_J A_J)w, w) \leq \begin{cases} C \cdot \dfrac{1}{1 - \left(\frac{2}{\beta^{1/2}}\right)^{\alpha_0}} \cdot \dfrac{h_J^{\alpha_0}}{m_J^{\alpha_0/2}}A(w, w), & \text{for } \beta > 4, \\ C \cdot J \cdot \dfrac{h_J^{\alpha_0}}{m_J^{\alpha_0/2}}A(w, w), & \text{for } \beta = 4, \end{cases}$$

for all $w \in V_J$.

3 Abstract cascading multigrid method with non-nested spaces and varying inner products and bilinear forms

In the previous sections we assumed that $V_0 \subset V_1 \subset \cdots \subset V_J \equiv V$ and $(\cdot, \cdot)$ and $A(\cdot, \cdot)$ are defined on all V_k. In this section, we allow the spaces V_k to be not necessarily nested and define a symmetric positive definite bilinear form $A_k(\cdot, \cdot)$ and an inner product $(\cdot, \cdot)_k$ corresponding to $0 \leq k \leq J$.

3.1 The cascading multigrid algorithm

We reset

(1) Let $V_0, \ldots, V_J$ be J finite dimensional spaces with $V_J \equiv V$, which are not necessarily nested.

(2) Define J linear operators

$$I_k : V_{k-1} \to V_k, \quad k = 1, \ldots, J,$$

which connect the spaces.

(3) Let $(\cdot,\cdot)_k$ be an inner product on $V_k \times V_k$, with induced norm $||\cdot||_k$, and let $A_k(\cdot,\cdot)$ be symmetric and elliptic bilinear form on $V_k \times V_k$ with $A_J(\cdot,\cdot) \equiv A(\cdot,\cdot)$ and induced norm $|||\cdot|||_k$.

(4) Define $A_k : V_k \to V_k$ by

$$(A_k u, v)_k := A_k(u, v), \quad \text{for all } v \in V_k.$$

(5) Define $P_{k-1} : V_k \to V_{k-1}$ and $Q_{k-1} : V_k \to V_{k-1}$ by

$$A_{k-1}(P_{k-1}u, v) := A_k(u, I_k v),$$

and

$$(Q_{k-1}u, v)_{k-1} := (u, I_k v)_k,$$

for all $v \in V_{k-1}$.

(6) Let R_k be a linear symmetric operator $R_k : V_k \to V_k$ with $R_0 := A_0^{-1}$.

We shall solve the equation (2.1.1) on each level $k := 0, 1, \ldots, J$ by the cascading multigrid algorithm.

Define the precondition operator $B \equiv B_J$ by

CMG Algorithm II:

0) $B_0 := A_0^{-1}$;

Define B_k in terms of B_{k-1}, for $k = 1, \ldots, J$:

1) For $l = 1, \ldots, m_k$, we set

$$y_k^l := y_k^{l-1} + R_k(Q_k g - A_k y_k^{l-1}).$$

Here $y_k^0 := I_k B_{k-1} Q_{k-1} g$.

3) $B_k Q_k g := y_k^{m_k}$.

Before we analyze CMG Algorithm II, we add the following three reasonable assumptions.

Assumption 3.1.1. *We suppose*

$$A_k(I_k u, I_k u) \le A_{k-1}(u,u), \quad \text{for all } u \in V_{k-1}. \tag{3.1.1}$$

Using the Cauchy-Schwartz inequality, we can easily prove that (3.1.1) holds true if and only if

$$A_{k-1}(P_{k-1}u, P_{k-1}u) \le A_k(u,u), \quad \text{for all } u \in V_k.$$

Assumption 3.1.2. *Let λ_k denote the maximum eigenvalue of A_k, i.e.,*

$$\lambda_k := \sup_{v \in V_k} \frac{(A_k v, v)_k}{(v,v)_k}.$$

For some $\alpha \in (0,1]$, there exists a constant C_α, independent of k, such that

$$(A_k^{1-\alpha}(I - I_k P_{k-1})u, (I - I_k P_{k-1})u)_k \le C_\alpha \lambda_k^{-\alpha} A_k(u,u), \quad \text{for all } u \in V_k$$

(compare with Assumption 2.1.1).

Assumption 3.1.3. *Let $K_k := I - R_k A_k$, $R_{k,\omega} := \omega \lambda_k^{-1} I$ and $K_{k,\omega} := I - R_{k,\omega} A_k$. There exists $\omega \in (0,1]$ such that*

$$A_k(K_k v, K_k v) \le A_k(K_{k,\omega/2} v, K_{k,\omega/2} v), \quad \text{for all } v \in V_k.$$

Hence it holds that

$$A_k(K_k^m v, K_k^m v) \le A_k(K_{k,\omega/2}^m v, K_{k,\omega/2}^m v), \quad \text{for nonnegative number } m \text{ and for } v \in V_k$$

(compare with Assumption 2.1.2).

By using the same argument as in the proof of Lemma 2.1.1, we can prove the following lemma.

Lemma 3.1.1. *For $w \in (0,1]$ and $\alpha \in (0,1]$, $K_{k,\omega}^{m_k}$ satisfies the following two properties:*

$$A_k(K_{k,\omega}^{m_k} v, K_{k,\omega}^{m_k} v) \le A_k(v,v), \tag{3.1.2}$$

and

$$A_k(K_{k,\omega}^{m_k} v, K_{k,\omega}^{m_k} v) \le C \frac{\lambda_k^{-\alpha}}{m_k^\alpha} (A_k^{1-\alpha} v, v)_k, \tag{3.1.3}$$

for $\omega \in (0,1]$ and $\alpha \in (0,1]$.

Now we are ready to present our main results.

Theorem 3.1.1. *Let B_J be defined by CMG Algorithm II and Assumptions {3.1.1, 3.1.2, 3.1.3, 2.1.3, 2.1.4} hold. Then we have*

$$A((I - B_J A_J)u, u) \leq \begin{cases} C \cdot \dfrac{1}{1-(b/\beta)^{\alpha/2}} \cdot \dfrac{\lambda_J^{-\alpha/2}}{m_J^{\alpha/2}} A(u,u), & \text{for } \beta > b, \\ C \cdot J \cdot \dfrac{\lambda_J^{-\alpha/2}}{m_J^{\alpha/2}} A(u,u), & \text{for } \beta = b, \end{cases}$$

for all $u \in V_J$.

Proof. We first estimate:

$$\begin{aligned} y_J - B_J A_J y_J &= y_J - B_J Q_J g = y_J - y^{m_J} \\ &= K_J^{m_J}(y_J - y^0) = K_J^{m_J}(y_J - I_J B_{J-1} Q_{J-1} g) \\ &= K_J^{m_J}(y_J - I_J y_{J-1}) + K_J^{m_J}(I_J y_{J-1} - I_J B_{J-1} Q_{J-1} g) \\ &\quad \ldots\ldots \\ &= K_J^{m_J}(y_J - I_J y_{J-1}) + (K_J^{m_J} I_J) K_{J-1}^{m_{J-1}}(y_{J-1} - I_{J-1} y_{J-2}) \\ &+ \ldots + (K_J^{m_J} I_J)(K_{J-1}^{m_{J-1}} I_{J-1}) \ldots (K_2^{m_2} I_2) K_1^{m_1}(y_1 - I_1 y_0). \end{aligned} \tag{3.1.4}$$

From (2.1.1):

$$A_k y_k = Q_k g,$$

we derive

$$(A_k y_k, I_k v)_k = (Q_k g, I_k v)_k, \quad \forall v \in V_{k-1}.$$

Hence,

$$A_k(y_k, I_k v) = (Q_k g, I_k v)_k, \quad \forall v \in V_{k-1}.$$

In view of the definitions of P_k and Q_k, we obtain

$$A_{k-1}(P_{k-1} y_k, v) = (Q_{k-1} g, v)_{k-1}, \quad \forall v \in V_{k-1},$$

which leads to

$$A_{k-1} P_{k-1} y_k = Q_{k-1} g.$$

Therefore, we have

$$y_{k-1} = P_{k-1} y_k, \quad \text{for } k = 0, \ldots, J-1. \tag{3.1.5}$$

Combining (3.1.5) and (3.1.4) we find

$$I - B_J A_J = K_J^{m_J}(I - I_J P_{J-1}) + \sum_{k=1}^{J-1} \prod_{i=1}^{J-k} (K_{J+1-i}^{m_{J+1-i}} I_{J+1-i}) \cdot K_k^{m_k} \cdot (I - I_k P_{k-1}) \cdot \prod_{i=1}^{J-k} P_{J-i}.$$

By the Cauchy-Schwartz inequality and Lemma 3.1.1,

$$
\begin{aligned}
& A((I - B_J A_J)u, u) && (3.1.6)\\
\leq\ & [A(K_J^{m_J}(I - I_J P_{J-1})u, K_J^{m_J}(I - I_J P_{J-1})u)]^{1/2} \cdot [A(u,u)]^{1/2} \\
+\ & \sum_{k=1}^{J-1} [A_k \left((K_k^{m_k} \cdot (I - I_k P_{k-1}) \cdot \prod_{i=1}^{J-k} P_{J-i})u, (K_k^{m_k} \cdot (I - I_k P_{k-1}) \cdot \prod_{i=1}^{J-k} P_{J-i})u \right)]^{1/2} \cdot [A(u,u)]^{1/2}.
\end{aligned}
$$

From Assumption 3.1.3, Lemma 3.1.1, and Assumption 3.1.2, we have

$$
\begin{aligned}
A\left(K_J^{m_J}(I - I_J P_{J-1})u, K_J^{m_J}(I - I_J P_{J-1})u\right) &\leq C \frac{\lambda_J^{-\alpha}}{m_J^{\alpha}} \left(A^{1-\alpha}(I - I_J P_{J-1})u, (I - I_J P_{J-1})u\right) \\
&\leq C \frac{\lambda_J^{-\alpha}}{m_J^{\alpha}} A(u,u), \qquad (3.1.7)
\end{aligned}
$$

and in addition, using Assumption 3.1.1,

$$
\begin{aligned}
& A_k \left((K_k^{m_k} \cdot (I - I_k P_{k-1}) \cdot \prod_{i=1}^{J-k} P_{J-i})u, (K_k^{m_k} \cdot (I - I_k P_{k-1}) \cdot \prod_{i=1}^{J-k} P_{J-i})u \right) \\
\leq\ & C \frac{\lambda_k^{-\alpha}}{m_k^{\alpha}} \left(A_k^{1-\alpha} (I - I_k P_{k-1}) \prod_{i=1}^{J-k} P_{J-i} u, (I - I_k P_{k-1}) \prod_{i=1}^{J-k} P_{J-i} u \right)_k \\
\leq\ & C \frac{\lambda_k^{-\alpha}}{m_k^{\alpha}} A_k \left(\prod_{i=1}^{J-k} P_{J-i} u, \prod_{i=1}^{J-k} P_{J-i} u \right) \leq C \frac{\lambda_k^{-\alpha}}{m_k^{\alpha}} A(u,u). \qquad (3.1.8)
\end{aligned}
$$

Hence (3.1.6), (3.1.7) and (3.1.8) lead to

$$
A((I - B_J A_J)u, u) \leq C \left(\sum_{k=1}^{J} \frac{\lambda_k^{-\alpha}}{m_k^{\alpha/2}} \right) A(u,u). \qquad (3.1.9)
$$

Combining (3.1.9) and Assumption 2.1.3 and Assumption 2.1.4, we obtain our desired result.

The computational cost estimate theorem is the same as Theorem 2.1.2.

3.2 Application for the interior penalty discontinuous Galerkin method

The V-cycle algorithm for the interior penalty discontinuous Galerkin method was presented in the paper by Gopalakrishnan and Kanschat [11], which was based on Arnold [1]. In this section, we analyze the CMG Algorithm II for the heat equation (2.2.1) with the interior penalty discontinuous Galerkin discretization. We will essentially use the notations in Section 3.1.

Let $\mathcal{T}_k$ ($k := 0, 1, \ldots, J$) be a quasi-uniform triangular partition of Ω with the mesh size $h_k = h_0 2^{-k}$. We define the multilevel spaces

$$
V_0 \subset V_1 \subset \ldots \subset V_J \equiv V,
$$

by

$$V_k := \{v : v|_T \in \mathcal{P}^{(m)}(T),\ \forall T \in \mathcal{T}_k\},$$

where $\mathcal{P}^{(m)}(T)$ denote the polynomial with degree not exceeding m on T.

To describe the interior penalty discontinuous Galerkin method, we need the spaces

$$H_0^1(\mathcal{T}_k) := \{v \in L^2(\Omega) : v|_T \in H^1(T) \text{ and } v|_{\partial\Omega\cap T} = 0,\ \forall T \in \mathcal{T}_k\}.$$

Let $\mathcal{E}_k$ denote the set of edges of the triangulation $\mathcal{T}_k$. If $e \in \mathcal{E}_k$ is an interior edge, denoting by n_e one of the two unit normal vectors at e, we define jumps and averages of normal derivatives (for $x \in e$) of $v \in H_0^1(\mathcal{T}_k)$ by

$$[v]_e(x) := \lim_{\delta\to 0+} [v(x - \delta n_e) - v(x + \delta n_e)],$$

and

$$\langle \partial_n v\rangle_e(x) := \frac{1}{2} \lim_{\delta\to 0+} [n_e \cdot \nabla v(x - \delta n_e) + n_e \cdot \nabla v(x + \delta n_e)],$$

while if $e \subseteq \partial\Omega$, we fix n_e to be the outward normal vector and let

$$[v]_e(x) := \lim_{\delta\to 0+} v(x - \delta n_e) \text{ and } \langle \partial_n v\rangle_e := \lim_{\delta\to 0+} n_e \cdot \nabla v(x - \delta n_e).$$

Define $B_k(\cdot,\cdot)$ on $H_0^1(\mathcal{T}_k) \times H_0^1(\mathcal{T}_k)$ by

$$\begin{aligned} &B_k(u, v) \\ := &\sum_{T\in\mathcal{T}_k} (\nabla u, \nabla v)_T + \sum_{e\in\mathcal{E}_k} \left(\frac{\sigma}{\ell_e}([u], [v])_e - (\langle\partial_n u\rangle, [v])_e - ([u], \langle\partial_n v\rangle)_e \right). \end{aligned} \tag{3.2.1}$$

Here ℓ_e denotes the length of the edge e and σ is a positive parameter to be chosen. The weak form of (2.2.1) is defined by: Find $u \in H_0^1(\mathcal{T}_k),\ u(x, 0) = u_0(x)$, such that

$$(\frac{\partial u}{\partial t}, v) + B_k(u, v) = (f, v), \quad \forall v \in H_0^1(\mathcal{T}_k),\ t \in [0, T]. \tag{3.2.2}$$

As in Section 2.2, we use the backward Euler scheme for the time-stepping of (3.2.2). Then we derive the weak form with time-stepping: To find $w \in H_0^1(\mathcal{T}_k)$ such that

$$A_{\tau,k}(w, v) := (g, v), \quad \forall v \in H_0^1(\mathcal{T}_k), \tag{3.2.3}$$

where

$$A_{\tau,k}(w, v) := \tau^{-1}(w, v) + B_k(w, v),$$

and

$$(g, v) := (f, v) - B_k(u^{n-1}, v).$$

The interior penalty discontinuous Galerkin approximation to (3.2.3) is then given by: To find $w_k \in V_k$ such that

$$A_{\tau,k}(w_k, v) = (g, v), \quad \forall v \in V_k. \tag{3.2.4}$$

Define

$$(A_k w_k, v) := A_k(w_k, v) := A_{\tau,k}(w_k, v), \quad \forall w_k,\ v \in V_k. \tag{3.2.5}$$

Then (3.2.4) can be expressed by

$$A_k w_k = g_k, \tag{3.2.6}$$

where $g_k \in V_k,\ (g_k, v) := (g, v),\ \forall v \in V_k$. From [1], we can easily obtain that

$$|||w - w_k|||_k \le C h_k^{\alpha}(1 + \tau^{-1} h_k)^{1/2} ||g||_{H^{1-\alpha}}. \tag{3.2.7}$$

From [11], we know that

$$c(\sigma + 1)h_k^{-2} + \tau^{-1} \le \lambda_k \le C(\sigma + 1)h_k^{-2} + \tau^{-1}.$$

Hence we have

$$c\left((\sigma+1)\frac{\lambda_J}{2^{2(J-k)}} + \frac{2^{2(J-k)}-1}{2^{2(J-k)}}\tau^{-1}\right) \le \lambda_k \le C\left((\sigma+1)\frac{\lambda_J}{2^{2(J-k)}} + \frac{2^{2(J-k)}-1}{2^{2(J-k)}}\tau^{-1}\right). \tag{3.2.8}$$

In [11] it was also proved that Assumption 3.1.2 holds true. Assumption 3.1.3 and Lemma 3.1.1 are guaranteed, since they only concern the smoother operator R_k. We assume τ satisfies (2.2.7). So by recalling (3.1.9), we derive that

$$\begin{aligned} & A((I - B_J A_J)w, w) \\ \le\ & C\left(\sum_{k=1}^{J} \frac{\lambda_k^{-\alpha}}{m_k^{\alpha/2}}\right) A(w, w) \\ \le\ & C\left(\frac{1}{1+\sigma+\gamma_0}\right)\frac{h_J^{\alpha}}{m_J^{\alpha/2}} \sum_{k=1}^{J}\left(\frac{2^{\alpha}}{\beta^{\alpha/2}}\right)^{k} A(w, w). \end{aligned}$$

The following theorem is therefore established.

Theorem 3.2.1. *Under Assumption 2.1.4, we have the convergence estimate of CMG Algorithm II for (3.2.6)*

$$A((I - B_J A_J)w, w) \le \begin{cases} C \cdot \dfrac{1}{1 - \left(\frac{2}{\beta^{1/2}}\right)^{\alpha}} \cdot \dfrac{h_J^{\alpha}}{m_J^{\alpha/2}} A(w, w), & \text{for } \beta > 4, \\ C \cdot J \cdot \dfrac{h_J^{\alpha}}{m_J^{\alpha/2}} A(w, w), & \text{for } \beta = 4, \end{cases}$$

for all $w \in V_J$.

The computational cost estimate is that of Theorem 2.1.2 with $a = 2^d$.

4 Concluding remarks

The foregoing discussion reveals that the abstract setting of the cascading multigrid method provides a more feasible way to establish the convergence theorem of the method for the time-dependent problems with mild regularity and for problems that are discretized by other new FEM methods. We believe it is also possible to extend the abstract method for the mesh-free method described in, e.g., [4, 14]. The applications of the method for integro-differential equations will be studied elsewhere.

Acknowledgement:
This research was supported in part by the Natural Science and Engineering Research Council of Canada (NSERC).

References

[1] D.N. Arnold, An interior penalty finite element method with discontinuous elements, *SIAM J. Numer. Anal.*, 19 (1982), 742–760.

[2] C. Bacuta, J.H. Bramble and J. Xu, Regularity estimates for elliptic boundary value problems in Besov spaces, *Math. Comp.*, 72 (2002), 1577–1595.

[3] R.E. Bank and T. Dupont, An optimal order process for solving finite element equations, *Math. Comp.*, 36 (1980), 35–51.

[4] T. Belytschko, Y. Krongauz, D. Organ, M. Fleming and P. Krysl, Meshless methods: an overview and recent developments, *Comput. Methods Appl. Mech. Engrg.*, 139 (1996), 3–47.

[5] F. Bornemann and P. Deuflhard, The cascadic multigrid method for elliptic problems, *Numer. Math.*, 75 (1996), 135–152.

[6] J.H. Bramble, *Multigrid Methods*, John Wiley & Sons, New York, 1993.

[7] A. Brandt, Multiscale scientific computation: review 2001, in *Multiscale and Multiresolution Methods*, 3–95, Lecture Notes in Comput. Sci. Eng., 20, Springer, Berlin, 2002.

[8] S.C. Brenner and L.R. Scott, *The Mathematical Theory of Finite Element Methods*, Springer-Verlag, New York, 2002.

[9] P. Ciarlet, *The Finite Element Method for Elliptic Problems*, North-Holland, Amsterdam, 1978.

[10] P. Deuflhard, Cascadic conjugate gradient methods for elliptic partial differential equations: algorithm and numerical results, *Domain Decomposition Methods in Scientific and Engineering Computing (University Park, PA, 1993)*, 29–42, Contemp. Math., 180, Amer. Math. Soc., Providence, RI, 1994.

[11] J. Gopalakrishnan and G. Kanschat, A multilevel discontinuous Galerkin method, *Numer. Math.*, 95 (2003), 527–550.

[12] W. Hackbusch, *Multi-Grid Methods and Applications*, Springer-Verlag, New York, 1985.

[13] C. Johnson, *Numerical Solution of Partial Differential Equations by the Finite Element Method*, Cambridge University Press, New York, 1987.

[14] G.R. Liu, *Mesh Free Methods: Moving beyond the Finite Element Method*, CRC Press, New York, 2003.

[15] Z. Shi and X. Xu, Cascadic multigrid for elliptic problems, *East-West J. Numer. Math.*, 6 (1998), 309–318.

[16] Z. Shi and X. Xu, Cascadic multigrid for elliptic problems, *East-West J. Numer. Math.*, 7 (1999), 199–211.

[17] Z. Shi and X. Xu, Cascadic multigrid for parabolic problems, *J. Comput. Math.*, 18 (2000), 450–459.

[18] V. Thomée, *Galerkin Finite Element Methods for Parabolic Problems*, Springer-Verlag, New York, 1997.

[19] U. Trottenberg, C.W. Oosterlee and A. Schüller, *Multigrid*, with contributions by A. Brandt, P. Oswald and K. Stüben, Academic Press, San Diego, 2001.