Convergence estimates for multigrid algorithms with SSC smoothers and applications to overlapping domain decomposition

E. Aulisa
Department of Mathematics and Statistics, Texas Tech University

G. Bornia
Department of Mathematics and Statistics, Texas Tech University

S. Calandrini
Department of Mathematics and Statistics, Texas Tech University

G. Capodaglio
Department of Mathematics and Statistics, Texas Tech University

Abstract
In this paper we study convergence estimates for a multigrid algorithm with smoothers of successive subspace correction (SSC) type, applied to symmetric elliptic PDEs. First, we revisit a general convergence analysis on a class of multigrid algorithms in a fairly general setting, where no regularity assumptions are made on the solution. In this framework, we are able to explicitly highlight the dependence of the multigrid error bound on the number of smoothing steps. For the case of no regularity assumptions, this represents a new addition to the existing theory. Then, we analyze successive subspace correction smoothing schemes for a set of uniform and local refinement applications with either nested or non-nested overlapping subdomains. For these applications, we explicitly derive bounds for the multigrid error, and identify sufficient conditions for these bounds to be independent of the number of multigrid levels. For the local refinement applications, finite element grids with arbitrary hanging nodes configurations are considered. The analysis of these smoothing schemes is cast within the far-reaching multiplicative Schwarz framework.

Keywords: Multigrid; SSC algorithm; Domain Decomposition; Hanging nodes; Local refinement; V-cycle.

*Corresponding Author
Email address: giacomo.capodaglio@ttu.edu (G. Capodaglio)

1108 Memorial Circle, Department of Mathematics and Statistics, Texas Tech University, Lubbock TX 79409, USA

Preprint submitted July 24, 2018
1. Introduction

Multigrid algorithms have been introduced in the literature since the 1960's with pioneering works such as [1, 2, 3, 4]. A wide literature of both theoretical and computational works has been developed ever since, driven by appealing features such as optimal computational complexity [4, 5]. Multigrid methods have been studied for the approximate solution of partial differential equations in various discretization schemes, starting with finite differences [1, 3, 4] and then moving to finite elements [6, 7] and other settings. The first results were obtained for elliptic operators of either symmetric [6] or non-symmetric type [8, 9, 10].

Convergence proofs of multigrid algorithms usually rely on two properties referred to as the smoothing and the approximation property [11, 12, 13]. The former is related to the definition of the smoothing operator involved in the algorithm, while the latter is usually proved assuming full elliptic regularity for the solution of the partial differential equation. A breakthrough in the convergence analysis took place with [14]. In this work the elliptic regularity assumption has been dropped. The error bound obtained is not optimal in the sense that it becomes worse as the number of multigrid levels increases; moreover, no dependence of the bound on the number of smoothing iterations is shown. Further work has been done in this direction by Bramble and Pasciak [15], where they showed that optimal convergence can be obtained provided that partial regularity assumptions are made. However, no dependence on the number of smoothing iterations was reported yet. In [16], convergence estimates were obtained by the same authors for the case of a multigrid algorithm with non-symmetric subspace correction smoothers under no regularity assumptions. The error bound obtained for applications to both uniform and local refinement showed quadratic dependence on the total number of multigrid spaces but not on the number of smoothing steps. An improvement in addressing this matter was made in [17], where the author showed that the multigrid error bound is optimal and can be improved when increasing the number of smoothing iterations, under partial regularity assumptions and using a Richardson relaxation scheme. More recently, further work on multigrid methods that rely on minimal regularity assumptions has been done in [18], where graded meshes obtained by a variant of the newest vertex bisection method are considered.

This work aims at first to further contribute to the description of multigrid methods by carrying out a general convergence analysis that does not require any regularity assumption. Only three clear assumptions on the smoothing error operator are identified which produce a multigrid error bound that shows dependence on the number of smoothing steps and on the continuity constants of the smoothing error operators in the topology of the energy norm. As mentioned in the abstract, the explicit dependence of the multigrid error bound on the number of smoothing iterations is a new result under no-regularity assumptions.
The setup of this framework is then used to analyze smoothing schemes of successive subspace correction (SSC) type. A unifying scheme encompassing successive subspace correction algorithms is given by Xu in [19]. See also [16] for an analysis of smoothers in a general framework to which subspace correction smoothers of either additive or multiplicative type belong. We will study both uniform and local refinement applications with arbitrary hanging nodes configurations, to show under what conditions on the subdomain solvers multigrid convergence is achieved, and when it is possible to obtain optimal multigrid error bounds, i.e., independent of the total number of levels. For the uniform refinement case, such results upgrade the ones in [16]. For the two local refinement applications, we derive ad-hoc decompositions of finite element spaces and set suitable choices of approximate subdomain solvers. These are needed when dealing with hanging nodes that are introduced by the local refinement procedure. In the first local refinement case, we construct a decomposition that has the advantage of being easy to implement and suitable for standard finite element codes, but it does not allow freedom in the choice of the subdomains on which the subspaces are built. A second decomposition requires additional work to ensure continuity of the finite element solution and so it requires a non-standard finite element implementation. However, it enables a choice of the subdomains that does not depend on the multigrid level. Numerical results for a similar choice of decomposition, together with a complexity analysis of the resulting algorithm, have been provided in [20]. In such a work, the smoothing procedure is carried out only locally, rather than at all nodes of a given multigrid level, as we do in this theory. A convergence analysis for this local smoothing approach is available in [16]. In both the aforementioned local refinement applications studied in this paper, the multigrid error bound shows a quadratic dependence on the multigrid level and this agrees with what was found in [16]. Furthermore, we explicitly show a dependence on the number of smoothing steps. This allows us to identify conditions on the number of smoothing iterations that guarantee convergence as well as optimality of the error bound. Basically, these conditions establish a balance between the action of the smoothing error and the number of smoothing steps. In the applications, we obtain smoothing error bounds that are either constant or increase tending to one with increasing level, thus corresponding to a poorer smoothing action with increasing level. In order to have convergence, only one smoothing iteration at each level is sufficient. Nevertheless, we find that optimality of the multigrid error bound may be obtained only with a quadratically increasing number of smoothing steps. Thus, the convergence deterioration of the smoother with an increasing number of levels may be compensated by an ad-hoc number of smoothing steps in order to obtain optimality. We remark that when a local smoothing procedure is conducted, increasing the number of smoothing iterations at a given multigrid level would only improve the multigrid error bound up to a given saturation value. As a consequence, a deterioration of the error bound that goes with the total number of levels could not be balanced by increasing the number of smoothing steps.

The outline of the paper is as follows. In Section 2 the multigrid algorithm is described together with a general convergence theory. Such theory is based
on three assumptions on the smoothing error operator and needs no regularity. Section 3 illustrates the algorithm used for the smoothing iteration and shows how it can be related to the multigrid convergence theory. Uniform and local refinement applications of the analysis described in the previous sections are presented in Section 4, where convergence bounds are obtained for the specific cases. Finally, we draw our conclusions.

2. The multigrid algorithm

In this section we describe the multigrid algorithm subject to our analysis. Throughout the paper, the total number of levels will be denoted as $J$. For $k = 0, \ldots, J$, let $V_k$ be a finite-dimensional vector space such that

$$V_0 \subset V_1 \subset \cdots \subset V_J,$$

and let $(\cdot, \cdot)$ and $a(\cdot, \cdot)$ be two symmetric positive definite (SPD) bilinear forms on $V_k$. Hence, both bilinear forms are inner products on $V_k$. Let $|| \cdot || = \sqrt{(\cdot, \cdot)}$ and $|| \cdot ||_E = \sqrt{a(\cdot, \cdot)}$ be the corresponding induced norms. Associated with these inner products, let us also define the operators $Q_k : V_J \to V_k$ and $P_k : V_J \to V_k$ as the orthogonal projections with respect to $(\cdot, \cdot)$ and $a(\cdot, \cdot)$ respectively, namely, for all $v \in V_J$ and all $w \in V_k$

$$(Q_k v, w) = (v, w), \quad a(P_k v, w) = a(v, w).$$

Note that from this definition it follows that

$$a((I - P_k)v, w) = 0 \quad \text{for all} \quad w \in V_k. \quad (3)$$

The multigrid algorithm seeks solutions of the following problem: given $f \in V_J$, find $u \in V_J$ such that

$$a(u, v) = (f, v) \quad \text{for all} \quad v \in V_J. \quad (4)$$

Before we can present the multigrid algorithm studied in this paper, we need to introduce a few operators that will be used in the description of the method.

For $k = 0, \ldots, J$, define the operators $A_k : V_k \to V_k$ as

$$(A_k u, v) = a(u, v) \quad \text{for all} \quad u, v \in V_k. \quad (5)$$

The operator $A_k$ is SPD with respect to $(\cdot, \cdot)$ as a consequence of the symmetry and positive definiteness of $a(\cdot, \cdot)$. If we set $f_k = Q_k f$, then at level $k$ the problem we want to solve consists in finding $u_k \in V_k$ such that

$$A_k u_k = f_k. \quad (6)$$

The prolongation $I_{k-1}^k : V_{k-1} \to V_k$ and restriction $I_k^{k-1} : V_k \to V_{k-1}$ operators are defined for all $v \in V_{k-1}$ and all $w \in V_k$ by

$$I_{k-1}^k v = v, \quad (I_k^{k-1} w, v) = (w, I_{k-1}^k v). \quad (7)$$
We are now ready to present the multigrid algorithm considered in this paper. Let $B_k : V_k \to V_k$ denote a smoothing operator. Associated to $B_k$ we can define a smoothing error operator $S_k : V_k \to V_k$ as $S_k = I - B_k A_k$, whose properties will be discussed later in detail.

For $k = 0, \ldots, J$, let $MG_k : V_k \times V_k \to V_k$ be the multigrid operators. The purpose of the operators $MG_k$ is to yield an approximate solution to (6). They are defined here in a recursive manner.

**Algorithm 1** (V-cycle multigrid). Let $z^{(0)}_k, f_k \in V_k$.

If $k = 0$, $MG_0(z^{(0)}_0, f_0) = A_0^{-1} f_0$ (namely, the exact solution is obtained).

For $k \geq 1$, $MG_k(z^{(0)}_k, f_k)$ is obtained recursively as follows.

1. **Pre-smoothing.** For $1 \leq i \leq m_k$, let
   \[ z^{(i+1)}_k = z^{(i)}_k + B_k (f_k - A_k z^{(i)}_k) \, . \]

2. **Error Correction.** Let $\bar{f}_k = t^{k-1}_k \left( f_k - A_k z^{(m_k)}_k \right)$, $q_{k-1} = MG_{k-1}(0, \bar{f}_k)$. Then,
   \[ z^{(m_k+1)}_k = z^{(m_k)}_k + t^{k}_k q_{k-1} \, . \]

3. **Post-smoothing.** For $m_k + 2 \leq i \leq 2 m_k + 1$, let
   \[ z^{(i+1)}_k = z^{(i)}_k + B_k (f_k - A_k z^{(i)}_k) \, . \]

Note that the total number of pre-smoothing iterations $m_k$ is assumed to be dependent on the level $k$. Also, we are assuming the same number $m_k$ of pre-smoothing and post-smoothing steps at each level. We also remark that we consider a symmetric version of the multigrid algorithm as in [14], in the sense that both pre-smoothing and post-smoothing are performed. Since we have only one iteration for the error correction step, this algorithm is referred to as V-cycle [12].

**2.1. Convergence analysis**

Here we present a general convergence analysis of the multigrid algorithm. We do so by introducing sufficient assumptions on the smoothing error operator $S_k$ for the derivation of the convergence results. It is then clear that the convergence properties of the multigrid algorithm are intimately dependent on the smoothing procedures.

Before listing the assumptions, we recall the expressions of the error operators associated to $B_k$ and $MG_k$. Let $z^{(i)}_k$ be the output of a pre- or post-smoothing iteration at level $k$. If we denote the associated error as $e^{(i)}_k = u_k - z^{(i)}_k$, then substituting for $z^{(i)}_k$ we have

\[ e^{(i)}_k = u_k - z^{(i-1)}_k - B_k (f_k - A_k z^{(i-1)}_k) = S_k e^{(i-1)}_k \, , \quad (8) \]
so that the effect of the smoothing step can be described as
\[ e_k^{(m_k)} = S_k^{m_k} e_k^{(0)}. \] (9)

The multigrid error operator \( E_k : V_k \rightarrow V_k \) associated to \( MG_k \) is defined recursively as
\[ E_0 = 0, \quad E_k = S_k^{m_k} \left[ I - (I - E_{k-1})P_{k-1} \right] S_k^{m_k}. \] (10)

Note that \( E_0 \) is assumed to be zero since we are using a direct solver at level \( k = 0 \). This means that \( B_0 = A_0^{-1} \) and \( S_0 = 0 \). Here we summarize the properties of the \( E_k \) operators. For a proof see [21] or [14] for the special case where \( z_k^{(0)} = 0 \).

**Proposition 1.** Let \( z_k^{(0)} \in V_k \), and let \( u_k \) be the exact solution to \( A_k u_k = f_k \). Then
\[ u_k - MG_k(z_k^{(0)}, f_k) = E_k \left( u_k - z_k^{(0)} \right), \quad k \geq 0. \]
Moreover, the \( E_k \)'s are symmetric positive semidefinite with respect to \( a(\cdot, \cdot) \) for \( k \geq 0 \).

We now state sufficient hypotheses for multigrid convergence. As the expression of the multigrid error operator (10) suggests, once the operators \( P_k \) are given by the differential problem at hand, multigrid convergence is affected by the properties of \( S_k \) and by the number of smoothing steps \( m_k \). These features are reflected in the following assumptions.

**Assumption 1.** For all \( k = 1, \ldots, J \), \( S_k \) is a symmetric positive semidefinite operator on \( V_k \) with respect to \( a(\cdot, \cdot) \). This means that for all \( v, w \in V_k \) we have
\[ a(S_kv, w) = a(v, S_kw) \quad \text{and} \quad a(S_kv, v) \geq 0. \] (11)

**Assumption 2.** For all \( k = 1, \ldots, J \) there exists a number \( \delta_k \) with \( 0 < \delta_k < 1 \) such that
\[ a(S_kv, v) \leq \delta_k a(v, v) \quad \text{for all} \quad v \in V_k. \] (12)

**Assumption 3.** For all \( k = 1, \ldots, J \), the finite sequence \( \psi_k = m_k(1 - \delta_k) \) is non-increasing, where \( m_k \) is the number of smoothing steps per level and \( \delta_k \) is the quantity in Assumption 2.

2.1.1. Smoothing and approximation properties

Assumptions 1 and 2 guarantee that the operators \( S_k \) satisfy certain monotonicity properties given by Lemmas 1 and 2 below. These properties will lead to the smoothing property of Lemma 3.

**Lemma 1.** Let Assumptions 1 and 2 hold. Let \( \alpha \) and \( \beta \) be two integers such that \( 0 \leq \alpha \leq \beta \). Then,
\[ a(S_k^\alpha v, v) \leq a(S_k^\beta v, v) \quad \text{for all} \quad v \in V_k. \] (13)
Proof. We will prove it for $\beta = \alpha + 1$ and the result will then follow by induction. By (11), $S_k$ is positive semidefinite with respect to $a(\cdot, \cdot)$, therefore also $S_k^{\frac{\alpha}{2}}$ is. Then there is a unique positive semidefinite square root operator $S_k^{\frac{\alpha}{2}}$, see [22]. By the symmetry of $S_k$ it follows that $S_k^{\frac{\alpha}{2}}$ is symmetric as well. Considering also (12), we have that for $v \in V_k$

$$a(S_k^{\alpha+1}v, v) = a(S_k^{\alpha}S_k v, v) = a(S_k^{\frac{\alpha}{2}} S_k^{\frac{\alpha}{2}} S_k v, v)$$

$$= a(S_k^{\frac{\alpha}{2}} S_k v, S_k^{\frac{\alpha}{2}} v) = a(S_k S_k^{\frac{\alpha}{2}} v, S_k^{\frac{\alpha}{2}} v)$$

$$\leq a(S_k^{\frac{\alpha}{2}} v, S_k^{\frac{\alpha}{2}} v) = a(S_k^{\alpha} v, v).$$

\[\square\]

Using the previous lemma, we can prove the next result.

**Lemma 2.** Let Assumptions 1 and 2 hold. Let $\alpha$ and $\beta$ be two integers such that $0 \leq \alpha \leq \beta$. Then,

$$a((I - S_k)S_k^{\beta} v, v) \leq a((I - S_k)S_k^{\alpha} v, v) \text{ for all } v \in V_k. \quad (14)$$

Proof. As we did before, we will prove it for $\beta = \alpha + 1$ and the result will follow by induction.

$$a((I - S_k)S_k^{\alpha+1} v, v) = a((I - S_k)S_k^{\alpha+1} v, (I - S_k + S_k)v)$$

$$= a((I - S_k)S_k^{\alpha+1} v, (I - S_k)v) + a((I - S_k)S_k^{\alpha+1} v, S_k v)$$

$$= a(S_k^{\alpha+1} (I - S_k)v, (I - S_k)v) + a((I - S_k)S_k^{\alpha+1} v, S_k v)$$

$$\leq a(S_k^{\alpha} (I - S_k)v, (I - S_k)v) + a((I - S_k)S_k^{\alpha+1} v, S_k v) \quad \text{(by (13))}$$

$$= a(S_k^{\alpha} (I - S_k)v, v) - a(S_k^{\alpha} (I - S_k)v, S_k v) + a((I - S_k)S_k^{\alpha+1} v, S_k v)$$

$$= a(S_k^{\alpha} (I - S_k)v, v) - a(S_k^{\alpha+1} (I - S_k)v, v) + a(S_k^{\alpha+1} (I - S_k)v, S_k v)$$

$$= a(S_k^{\alpha+1} (I - S_k)v, v) - a(S_k^{\alpha+1} (I - S_k)v, (I - S_k)v)$$

$$\leq a(S_k^{\alpha+1} (I - S_k)v, v) \quad \text{($S_k^{\alpha+1}$ is positive semidefinite, see [22])}$$

$$= a((I - S_k)S_k^{\alpha} v, v).$$

\[\square\]

Now we are ready to prove the smoothing property of the operator $S_k$.

**Lemma 3** (Smoothing property). Let Assumptions 1 and 2 hold. Let $v \in V_k$. Then,

$$a((I - S_k)S_k^{2m_k} v, v) \leq \frac{1}{2m_k} a((I - S_k^{2m_k}) v, v). \quad (15)$$
Proof. By Lemma 2 we get

\[(2m_k)a((I - S_k)S_k^{2m_k}v, v) = a((I - S_k)S_k^{2m_k}v, v) + \cdots + a((I - S_k)S_k^{2m_k}v, v) \]

2m_k times

\[\leq a((I - S_k)v, v) + a((I - S_k)v, v) + \cdots + a((I - S_k)v, v)\]

\[= a((I - S_k + S_k - S_k + \cdots + S_k^{2m_k-1} - S_k^{2m_k})v, v)\]

\[= a((I - S_k^{2m_k})v, v).\]

\[\square\]

Before we can show a bound on the error \(E_k\), we need to establish the approximation property.

Lemma 4 (Approximation property). Let Assumptions 1 and 2 hold, and let \(w \in V_k\). Then,

\[a((I - P_{k-1})w, w) \leq \left( \frac{1}{1 - \delta_k} \right) a((I - S_k)w, w). \tag{16}\]

Proof. Let \(y = (I - P_{k-1})w\). Note that from the definition of \(P_{k-1}\) and from the nestedness of the spaces (1) we have that

\[a(y, P_{k-1}w) = 0. \tag{17}\]

By Assumption 2 we have

\[a((I - S_k)w, w) \geq (1 - \delta_k) a(w, w)\]

\[= (1 - \delta_k) a(y + P_{k-1}w, y + P_{k-1}w)\]

\[= (1 - \delta_k) (a(y, y) + a(P_{k-1}w, P_{k-1}w)) \tag{by (17)}\]

\[\geq (1 - \delta_k) a((I - P_{k-1})w, (I - P_{k-1})w)\]

\[= (1 - \delta_k) a((I - P_{k-1})w, w) \tag{by (17)}.\]

\[\square\]

Notice that in this setting the approximation property given by Lemma 4 is dependent on the smoothing property. This is in contrast with other analyses of multigrid methods in which the smoothing and the approximation properties are derived independently [21]. As a consequence of the approximation property, we have the following result.

Lemma 5. Let Assumptions 1 and 2 hold. Let \(v \in V_k\). Then,

\[a((I - P_{k-1})S_k^{m_k}v, (I - P_{k-1})S_k^{m_k}v) \leq \left( \frac{1}{1 - \delta_k} \right) \frac{1}{2m_k} a((I - S_k^{2m_k})v, v). \tag{18}\]
Proof. We have
\[
a((I - P_{k-1})S_k^m v, (I - P_{k-1})S_k^m v) \\
= a((I - P_{k-1})S_k^m v, S_k^m v) \quad \text{(from the definition of } P_{k-1}) \\
\leq \left( \frac{1}{1 - \delta_k} \right) a((I - S_k)S_k^m v, S_k^m v) \quad \text{(by Lemma 4)} \\
= \left( \frac{1}{1 - \delta_k} \right) a((I - S_k)S_k^{2m} v, v) \\
\leq \left( \frac{1}{1 - \delta_k} \right) \frac{1}{2m_k} a((I - S_k^{2m}) v, v) \quad \text{(by Lemma 3)}.
\]

\[\square\]

2.1.2. Error bound

We are now in a position to obtain a bound on the multigrid error operator $E_J$ that gives convergence.

**Theorem 1.** Let Assumptions 1, 2 and 3 hold. For $k = 0, 1, \ldots, J$ let

\[
\gamma_k = \frac{1}{1 + 2m_k(1 - \delta_k)}. \quad (19)
\]

Then, if $v \in V_J$,

\[
a(E_J v, v) \leq \gamma_J a(v, v). \quad (20)
\]

**Proof.** By Assumption 3, we have

\[
\gamma_{k-1} \leq \gamma_k. \quad (21)
\]

The proof will be done by induction as in [21]. For $k = 0$, $E_0 = 0$ so the result is obvious. By induction assume that

\[
a(E_{J-1} v, v) \leq \gamma_{J-1} a(v, v) \quad \forall v \in V_{J-1}.
\]
Now consider \( v \in V_j \), then
\[
\begin{align*}
    a(E_J v, v) &= a(S^m_J v, S^m_J v) - a(P_{J-1} S^m_J v, P_{J-1} S^m_J v) \\
    &= a(E_J v, S^m_J v) - a(E_J v, P_{J-1} S^m_J v) \\
    &= a((I - P_{J-1}) S^m_J v, (I - P_{J-1}) S^m_J v) \\
    &+ a(E_J v, P_{J-1} S^m_J v) \\
    &
    \leq a((I - P_{J-1}) S^m_J v, (I - P_{J-1}) S^m_J v) \\
    &+ \gamma_{J-1} a(P_{J-1} S^m_J v, P_{J-1} S^m_J v) \\
    &\leq a((I - P_{J-1}) S^m_J v, (I - P_{J-1}) S^m_J v) \\
    &+ \gamma_{J-1} a(P_{J-1} S^m_J v, P_{J-1} S^m_J v) \\
    &= (1 - \gamma_J) a((I - P_{J-1}) S^m_J v, (I - P_{J-1}) S^m_J v) \\
    &+ \gamma_J a((I - P_{J-1}) S^m_J v, (I - P_{J-1}) S^m_J v) \\
    &+ \gamma_J a(P_{J-1} S^m_J v, P_{J-1} S^m_J v) \\
    &= (1 - \gamma_J) a((I - P_{J-1}) S^m_J v, (I - P_{J-1}) S^m_J v) \\
    &+ \gamma_J a(S^m_J v, S^m_J v) \\
    &\leq \left( \frac{1}{2m_J (1 - \delta_J)} \right) (1 - \gamma_J) a((I - S^{2m_J}) v, v) \\
    &+ \gamma_J a(S^m_J v, S^m_J v) \\
    &= \gamma_J a((I - S^{2m_J}) v, v) + \gamma_J a(S^m_J v, S^m_J v) \\
    &= \gamma_J a(v, v) - \gamma_J a(S^m_J v, S^m_J v) + \gamma_J a(S^m_J v, S^m_J v) \\
    &= \gamma_J a(v, v)
\end{align*}
\]

It is evident that the convergence of the multigrid algorithm is dependent on Assumption 3, which lies both on the number of smoothing steps \( m_k \) and on the constants \( \delta_k \) of the smoothing error operator, jointly. No other parameters affect the convergence rate. Since the behavior of \( \delta_k \) is determined by the choice of \( S_k \), different choices on \( m_k \) can be taken subsequently so that (21) holds. For instance, if the \( S_k \) are such that \( \delta_k \) is non-decreasing, then it is sufficient to take \( m_k \) to be non-increasing. In fact, if \( \delta_k \geq \delta_{k-1} \) and \( m_k \leq m_{k-1} \), Assumption 3 holds since
\[
m_k (1 - \delta_k) \leq m_{k-1} (1 - \delta_{k-1}).
\]  

Although sufficient, a non-increasing \( m_k \) is not necessary. Also, observe that a particular case of non-increasing \( m_k \) is \( m_1 = m_2 = \cdots = m_J = m \). In this case it is directly visible how an increasing \( m \) can lead to better convergence rates, as well as an increasing number of multilevel spaces \( J \) can lead to worse convergence. While this last situation was shown in [14], to the best of our knowledge
the first feature was not shown in similar multigrid frameworks without, or with minimal regularity assumptions (see, e.g., [14, 18, 17]).

In Section 3 a characterization of $\delta_k$ will be given when the smoothing process is chosen to be a successive subspace correction algorithm. From this characterization, proper choices of the number of smoothing steps $m_k$ can lead to convergence and, in addition, to optimal (i.e., with a value of $\gamma_J$ that is independent of $J$) multigrid error bounds. Since $m_k \in \mathbb{N}$ and $\delta_k \in \mathbb{R}$, our analysis suggests that the determination of an optimal multigrid error bound may take place by a proper choice of $m_k$ if and only if the quantity $1 - \delta_k$ is inversely proportional to an integer-valued function of $k$. The achievement of an optimal convergence bound seems otherwise impossible, unless a radically new setup of the multigrid algorithm is formulated that is oriented to that purpose.

3. Successive Subspace Correction (SSC) algorithms

Now we describe the Successive Subspace Correction (SSC) algorithm. The SSC algorithm is an iterative method to approximate the solution of SPD linear systems [19]. We link the multigrid convergence theory of the previous section with the SSC theory by using smoothers of subspace correction type for the multigrid algorithm 1. The SSC algorithm yields an approximate solution to (6) and is based on a decomposition of the finite-dimensional space as an algebraic sum of subspaces. In the multigrid algorithm presented above, smoothing is performed at each level $k = 1, \ldots, J$, therefore we decompose each $V_k$ using subspaces $V^i_k \subset V_k$ such that

$$V_k = \sum_{i=0}^{p_k} V^i_k = \left\{ v \mid v = \sum_{i=0}^{p_k} v^i_k, v^i_k \in V^i_k \right\}. \quad (23)$$

Notice that the number of subspaces $p_k$ is in general different for each level. In order to present the algorithm, we first define for all $i$, with $u \in V_k$ and $u^i_k, v^i_k \in V^i_k$, the operators

$$Q^i_k : V_k \to V^i_k, \quad P^i_k : V_k \to V^i_k, \quad A^i_k : V^i_k \to V^i_k,$$

$$(Q^i_k u, v^i_k) = (u, v^i_k), \quad a(P^i_k u, v^i_k) = a(u, v^i_k),$$

$$(A^i_k u^i_k, v^i_k) = (A_k u^i_k, v^i_k).$$

It follows from its definition that $A^i_k$ is an SPD operator. Moreover, as a consequence of the above definitions we have that

$$A^i_k P^i_k = Q^i_k A_k. \quad (24)$$

Hence if $u_k$ is the exact solution of (6), then $P^i_k u_k = u^i_k$ will be the solution of

$$A^i_k u^i_k = f^i_k. \quad (25)$$
where \( f_k^i = Q_k^i f_k \). Equation (25) is in general solved approximately, therefore we introduce for all \( i \) the operators
\[
R_k^i : V_k^i \to V_k^i, \quad T_k^i : V_k \to V_k^i, \quad T_k^i := R_k^i Q_k^i A_k = \bar{R}_k^i \bar{A}_k^i \bar{P}_k^i.
\]
The operators \( R_k^i \) act as approximate inverses of \( A_k^i \). If \( R_k^i \) is taken to be an exact solver then \( T_k^i = \bar{P}_k^i \). When no confusion arises, we drop the subscript \( k \) for \( p_k \) as well as for the operators \( Q_k^i, P_k^i, A_k^i, R_k^i \) and \( T_k^i \). We now define the SSC algorithm.

**Algorithm 2** (Successive Subspace Correction Algorithm.). Let \( z^0 \in V_k \) be given. Then \( z^{\alpha+1} \) is obtained in \( p + 1 \) substeps starting from \( z^\alpha \) by
\[
z^{\alpha+1} - \frac{1}{p+1} = z^{\alpha+1} - \frac{1}{p+1} + R^i Q^i (f_k - A_k z^{\alpha+1} - \frac{1}{p+1}),
\]
for \( i = p, \ldots, 0 \).

The error operator associated to this algorithm is denoted as \( \hat{E}_p \). If \( u_k \) is the exact solution of (6), then for \( i = p, \ldots, 0 \) we have
\[
(u_k - z^{\alpha+1} - \frac{1}{p+1}) = (I - T^i)(u_k - z^{\alpha+1} - \frac{1}{p+1})
\]
This yields
\[
(z_k - z^{\alpha+1}) = \hat{E}_p(z_k - z^\alpha), \quad \hat{E}_p = (I - T^0)(I - T^1) \cdots (I - T^p).
\]
The symmetric version of the SSC algorithm is given here.

**Algorithm 3** (Symmetric Successive Subspace Correction Algorithm.). Let \( z^0 \in V_k \) be given. First, \( z^{\alpha+\frac{1}{2}} \) is obtained in \( p + 1 \) substeps starting from \( z^\alpha \) by
\[
z^{\alpha+\frac{1}{2}} = z^{\alpha+\frac{1}{2}} + R^i Q^i (f_k - A_k z^{\alpha+\frac{1}{2}} + \frac{1}{p+1}),
\]
for \( i = p, \ldots, 0 \). Then, \( z^{\alpha+1} \) is obtained in \( p + 1 \) substeps starting from \( z^{\alpha+\frac{1}{2}} \) by
\[
z^{\alpha+1} + \frac{1}{2} \frac{1}{p+1} = z^{\alpha+1} + \frac{1}{2} \frac{1}{p+1} + R^i Q^i (f_k - A_k z^{\alpha+1} + \frac{1}{2} \frac{1}{p+1}),
\]
for \( i = 0, \ldots, p \).

The error operator of this algorithm is denoted as \( \hat{E}_p^s \) and is given by
\[
\hat{E}_p^s = (I - T^p) \cdots (I - T^1)(I - T^0)^2(I - T^1) \cdots (I - T^p).
\]
Because of the symmetry requirements for the smoother in Assumption 1, we will fit the smoother within the symmetric SSC framework.
3.1. Convergence analysis

We recall the main convergence result about the SSC algorithm, whose proof can be found in [19]. First, we introduce sufficient assumptions.

**Assumption 4 (Bound on \( w_1 \)).** The operators \( R^i \) are SPD with respect to \( (\cdot, \cdot) \) and satisfy \( w_1 < 2 \), where \( w_1 = \max_{i=0,\ldots,p} \rho(R^i A^i) \), \( \rho(R^i A^i) \) being the spectral radius of \( R^i A^i \) and \( p \) being the number of subspaces in the decomposition (23).

**Assumption 5 (Existence of \( K_0 \)).** There exists \( K_0 \) such that for any \( v \in V_k \) there exists a decomposition \( v = \sum_{i=0}^{p} v_i \), with the property
\[
\sum_{i=0}^{p} ((R^i)^{-1}v_i, v_i) \leq K_0 (A^k v, v). \tag{32}
\]

**Assumption 6 (Existence of \( K_1 \)).** Given the same \( p \) as in Assumption 5, there exists \( K_1 \) such that for any \( S \subset \{0,1,\ldots,p\} \times \{0,1,\ldots,p\} \) and \( u_i, v_i \in V_k \) for \( i = 0, 1, \ldots, p \) we have
\[
\sum_{(i,j) \in S} |a(T^i u_i, T^j v_j)| \leq K_1 \left( \sum_{i=0}^{p} a(T^i u_i, u_i) \right)^{\frac{1}{2}} \left( \sum_{j=0}^{p} a(T^j v_j, v_j) \right)^{\frac{1}{2}}. \tag{33}
\]

Notice that all assumptions are related to the choice of the operators \( R^i \). Assumption 6 involves only functions in \( V_k \), without using the decomposition in Assumption 5. We remark that the absolute value in Equation (33) is sufficient but not necessary for convergence, see [19]. Due to Assumption 4, \( R^i \) is invertible so that (32) is well-defined. Also, recall the following property.

**Lemma 6.** Let Assumption 4 hold. The operator \( T^i \) is symmetric and positive semi-definite with respect to \( a(\cdot, \cdot) \).

**Proof.** Let \( u \) and \( v \) be in \( V_k \). Using the fact that \( A_k \) is symmetric with respect to \( (\cdot, \cdot) \) in \( V_k \), \( R^i_k \) is symmetric with respect to \( (\cdot, \cdot) \) in \( V_k^i \), together with the definition of \( Q^i_k \), we have
\[
a(T^i_k u, v) = a(R^i_k Q^i_k A_k u, v) = (R^i_k Q^i_k A_k u, A_k v) \\
= (R^i_k Q^i_k A_k u, Q^i_k A_k v) = (Q^i_k A_k u, R^i_k Q^i_k A_k v) \\
= (Q^i_k A_k u, T^i_k v) = (A_k u, T^i_k v) = a(u, T^i_k v).
\]

This shows \( T^i_k \) is symmetric with respect to \( a(\cdot, \cdot) \). To see that it is also positive semi-definite, we use the same properties as above and get
\[
a(T^i_k u, u) = (R^i_k Q^i_k A_k u, A_k u) = (R^i_k Q^i_k A_k u, Q^i_k A_k u).
\]

Since \( R^i_k \) is SPD with respect to \( (\cdot, \cdot) \) by Assumption 4, the result follows. \( \square \)
By the symmetry of $T^i$ with respect to $a(\cdot, \cdot)$, $I - T^i$ is symmetric with respect to $a(\cdot, \cdot)$. Hence $\hat{E}^*_p = (I - T^p) \cdots (I - T^1)(I - T^0)$ is the adjoint of $\hat{E}_p$ with respect to $a(\cdot, \cdot)$, so that

$$\hat{E}^*_p = \hat{E}^*_p \hat{E}_p.$$  \hfill (34)

Thus, we have

$$a(\hat{E}^*_p v, v) = ||\hat{E}_p v||^2_E \quad \forall v \in V_k.$$  \hfill (35)

We now state the convergence result.

**Theorem 2.** Let Assumptions 4, 5 and 6 hold. Then, we have

$$||\hat{E}_p||^2_E \leq 1 - \frac{2 - w_1}{K_0(1 + K_1)^2},$$  \hfill (36)

where $w_1$ was defined in Assumption 4 and $K_0$ and $K_1$ are constants related to the ones in Assumptions 5 and 6.

**Proof.** See [19] for a proof. Let us point out that the quantity in the right-hand side of (36) has a nonzero denominator, is larger than 0 and less than 1. In fact, notice that Assumption 4 implies that $(R^i)^{-1}$ is also SPD with respect to $(\cdot, \cdot)$, which, together with the fact that $A$ is SPD with respect to $(\cdot, \cdot)$, implies $K_0 > 0$. Then, $2 - w_1 \frac{K_0}{K_0(1 + K_1)^2}$ is well-defined and by Assumption 4 we have $\frac{2 - w_1}{K_0(1 + K_1)^2} > 0$. Finally, the constant $K_0$ can be majorized by another constant such that (32) still holds and $\frac{2 - w_1}{K_0(1 + K_1)^2} < 1$. \hfill $\square$

The convergence of the symmetric version of the SSC algorithm is then a direct consequence of (34).

**Remark 1.** As we dropped the subscript $k$ to make the notation more readable, we point out that the quantities $w_1$, $K_0$ and $K_1$ in general depend on $k$. Moreover, also the quantity $p$ in Algorithms 2 or 3 can be chosen differently for different multigrid levels.

### 3.2. Sufficient conditions for multigrid convergence with smoothers of SSC type

Our intent is to fit the properties of the SSC smoother to the sufficient conditions needed for the convergence of the multigrid algorithm 1. Choosing the symmetric SSC iteration as the smoother for our multigrid algorithm, we then have

$$S_k = \hat{E}^*_p.$$  \hfill (37)

Our purpose is to consider a set of choices of $V_k$ in the multigrid algorithm and of smoothers $S_k = \hat{E}^*_p$ for which Assumptions 1, 2 and 3 are satisfied. First,
the existence of a suitable error operator $\tilde{E}_{sp}$ with norm less than one is given by the fulfillment of Assumptions 4, 5 and 6. Once the existence of this operator is granted, Assumptions 1 and 2 are true.

In fact, Assumption 1 is a consequence of the fact that $\tilde{E}_{sp}$ is SPD with respect to $a(\cdot, \cdot)$. Concerning Assumption 2, we set $\delta_k$ as

$$\delta_k = 1 - \frac{2 - w_1}{K_0(1+K_1)^2}. \quad (38)$$

It then follows by (35) and (36) that Assumption 2 holds.

The only assumption that remains to be checked is Assumption 3, which again depends on all the Assumptions 4, 5 and 6. This is due to the fact that $\delta_k$ in (38) is determined by $w_1$, $K_0$ and $K_1$, all of which in general depend on $k$. Different definitions of $V_k$ and $S_k$ correspond to multigrid algorithms on different spaces with different subspace correction smoothing schemes. Some examples will be described in Section 4. In these, a verification of Assumptions 4, 5, 6 is provided, along with a characterization of the constants $w_1$, $K_0$ and $K_1$ in terms of $k$. This characterization leads to the identification of the conditions for Assumption 3 to hold. The conditions for the optimality of the multigrid error bound are also determined.

4. Refinement applications with subspace correction smoothing schemes

In this section, we apply the multigrid algorithm with SSC-type smoother described earlier, to applications involving uniform and local refinement and also domain decomposition smoothing. Multiplicative domain decomposition algorithms can in fact be seen as instances of SSC algorithms [19]. In the following, we introduce the model problem and its finite element discretization. The applications of the theory that we consider here differ in the definition of $V_k$, in its decomposition into appropriate subspaces $V_i^k$ and in the choice of the operators $R_i^k$. We first address a case of uniform refinement with exact subsolvers. Then, we present two local refinement applications that deal with two possible ways of enforcing continuity, corresponding to appropriate choices of the subspace decomposition of the multigrid spaces $V_k$.

4.1. Model problem and finite element discretization

To fix the ideas, let $\Omega$ be a polygonal subset of $\mathbb{R}^n$, let $\Theta = (\theta_{ij})$ be a symmetric matrix and consider the elliptic boundary value problem

$$-\sum_{i=1}^n \sum_{j=1}^n \frac{\partial}{\partial x_i} \left( \theta_{ij} \frac{\partial u}{\partial x_j} \right) = f \quad \text{in} \quad \Omega$$

$$u = 0 \quad \text{on} \quad \partial \Omega,$$

then $u$ is a weak solution of the above problem if and only if

$$a(u, v) = (f, v) \quad \text{for all} \quad v \in H_0^1(\Omega) \quad (39)$$
where $(\cdot, \cdot)$ denotes the $L^2(\Omega)$ inner product and
\[
a(u, v) = -\sum_{i=1}^{n} \sum_{j=1}^{n} \int_{\Omega} \theta_{ij} \frac{\partial u}{\partial x_j} \frac{\partial v}{\partial x_i} \, dx.
\] (40)

We assume that there exist $C_{\Theta,1}$ and $C_{\Theta,2}$ depending on $\Theta$ for which
\[
C_{\Theta,1} \| u \|_{H^1}^2 \leq a(u, u) \leq C_{\Theta,2} \| u \|_{H^1}^2, \quad u \in V_k \subset H^1_0(\Omega). \tag{41}
\]

This means that both $(\cdot, \cdot)$ and $a(\cdot, \cdot)$ are SPD bilinear forms on $V_k$, as needed in the previous convergence theory. Moreover, since the trace of $V_k$ is zero on the boundary of $\Omega$, we have that $a(\cdot, \cdot)$ is also equivalent to $|\cdot|_{H^1(\Omega)}$ on $V_k$ due to the Poincaré inequality.

Let $P_1$ be the space of linear polynomials, then the multigrid spaces $V_k$ in (1) will be considered to be the finite element spaces of continuous piecewise-linear functions built on triangulations $T_k$ of $\Omega$,
\[
V_k = \{ v \in H^1_0(\Omega) : v|_\tau \in P_1, \ \forall \tau \in T_k \} \quad k = 0, \ldots, J. \tag{42}
\]

Such triangulations will be defined by using either uniform or local midpoint refinement. In the case where such refinement procedure is performed only on a subdomain of $\Omega$ (local refinement), hanging nodes will be introduced in the mesh and the triangulation will be referred to as irregular (or non-conforming). Hanging nodes (also called slave nodes by some authors) are vertices of some element $\tau_1 \in T_k$ that lie on the interior of an edge of some other element $\tau_2 \in T_k$ without being a node for $\tau_2$. A more formal description of hanging nodes can be found in [23, 24, 25, 26]. Continuity constraints can be added in the definition of the finite element spaces to make sure that no additional degrees of freedom are introduced for the hanging nodes. Therefore for all $k = 0, \ldots, J$, $V_k$ has a nodal basis that consists of functions associated to all vertices of $T_k$ excluding the hanging nodes. In practice, a possible way to obtain a continuous nodal basis is given in [25], where shape functions of elements with hanging nodes in the element corners are modified. In [25], the support of a basis function associated to a regular node $n$ is the union of elements that share this node or the potential hanging nodes on the edges that have node $n$. We point out that the local refinement applications covered by our theory allow the presence of edges with an arbitrary number of hanging nodes. Usually, only 1-irregular meshes are considered [27], namely meshes where at most one hanging node is allowed on any edge of the triangulation.

4.2. Uniform refinement: overlapping non-nested subdomains, subproblems on regular grids and exact subsolvers

We first describe a case of uniform refinement by defining the triangulations and the corresponding subdomains on each of them.

**Definition 1 (Triangulations $T_k$).** Let $T_0$ be a quasi-uniform coarse triangulation of $\Omega$ of size $h_0 \in (0, 1]$. Assume $T_{k-1}$ has been obtained, then $T_k$ is derived
from $\mathcal{T}_{k-1}$ by means of midpoint refinement. It follows that the size $h_k$ of $\mathcal{T}_k$ will be $h_k = 2^{-k}h_0$ and that

$$\mathcal{T}_0 \subset \mathcal{T}_1 \subset \cdots \mathcal{T}_k,$$

in the sense that any $\tau \in \mathcal{T}_k$ can be written as the union of elements in $\mathcal{T}_{k+1}$ [28].

**Definition 2** (Subdomains $\hat{\Omega}_k^i$). Let $\{\Omega_k^i\}_{i=1}^{p_k}$ be a collection of non-overlapping open subdomains of $\Omega$ whose boundaries align with the mesh triangulation $\mathcal{T}_k$, such that $\Omega = \bigcup_{i=1}^{p_k} \Omega_k^i$. For $i = 1, \ldots, p_k$, let $\hat{\Omega}_k^i$ be overlapping subsets of $\Omega$ whose boundaries still align with the triangulation and are defined by

$$\hat{\Omega}_k^i = \{x \in \Omega \mid \text{dist}(x, \Omega_k^i) \leq h_0\}.$$  

(43)

Notice that the number of subdomains $p_k$ varies with the level. An example of a subdomain described in the above definition is shown in Figure 1.

For this application, the multigrid spaces $V_k$ in (42), the subspaces $V_k^i$ and the subsolvers $R_k^i$ are defined as follows.

**Definition 3.** Given the triangulations $\mathcal{T}_k$ in Definition 1 and the overlapping subdomains $\hat{\Omega}_k^i$ in Definition 2, we set for $k = 0, \ldots, J$ and for $i = 0, \ldots, p_k$

$$\begin{cases}
V_k \text{ in (42), built on } \mathcal{T}_k \text{ as in Definition 1}, \\
V_k^i := \left\{v \in V_k \mid \text{supp}(v) \subseteq \hat{\Omega}_k^i\right\} & \text{for } i = 0 \\
R_k^i := (A_k^i)^{-1} & \text{for } i = 1, \ldots, p_k
\end{cases}$$

(44)

We point out that the $V_k$ defined in (44) satisfy the nestedness condition (1). The following lemma describes a decomposition of $V_k$. 

17
Lemma 7. Given $V_k$ and $V_k^i$ in Definition 3, we have

$$V_k = \sum_{i=0}^{p_k} V_k^i.$$  

Moreover, if we denote with $v_i \in V_k^i$ the components of any $v \in V_k$ (i.e., such that $v = \sum_{i=0}^{p_k} v_i$), there is a constant $C_0$ independent of $h_0$, $h_k$ and $p_k$ such that

$$\sum_{i=0}^{p_k} a(v_i, v_i) = C_0 a(v, v) \quad \forall v \in V_k. \quad (45)$$

Proof. A proof of this result can be found in [19] and [29]. \qed

The choice of $R_k^i$ implies that $R_k^i A_k^i = I$ for all $i = 0, \ldots, p_k$ and $k = 0, \ldots, J$ and so we have $w_{1,k} = w_{1} = 1$. Assumption 4 is then satisfied. Now we can look at Assumptions 5 and 6 by showing the existence of the parameters $K_0$ and $K_1$ for this application.

Lemma 8. Let $V_k$ be as in Definition 3. Then, there exists a constant $K_0$ satisfying Assumption 5.

Proof. Let $v \in V_k$ and consider the decomposition of $V_k$ provided by Lemma 7. Then we have by (45)

$$\sum_{i=0}^{p_k} ((R_k^i)^{-1} v_i, v_i) = \sum_{i=0}^{p_k} (A_k^i v_i, v_i) = \sum_{i=0}^{p_k} a(v_i, v_i) \leq C_0 a(v, v).$$

This shows that $K_0$ exists and $K_0 = C_0$. \qed

Lemma 9. Let $V_k^i$ and $R_k^i$ as in Definition 3. Then, there exists a constant $K_1$ satisfying Assumption 6.

Proof. Here we follow a variation of a procedure in [30], Section 2.5. Given the subdomains $\tilde{\Omega}_k^1, \ldots, \tilde{\Omega}_k^{p_k}$ as in Definition 2, we define the symmetric $p_k \times p_k$ matrix $G$ by

$$G_{ij} = \begin{cases} 1 & \text{if } \tilde{\Omega}_k^i \cap \tilde{\Omega}_k^j \neq \emptyset, \\ 0 & \text{if } \tilde{\Omega}_k^i \cap \tilde{\Omega}_k^j = \emptyset \end{cases} \quad g_0 = \max_{i=1,\ldots,p_k} \left( \sum_{j=1}^{p_k} G_{ij} \right) = ||G||_\infty \quad (46)$$

Note that $g_0$ represents the maximum number of neighbors intersecting a subdomain (counting self intersections) and it does not depend on $p_k$ but only on the geometry of the triangulation. Unlike [30], the summation in the definition of the constant $g_0$ does not exclude the $i$ term. Also, by the choice of the subdomains, $g_0$ will be uniformly bounded. Let $S \subset \{0, 1, \ldots, p_k\} \times \{0, 1, \ldots, p_k\}$,
and consider the decomposition as in [30], namely

\[ S = S_{00} \cup S_{10} \cup S_{01} \cup S_{11}, \]
\[ S_{00} = \{(i, j) \in S | i = 0, j = 0\}, \]
\[ S_{10} = \{(i, j) \in S | 1 \leq i \leq p_k, j = 0\}, \]
\[ S_{01} = \{(i, j) \in S | i = 0, 1 \leq j \leq p_k\}, \]
\[ S_{11} = \{(i, j) \in S | 1 \leq i, j \leq p_k\}. \]

Let \( u_i, v_i \in V_k \) for \( i = 0, 1, \ldots, k \), then the sum over \( S \) can be split as

\[
\sum_{(i,j) \in S} |a(T^i u_i, T^j v_j)| = \sum_{(i,j) \in S_{00}} |a(T^i u_i, T^j v_j)| + \sum_{i(i,0) \in S_{10}} |a(T^i u_i, T^0 v_0)|
\]
\[ + \sum_{j(0,j) \in S_{01}} |a(T^0 u_0, T^j v_j)| + \sum_{(i,j) \in S_{11}} |a(T^i u_i, T^j v_j)| \quad (47) \]

Let us consider one summand at a time. By the Cauchy-Schwarz inequality and Lemma 6 we have that

\[
\left( \sum_{(i,j) \in S_{00}} |a(T^i u_i, T^j v_j)| \right)^2 \leq \left( \sum_{i=0}^{p_k} a(T^i u_i, u_i) \right) \left( \sum_{j=0}^{p_k} a(T^j v_j, v_j) \right).
\]

If for given \( i \) and \( j \), \( \hat{\Omega}_k^i \cap \hat{\Omega}_k^j = \emptyset \), then \( a(T^i u_i, T^j v_j) = 0 \). Hence for the last summand we have

\[
\left( \sum_{(i,j) \in S_{11}} |a(T^i u_i, T^j v_j)| \right)^2 = \left( \sum_{(i,j) \in S_{11}} G_{ij} |a(T^i u_i, T^j v_j)| \right)^2 \leq \left( \sum_{(i,j) \in S_{11}} G_{ij} \sqrt{a(T^i u_i, T^j v_j) \sqrt{a(T^j v_j, T^j v_j)}} \right)^2
\]
\[ = \left( \sum_{(i,j) \in S_{11}} G_{ij} \sqrt{a(T^i u_i, T^j v_j) \sqrt{a(T^j v_j, v_j)}} \right)^2 \quad \text{(by def of } P_k^i) \]
\[ \leq \left( \sum_{i=1}^{p_k} \sum_{j=1}^{p_k} G_{ij} \sqrt{a(T^i u_i, T^j v_j) v_j} \right)^2 \leq \rho(G)^2 \left( \sum_{i=1}^{p_k} a(T^i u_i, u_i) \right) \left( \sum_{j=1}^{p_k} a(T^j v_j, v_j) \right) \quad \text{(by (4.12) in [31])}
\]
\[ \leq g_0^2 \left( \sum_{i=0}^{p_k} a(T^i u_i, u_i) \right) \left( \sum_{j=0}^{p_k} a(T^j v_j, v_j) \right), \]

19
where \( \rho(G) \) denotes the spectral radius of \( G \) which satisfies \( \rho(G) \leq ||G||_{\infty} \).

Considering that \( a(T^i u_i, T^0 v_0) = 0 \) anytime \( \hat{\Omega}_k \cap \hat{\Omega}_k = \emptyset \), for the second term of the sum in (60) we have

\[
\left( \sum_{i: (i, 0) \in S_{10}} |a(T^i u_i, T^0 v_0)| \right)^2 \leq \left( \sum_{i: (i, 0) \in S_{10}} G_{0i} \sqrt{a(T^i u_i, T^i u_i)} \sqrt{a(T^0 v_0, T^0 v_0)} \right)^2 = \left( \sum_{i: (i, 0) \in S_{10}} G_{0i} \sqrt{a(T^i u_i, T^i u_i)} \right)^2 a(T^0 v_0, T^0 v_0) \\
\leq \left( \sum_{i=1}^{p_k} G_{0i} \right) \left( \sum_{i: (i, 0) \in S_{10}} a(T^i u_i, T^i u_i) \right) a(T^0 v_0, T^0 v_0) \\
\leq g_0 \left( \sum_{i=0}^{p_k} a(T^i u_i, u_i) \right) a(T^0 v_0, v_0) \quad \text{(by def of } P^t_k \text{ and } g_0) \\
\leq g_0 \left( \sum_{i=0}^{p_k} a(T^i u_i, u_i) \right) \left( \sum_{j=0}^{p_k} a(T^j v_j, v_j) \right).
\]

Similarly, for the last term of the sum we have

\[
\left( \sum_{j: (0, j) \in S_{01}} |a(T^0 u_0, T^j v_j)| \right)^2 \leq g_0 \left( \sum_{i=0}^{p_k} a(T^i u_i, u_i) \right) \left( \sum_{j=0}^{p_k} a(T^j v_j, v_j) \right).
\]

Combining these four inequalities, it follows that

\[
\left( \sum_{(i, j) \in S} |a(T^i u_i, T^j v_j)| \right)^2 \leq 4 (1 + 2 g_0 + g_0^2) \left( \sum_{i=0}^{p_k} a(T^i u_i, u_i) \right) \left( \sum_{j=0}^{p_k} a(T^j v_j, v_j) \right).
\]

This shows that \( K_1 \) exists and

\[ K_1 = 2 (1 + g_0). \quad (49) \]

The next result follows immediately from Lemmas 8 and 9. It shows how the assumption about the non-increasing behavior of \( \psi_k \) in (38) is satisfied.

**Lemma 10.** Let \( V^i_k \) and \( R^i_k \) as in Definition 3. Then Assumption 3 is satisfied with

\[
\delta_k = 1 - \frac{1}{C_0(3 + 2 g_0)^2}, \quad \psi_k = \frac{m_k}{C_0(3 + 2 g_0)^2}, \quad (50)
\]

if and only if \( m_k \) is non-increasing. Here, \( C_0 \) is the constant from Lemma 7 and \( g_0 \) is defined in (46).
Notice that the constant $\delta_k$ is independent of $k$. Hence, we have the convergence result.

**Theorem 3.** If $m_k$ is non-increasing, the multigrid algorithm 1 converges with

$$
\gamma_k = \frac{C_0(3 + 2g_0)^2}{C_0(3 + 2g_0)^2 + 2m_k},
$$

where $\gamma_k$ are the constants defined in (19).

Moreover, if $m_1 = m_2 = \ldots = m_J$, the error bound is optimal in the sense that it does not deteriorate as the number of multigrid spaces $J$ increases.

Notice that convergence can be achieved even by performing only one smoothing iteration, but a larger $m_k$ can further lower the error bound.

### 4.3. Local refinement: overlapping nested subdomains, subproblems on regular grids and approximate subsolvers

Now we move to an application involving a locally refined grid.

**Definition 4 (Triangulations $T_k$).** Let $\{\Omega_k\}_{k=0}^J$ be a collection of closed subdomains of $\Omega$ such that $\Omega_J \subset \Omega_{J-1} \subset \cdots \Omega_0 \equiv \Omega$.

Let $T_0$ be a coarse quasi-uniform triangulation of $\Omega$ of size $h_0 \in (0, 1]$. Assume $T_{k-1}$ has been defined, then $T_k$ is obtained performing midpoint refinement only on those elements of $T_{k-1}$ that belong to $\Omega_k$.

This process introduces hanging nodes, causing the grid $T_k$ to become irregular, for all $k = 1, \ldots, J$. However, restricted to $\Omega_k$, $T_k$ is a regular grid without hanging nodes and size $h_k = 2^{-k} h_0$. We observe that the sequence $\{T_k\}_{k=0}^J$ is nested in the sense that an element $T \in T_{k-1}$ can be written as the union of elements in $T_k$ [28]. Moreover, by construction we have that $h_0 = \max_{T \in T_k} h_T$, where $h_T$ denotes the size of one element $T \in T_k$. Figure 2 sketches an example of triangulation for this case. Concerning the spaces $V_k$, the subspaces $V^i_k$ and the corresponding subsolvers $R^i_k$ we choose the following.

**Definition 5.** Given the overlapping subdomains $\Omega_i$ and the triangulations $T_k$ in Definition 4, we set for $k = 0, \ldots, J$ and for $i = 0, \ldots, k$

$$
\begin{align*}
V_k &= \{ v \in H^1_0(\Omega) \cap C^0(\Omega) : v|_\tau \in P_1, \ \forall \tau \in T_k, \}, \\
V^i_k &= \begin{cases} 
V_0, & i = 0, \\
\{ v \in V_i \mid \text{supp}(v) \subseteq \Omega_i \}, & i = 1, \ldots, k,
\end{cases}
\end{align*}
$$

$$
R^i_k := \begin{cases} 
A_0^{-1}, & i = 0, \\
\frac{1}{\lambda^i} I, & i = 1, \ldots, k,
\end{cases}
$$

where $\lambda^i$ denotes the spectral radius of $A^i$. 

We point out that the $V_k$ satisfy by construction the nestedness condition (1). Moreover, the continuity requirement in the definition of $V_k$ implies that its nodal basis will have no function associated to hanging nodes of $T_k$. Also, since the support of the functions in each $V^i_k$ is contained in $\Omega_i$, the subproblems are all defined on uniformly refined grids without hanging nodes, although $T_k$ is irregular. This considerably simplifies the implementation since no actual constraints have to be added and no change in the nodal basis is required to obtain a continuous numerical solution. If $v \in V^i_k$, it will be a linear combination of the basis functions associated with the interior nodes of $\Omega_i$. The following lemma is a consequence of the choice of the subspaces introduced in Definition 5. See also [32] for more on this decomposition.

**Lemma 11.** Given $V_k$ and $V^i_k$ in Definition 5, we have

$$V_k = \sum_{i=0}^{k} V^i_k.$$  

**Proof.** The result follows if for given $v \in V_k$ we can find a decomposition $v = \sum_{i=0}^{k} v_i$ such that $v_i \in V^i_k$. To do this, we will consider a result from [14] that relies on the construction of a sequence of operators $\bar{Q}_i : V_k \rightarrow V_i$. Let $V_i$ be the space obtained by taking $\Omega_0 = \Omega_1 = \cdots = \Omega_i$, namely the space built over a uniformly refined triangulation of size $h_i = h_{0}2^{-i}$ and let $\bar{Q}_i$ be the $L^2(\Omega)$ projection operator onto $V_i$. Set $\bar{Q}_k = I$ and for $i = 0, \ldots, k-1$ define $\bar{Q}_i v = w$ as the unique function on $V_i$ that satisfies

$$w = \begin{cases} 
\bar{Q}_i v & \text{at the nodes of } V_i \text{ in the interior of } \Omega_{i+1}, \\
v & \text{at the remaining nodes of } V_i.
\end{cases}$$
It has been shown in [14] that \((\hat{Q}_i - \hat{Q}_{i-1})v\) is a function in \(V_k^i\) for all \(i = 1, \ldots, k\) and that
\[
(\hat{Q}_i - \hat{Q}_{i-1})v, (\hat{Q}_i - \hat{Q}_{i-1})v) \leq \hat{C}_1 h_0^2 a(v, v) \quad \text{for } i = 1, \ldots, k,
\]
\[
(\hat{Q}_i - \hat{Q}_{i-1})v, (\hat{Q}_i - \hat{Q}_{i-1})v \leq C_1 \lambda_i^{-1} a(v, v) \quad \text{for } i = 1, \ldots, k,
\]
\[
a(\hat{Q}_i v, \hat{Q}_i v) \leq C_2 a(v, v) \quad \text{for } i = 0, \ldots, k - 1,
\]
where \(\lambda_i\) denotes the spectral radius of the operator \(A_i\) and \(\hat{C}_1, C_1\) and \(C_2\) do not depend on \(i\). It then follows that for all \(v \in V_k\)
\[
v = \hat{Q}_0 v + \sum_{i=1}^k (\hat{Q}_i - \hat{Q}_{i-1})v = \sum_{i=0}^k v_i,
\]
where
\[
u_i := \begin{cases} \hat{Q}_0 v, & i = 0, \\ (\hat{Q}_i - \hat{Q}_{i-1})v, & i = 1, \ldots, k, \end{cases}
\]
with \(v_i \in V_k^i\) for all \(i\).

**Remark 2.** In this case \(p_k = k\). This means that at each level \(k\), the number of subdomains is fixed and equal to \(k\) as well. At the given level \(k\), notice that \(V_k^i \subsetneq V_k^j, \forall i > j\), since the trace of \(V_k^i\) on \(\partial \Omega_4\) is zero while the trace of \(V_k^j\) is not. Moreover, it follows from Definition 5 that the \(V_k^i\) are independent of \(k\). Consequently so will be \(A_i\), in the sense that \(A_i^i = A_i^{i+1} = \ldots = A_k^i\).

Note that with the choice of \(R_k^i\) in (52) we have \(\rho(R_k^i A_0) = 1\) and \(\rho(R_k^i A_i) = 1/k\) for all \(i = 1, \ldots, k\). This implies that \(w_{1,k} = w_1 = 1\), so that Assumption 4 is satisfied. Now we can show the existence of the parameters \(K_0\) and \(K_1\).

**Lemma 12.** Let \(V_k^i\) and \(R_k^i\) as in Definition 5. Then, there exists a constant \(K_0\) satisfying Assumption 5.

**Proof.** Using the definition of \(R_k^i\) together with (53) and (55) we have
\[
\sum_{i=0}^k ((R_k^i)^{-1} v_i, v_i) = (A_0 v_0, v_0) + k \sum_{i=1}^k \lambda^i ((\hat{Q}_i - \hat{Q}_{i-1})v, (\hat{Q}_i - \hat{Q}_{i-1})v)
\]
\[
\leq a(v_0, v_0) + k \sum_{i=1}^k C_1 \lambda_i a(v, v)
\]
\[
\leq C_2 a(v, v) + k^2 C_1 a(v, v)
\]
\[
\leq \max\{C_1, C_2\} (1 + k^2) a(v, v) = C_3 (1 + k^2) a(v, v).
\]
This shows that \(K_0\) exists and
\[
K_0 = C_3 (1 + k^2).
\]

\[\square\]
Let us now show the existence of $K_1$ for this application.

**Lemma 13.** Let $V^i_k$ and $R^i_k$ as in Definition 5. Then, there exists a constant $K_1$ satisfying Assumption 6.

**Proof.** For $i = 1, \ldots, k$ and $u \in V_k$ we have

$$0 \leq a(T^i u, T^i u) = a(R^i_k Q^i A_k u, R^i_k Q^i A_k u)$$

$$= \left( \frac{1}{\lambda_i k} \right)^2 a(Q^i A_k u, Q^i A_k u)$$

$$= \left( \frac{1}{\lambda_i k} \right)^2 (Q^i A_k u, A^i Q^i A_k u) \quad \text{(def. of $A^i$)}$$

$$\leq \left( \frac{1}{\lambda_i k} \right)^2 \lambda_i (Q^i A_k u, Q^i A_k u) \quad \text{(def. of $Q^i$)}$$

$$= \left( \frac{1}{\lambda_i k^2} \right) (Q^i A_k u, A_k u) \quad \text{(def. of $Q^i$)}$$

$$= \frac{1}{k} (R^i_k Q^i A_k u, A_k u) \quad \text{(def. of $R^i_k$)}$$

$$= \frac{1}{k} (T^i u, A_k u) \quad \text{(def. of $T^i$)}$$

$$= \frac{1}{k} a(T^i u, u) \quad \text{(def. of $A_k$)}.$$  

For $i = 0$ we have $R^0_k = A_0^{-1}$ so that $T^0 = P^0$ and

$$a(T^0 u, T^0 u) = a(P^0 u, P^0 u) = a(P^0 u, u) = a(T^0 u, u). \quad (57)$$

In summary

$$a(T^i u, T^i u) \begin{cases} = a(T^i u, u), & i = 0, \\ \leq \frac{1}{k} a(T^i u, u), & i = 1, \ldots, k. \end{cases} \quad (58)$$

Let $S \subset \{0, 1, \ldots, k\} \times \{0, 1, \ldots, k\}$, and consider the decomposition of such set as before with $p_k = k$, namely

$$S = S_{00} \cup S_{10} \cup S_{01} \cup S_{11},$$

$$S_{00} = \{(i, j) \in S \mid i = 0, j = 0\},$$

$$S_{10} = \{(i, j) \in S \mid 1 \leq i \leq k, j = 0\},$$

$$S_{01} = \{(i, j) \in S \mid i = 0, 1 \leq j \leq k\},$$

$$S_{11} = \{(i, j) \in S \mid 1 \leq i, j \leq k\}. \quad (59)$$

Let $u_i, v_i \in V_k$ for $i = 0, 1, \ldots, k$, then

$$\sum_{(i, j) \in S} |a(T^i u_i, T^j v_j)| = \sum_{(i, j) \in S_{00}} |a(T^i u_i, T^j v_j)| + \sum_{i : (i, 0) \in S_{10}} |a(T^i u_i, T^0 v_0)|$$

$$+ \sum_{j : (0, j) \in S_{01}} |a(T^0 u_0, T^j v_j)| + \sum_{(i, j) \in S_{11}} |a(T^i u_i, T^j v_j)|. \quad (60)$$

24
Let us consider one summand at a time. By the Cauchy-Schwarz inequality with the $a(\cdot, \cdot)$ inner product, (58) and Lemma 6 we have that
\[
\left( \sum_{(i,j) \in S_{10}} |a(T^i u_i, T^j v_j)| \right)^2 \leq a(T^0 u_0, u_0) a(T^0 v_0, v_0)
\]
\[
\leq \left( \sum_{i=0}^{k} a(T^i u_i, u_i) \right) \left( \sum_{j=0}^{k} a(T^j v_j, v_j) \right).
\]
For the last summand we have, using again the same properties,
\[
\left( \sum_{(i,j) \in S_{11}} |a(T^i u_i, T^j v_j)| \right)^2 \leq \left( \sum_{(i,j) \in S_{11}} \sqrt{a(T^i u_i, T^i u_i)} \sqrt{a(T^j v_j, T^j v_j)} \right)^2
\]
\[
\leq \frac{1}{k^2} \left( \sum_{i:(i,j) \in S_{11}} \sqrt{a(T^i u_i, u_i)} \right)^2 \left( \sum_{j:(i,j) \in S_{11}} \sqrt{a(T^j v_j, v_j)} \right)^2
\]
\[
\leq \frac{1}{k^2} \left( \sum_{i=1}^{k} a(T^i u_i, u_i) \right) \left( \sum_{j=1}^{k} a(T^j v_j, v_j) \right)
\]
\[
\leq \left( \sum_{i=0}^{k} a(T^i u_i, u_i) \right) \left( \sum_{j=0}^{k} a(T^j v_j, v_j) \right).
\]
For the second term of the sum in (60) use the Cauchy-Schwarz inequality with the $a(\cdot, \cdot)$ inner product, and (58),
\[
\left( \sum_{i:(i,0) \in S_{10}} |a(T^i u_i, T^0 v_0)| \right)^2 \leq \left( \sum_{i:(i,0) \in S_{10}} \sqrt{a(T^i u_i, T^i u_i)} \sqrt{a(T^0 v_0, T^0 v_0)} \right)^2
\]
\[
= \left( \sum_{i:(i,0) \in S_{10}} \sqrt{a(T^i u_i, T^i u_i)} \right)^2 a(T^0 v_0, T^0 v_0)
\]
\[
\leq \left( \sum_{i:(i,0) \in S_{10}} 1 \right) \left( \sum_{i:(i,0) \in S_{10}} a(T^i u_i, T^i u_i) \right) a(T^0 v_0, v_0)
\]
\[
\leq k \frac{1}{k} \left( \sum_{i:(i,0) \in S_{10}} a(T^i u_i, u_i) \right) a(T^0 v_0, v_0)
\]
\[
\leq \left( \sum_{i=0}^{k} a(T^i u_i, u_i) \right) \left( \sum_{j=0}^{k} a(T^j v_j, v_j) \right)
\]
Similarly, for the last term of the sum we have
\[
\left( \sum_{j:(0,j) \in S_{01}} |a(T^0 u_0, T^j v_j)| \right)^2 \leq \left( \sum_{i=0}^{k} a(T^i u_i, u_i) \right) \left( \sum_{j=0}^{k} a(T^j v_j, v_j) \right).
\]
Combining these four inequalities, it follows that

$$\left( \sum_{(i,j) \in S} |a(T^i u_i, T^j v_j)| \right)^2 \leq 4 \left( \sum_{i=0}^{k} a(T^i u_i, u_i) \right) \left( \sum_{j=0}^{k} a(T^j v_j, v_j) \right). \quad (61)$$

This shows that $K_1$ exists and

$$K_1 = 2. \quad (62)$$

**Lemma 14.** Let $V^k_i$ and $R^k_i$ as in Definition 5. Then, Assumption 3 is satisfied with

$$\delta_k = 1 - \frac{1}{C_4(1 + k^2)}, \quad \psi_k = \frac{m_k}{C_4(1 + k^2)}, \quad (63)$$

if and only if $m_k$ is chosen so that $\psi_k$ is non-increasing. Here, $C_4 = 9 C_3$ and $C_3$ is the constant in (56).

Note that $\delta_k$ is now increasing. Various choices of $m_k$ guarantee Assumption 3: constant $m_k = 1$, decreasing $m_k = J + 1 - k$, increasing $m_k = 1 + k$. We now state the convergence result.

**Theorem 4.** If $m_k$ is chosen so that $\psi_k$ is non-increasing, the multigrid algorithm 1 converges with

$$\gamma_k = \frac{C_4(1 + k^2)}{C_4(1 + k^2) + 2 m_k}, \quad (64)$$

where $\gamma_k$ is defined in Theorem 1, $k = 0, 1, \ldots, J$.

Moreover, the error bound is optimal (in the sense that it does not depend on the number of multigrid spaces $J$) if and only if $m_k = q(1 + k^2)$ for some $q \in \mathbb{N}$, and is given by

$$\gamma_1 = \gamma_2 = \cdots = \gamma_J = \frac{C_4}{2q + C_4}. \quad (65)$$

We observe that the number of smoothing iterations appears in the error bound and this was not shown in [16]. Although the choice $m_k = q(1 + k^2)$ is not optimal in terms of computational cost, since more smoothing steps are needed on finer grids, nevertheless it guarantees that the error bound is independent of the number of levels.

### 4.4. Local refinement: overlapping non-nested subdomains, subproblems on irregular grids and exact subsolvers

We now describe another local refinement application. We keep the same triangulations as in Definition 4 and the same definition of $V_k$ as in the previous local refinement application. However, the subdomains are chosen as in Definition 2. This will lead to a different characterization of the space $V_k$. 26
A sketch of the subdomains involved in this application is visible in Figure 3. Moreover, unlike Section 4.3, the overlapping subdomains $\hat{\Omega}_k^i$ at each level $k$ are not nested. Let us now define the spaces $V_k$ and choose the subspaces for its decomposition, and the subsolvers $R_k^i$.

Definition 6. Given the triangulations $T_k$ in Definition 4 and the subdomains $\Omega_i$ in Definition 2, we set for $k = 0, \ldots, J$ and for $i = 0, \ldots, p_k$

\[
\begin{align*}
V_k &= \{ v \in H^1_0(\Omega) \cap C^0(\Omega) : v|_\tau \in P_1, \ \forall \tau \in T_k, \} \\
where T_k & \text{ is as in Definition 4,} \\
V_k^i &= \begin{cases} 
V_0, & i = 0, \\
\{ v \in V_k | \text{supp}(v) \subseteq \hat{\Omega}_k^i \}, & i = 1, \ldots, p_k,
\end{cases} \\
R_k^i &= (A_k^i)^{-1}.
\end{align*}
\]  

(66)

Since the definition of $V_k$ in Definition 6 coincides with Definition 5, the $V_k$ again satisfy the nestedness condition (1) and the nodal basis does not have any function associated to the hanging nodes, because of the continuity requirement. Here, we give another characterization of $V_k$ based on the non-nested subdomains.

Lemma 15. Given $V_k$ and $V_k^i$ in Definition 6, we have

\[
V_k = \sum_{i=0}^{p_k} V_k^i.
\]

Moreover, if we denote with $v_i \in V_k^i$ the components of any $v \in V_k$ (such that

27
Figure 4: An example of construction of overlapping subdomains obtained from the non-overlapping subdomains in Figure 3.

\[ v = \sum_{i=0}^{p_k} v_i, \] then there is a constant \( C_k \) dependent only on \( k \) such that

\[ \sum_{i=0}^{p_k} a(v_i, v_i) \leq C_k a(v, v) \quad \forall v \in V_k. \]

**Proof.** We are going to construct a set of functions \( \{v_i\}_{i=0}^{p_k} \in V_k^i \subseteq V_k \) such that every \( v \in V_k \) can be expressed as their sum. To this end, let \( \{\theta_i^k\}_{i=1}^{p_k} \) be a smooth partition of unity subordinate to the cover \( \{\Omega_i^k\}_{i=1}^{p_k} \). This means that \( \sum_{i=1}^{p_k} \theta_i^k = 1, \ 0 \leq \theta_i^k(x) \leq 1 \) for all \( x \in \Omega_i^k \) and \( \text{supp}(\theta_i^k) \subset \Omega_i^k \), for all \( i = 1, \ldots, p_k \). Let \( \hat{V}^j \) be the subspace of \( V_k \) defined in Definition 5 in the previous local refinement application. Then we know that \( V_k = \sum_{j=0}^{k} \hat{V}^j \), so that any \( v \) in \( V_k \) can be written as \( v = \sum_{j=0}^{k} \hat{v}_j \), where \( \hat{v}_j \in \hat{V}^j \) are given by (55). Define \( \mathcal{I}_h^j \) to be the standard nodal interpolant of the finite element space \( \hat{V}^j \) for all \( j = 1, \ldots, k \). Note that this is well defined since each \( \hat{V}^j \) is built on a quasi-uniform grid. Then, for \( v \in V_k \), set

\[ v_0 = \hat{v}_0, \quad v_i = \sum_{j=1}^{k} \mathcal{I}_h^j(\theta_i^k \hat{v}_j), \quad i = 1, \ldots, p_k. \]

Notice that all the terms in the sum that defines \( v_i \) are functions in \( \hat{V}^j \subset V_k \) and have support in \( \Omega_i^k \), therefore they all belong to \( V_k^i \). Moreover, using the
fact that the $\mathcal{I}_h^j$ are linear and projections we have

\[
v = \sum_{j=0}^{k} \hat{v}_j = \hat{v}_0 + \sum_{j=1}^{k} \hat{v}_j = \sum_{j=1}^{k} \mathcal{I}_h^j(\hat{v}_j) = \hat{v}_0 + \sum_{j=1}^{k} \mathcal{I}_h^j(\sum_{i=1}^{p_k} \theta_i^k \hat{v}_j)
\]

\[= \hat{v}_0 + \sum_{j=1}^{k} \sum_{i=1}^{p_k} \mathcal{I}_h^j(\theta_i^k \hat{v}_j) = \sum_{i=0}^{p_k} v_i.
\]

To prove the second part of the lemma, let us proceed one summand at a time. If $T \in \mathcal{T}_k \cap \hat{\Omega}_i^k$, then using an inverse estimate (see [21]) we get

\[
|\mathcal{I}_h^j(\theta_i^k \hat{v}_j)|^2_{H^1(T)} \leq h_0^{-2} |\mathcal{I}_h^j(\theta_i^k \hat{v}_j)|^2_{L^2(T)} \\
\leq h_0^{-2} C |\theta_i^k \hat{v}_j|^2_{L^2(T)} \\
\leq h_0^{-2} C |\hat{v}_j|^2_{L^2(T)},
\]

where the constant $C$ is the bound for the operator norm of $\mathcal{I}_h^j$ and it only depends on the reference element [21]. Summing over all $T \in \mathcal{T}_k \cap \hat{\Omega}_i^k$ (remember that we assumed the subdomains align with the triangulation) we obtain

\[
|v_i|^2_{H^1(\Omega)} = |v_i|^2_{H^1(\hat{\Omega}_i^k)} \leq \sum_{T \in \mathcal{T}_k \cap \hat{\Omega}_i^k} \left( \sum_{j=1}^{k} |\mathcal{I}_h^j(\theta_i^k \hat{v}_j)|^2_{H^1(T)} \right)^2 \\
\leq \sum_{T \in \mathcal{T}_k \cap \hat{\Omega}_i^k} k \sum_{j=1}^{k} |\mathcal{I}_h^j(\theta_i^k \hat{v}_j)|^2_{H^1(T)} \\
\leq k \sum_{j=1}^{k} h_0^{-2} C |\hat{v}_j|^2_{L^2(\hat{\Omega}_i^k)}.
\]

Summing over the subdomains $\hat{\Omega}_i^k$, and considering that each point in $\Omega$ is covered only a finite number of times [29] we obtain

\[
\sum_{i=1}^{p_k} |v_i|^2_{H^1(\Omega)} \leq k \sum_{j=1}^{k} h_0^{-2} C |\hat{v}_j|^2_{L^2(\hat{\Omega}_i^k)}.
\]

(68)

Thanks to (53) we can say that

\[
||\hat{v}_j||^2_{L^2(\Omega)} \leq \widetilde{C}_1 h_0^2 |v|^2_{H^1(\Omega)}.
\]

(69)

Therefore, using the previous results and the Poincaré inequality we have

\[
\sum_{i=1}^{p_k} a(v_i, v_i) \leq \overline{C} \sum_{i=1}^{p_k} |v_i|^2_{H^1(\Omega)} \leq k^2 \overline{C} |v|^2_{H^1(\Omega)} \leq k^2 \overline{C} a(v, v).
\]
Again by (53) we know that $a(\vec{v}_0, \vec{v}_0) \leq C_2 a(v, v)$, hence if we let $\hat{C}_0 = \max\{C_2, \tilde{C}\}$ we can conclude with
\[
\sum_{i=0}^{p_k} a(v_i, v_i) \leq \hat{C}_0 (1 + k^2) a(v, v).
\]

The proof of this lemma for a uniform refinement case relies on the uniform boundedness of the standard nodal interpolator on $V_k$. In the case where an irregular grid is employed a nodal interpolator in the classical sense cannot be defined on $V_k$. An alternative to the solution we adopted in our proof could be to use interpolation operators specifically designed for irregular grids as in [33].

For the subsolvers we clearly have that $R^*_k A^i = I$ for all $i = 0, \ldots, p_k$ and so again we have $w_{1,k} = w_1 = 1$. Assumption 4 is then true. However, from a practical point of view, defining problems on irregular grids actually requires the implementation of the constraints that make the nodal basis of $V_k$ continuous, as in [25]. Now we can show the existence of $K_0$ and $K_1$.

**Lemma 16.** Let $V_k^i$ and $R_k^i$ as in Definition 6. Then, there exists a constant $K_0$ satisfying Assumption 5.

**Proof.** Considering the decomposition of $v$ given by Lemma 15, we have
\[
\sum_{i=0}^{p_k} ((R_k^i)^{-1} v_i, v_i) = \sum_{i=0}^{p_k} (A_k^i v_i, v_i) = \sum_{i=0}^{p_k} a(v_i, v_i) \leq \hat{C}_0 (1 + k^2) a(v, v).
\]
This shows that $K_0$ exists and
\[
K_0 = \hat{C}_0 (1 + k^2).
\]

**Lemma 17.** Let $V_k^i$ and $R_k^i$ as in Definition 6. Then, there exists a constant $K_1$ satisfying Assumption 6.

**Proof.** The existence of $K_1$ can be carried out exactly as for Lemma 9 concerning the case of uniform refinement. Therefore $K_1$ exists and
\[
K_1 = 2 (1 + g_0).
\]

The next lemma immediately follows.

**Lemma 18.** Let $V_k^i$ and $R_k^i$ as in Definition 6. Then, Assumption 3 is satisfied with
\[
\delta_k = 1 - \frac{1}{C_0(1 + k^2)(3 + 2 g_0)^2}, \quad \psi_k = \frac{m_k}{C_0(1 + k^2)(3 + 2 g_0)^2},
\]
if and only if $m_k$ is chosen so that $\psi_k$ is non-increasing. Here, $\hat{C}_0$ is the constant in (70) and $g_0$ is defined in (46).
The constant $\delta_k$ is again increasing. Consequently, the convergence bound for the multigrid algorithm is obtained.

**Theorem 5.** If $m_k$ is chosen so that $\psi_k$ is non-increasing, the multigrid algorithm 1 converges with

$$\gamma_k = \frac{C_5(1 + k^2)}{C_5(1 + k^2) + 2m_k},$$

where $\gamma_k$ is defined in Theorem 1 and $C_5 = \tilde{C}_0(3 + 2g_0)^2$.

Moreover, the error bound is optimal (in the sense that it does not depend on the number of multigrid spaces $J$) if and only if $m_k = q(1 + k^2)$ for some $q \in \mathbb{N}$, and is given by

$$\gamma_1 = \gamma_2 = \cdots = \gamma_J = \frac{C_5}{2q + C_5}. \quad (74)$$

---

**5. Conclusions**

In this paper, we performed a convergence analysis of a multigrid algorithm for symmetric elliptic PDEs under no regularity assumptions with smoothers of SSC type. In particular, we focused on the dependence of the multigrid error bound on the number of smoothing steps. This represents a novel result for the case of no-regularity assumptions. We provided an analysis that can be used for any smoothing procedure of symmetric SSC type. We then utilized this framework to address uniform and local refinement applications and study convergence bounds for the multigrid error. Our theory allows an arbitrary number of hanging nodes on a given edge of the triangulation. A judicious choice of the subdomain solvers and of the number of smoothing steps at each level can avoid the dependence of the multigrid error bound on the total number of multigrid levels. To this end, proper decompositions of the finite element spaces had to be derived in the analysis. For the uniform refinement case, a uniform bound for the multigrid error can be obtained, even regardless of the choice of the smoothing steps. For the local refinement applications, we described two different subspace decompositions of the multigrid space using overlapping nested or non-nested subdomains that correspond to different ways of enforcing the continuity of the finite element space when hanging nodes are present. In both cases, we show that convergence can be obtained and optimality can be guaranteed by appropriately choosing the number of smoothing steps for each refinement level. A computational analysis of the methods proposed in this paper will be subject to future investigation.

---

**6. Acknowledgments**

This work was supported by the National Science Foundation grant DMS-1412796.
References

[1] R. P. Fedorenko, A relaxation method for solving elliptic difference equations, USSR Computational Mathematics and Mathematical Physics 1 (4) (1962) 1092–1096.

[2] N. S. Bakhvalov, On the convergence of a relaxation method with natural constraints on the elliptic operator, USSR Computational Mathematics and Mathematical Physics 6 (5) (1966) 101–135.

[3] R. Nicolaides, On multiple grid and related techniques for solving discrete elliptic systems, Journal of Computational Physics 19 (4) (1975) 418–431.

[4] A. Brandt, Multi-level adaptive solutions to boundary-value problems, Mathematics of computation 31 (138) (1977) 333–390.

[5] H. Yserentant, Old and new convergence proofs for multigrid methods, Acta Numerica 2 (1993) 285–326.

[6] R. Nicolaides, On the $l^2$ convergence of an algorithm for solving finite element equations, Mathematics of Computation 31 (140) (1977) 892–906.

[7] R. Nicolaides, On some theoretical and practical aspects of multigrid methods, Mathematics of Computation 33 (147) (1979) 933–952.

[8] J. H. Bramble, D. Y. Kwak, J. E. Pasciak, Uniform convergence of multigrid V-cycle iterations for indefinite and nonsymmetric problems, SIAM Journal on Numerical Analysis 31 (6) (1994) 1746–1763.

[9] J. Wang, Convergence analysis of multigrid algorithms for nonselfadjoint and indefinite elliptic problems, SIAM journal on numerical analysis 30 (1) (1993) 275–285.

[10] M. A. Olshanskii, A. Reusken, Convergence analysis of a multigrid method for a convection-dominated model problem, SIAM journal on numerical analysis 42 (3) (2004) 1261–1291.

[11] W. Hackbusch, Multi-grid methods and applications, Vol. 4, Springer Science & Business Media, 2013.

[12] D. Braess, Finite elements: Theory, fast solvers, and applications in solid mechanics, Cambridge University Press, 2007.

[13] J. H. Bramble, J. E. Pasciak, New convergence estimates for multigrid algorithms, Mathematics of computation 49 (180) (1987) 311–329.

[14] J. H. Bramble, J. E. Pasciak, J. P. Wang, J. Xu, Convergence estimates for multigrid algorithms without regularity assumptions, Mathematics of Computation 57 (195) (1991) 23–46.
[15] J. H. Bramble, J. E. Pasciak, New estimates for multilevel algorithms including the V-cycle, Mathematics of computation 60 (202) (1993) 447–471.

[16] J. H. Bramble, J. E. Pasciak, The analysis of smoothers for multigrid algorithms, Mathematics of Computation 58 (198) (1992) 467–488.

[17] S. Brenner, Convergence of the multigrid V-cycle algorithm for second-order boundary value problems without full elliptic regularity, Mathematics of Computation 71 (238) (2002) 507–525.

[18] L. Chen, R. H. Nochetto, J. Xu, Optimal multilevel methods for graded bisection grids, Numerische Mathematik 120 (1) (2012) 1–34.

[19] J. Xu, Iterative methods by space decomposition and subspace correction, SIAM review 34 (4) (1992) 581–613.

[20] B. Janssen, G. Kanschat, Adaptive multilevel methods with local smoothing for $H^1$- and $H^{curl}$-conforming high order finite element methods, SIAM Journal on Scientific Computing 33 (4) (2011) 2095–2114.

[21] S. C. Brenner, L. R. Scott, The Mathematical Theory of Finite Element Methods, Springer, 2008.

[22] R. A. Horn, C. R. Johnson, Matrix analysis, Cambridge university press, 2012.

[23] C. Carstensen, J. Hu, Hanging nodes in the unifying theory of a posteriori finite element error control, Journal of Computational Mathematics (2009) 215–236.

[24] V. Heuveline, F. Schieweck, On the inf-sup condition for higher order mixed FEM on meshes with hanging nodes, ESAIM: Mathematical Modelling and Numerical Analysis 41 (1) (2007) 1–20.

[25] T.-P. Fries, A. Byfut, A. Alizada, K. W. Cheng, A. Schröder, Hanging nodes and XFEM, International Journal for Numerical Methods in Engineering 86 (4-5) (2011) 404–430.

[26] P. Di Stolfo, A. Schröder, N. Zander, S. Kollmannsberger, An easy treatment of hanging nodes in hp-finite elements, Finite Elements in Analysis and Design 121 (2016) 101–117.

[27] X. Zhao, S. Mao, Z.-C. Shi, Adaptive quadrilateral and hexahedral finite element methods with hanging nodes and convergence analysis, Journal of Computational Mathematics (2010) 621–644.

[28] P. G. Ciarlet, Finite Element Method for Elliptic Problems, Society for Industrial and Applied Mathematics, Philadelphia, PA, USA, 2002.
[29] M. Dryja, O. B. Widlund, Iterative Methods for Large Linear Systems, Academic Press Professional, Inc., San Diego, CA, USA, 1990, Ch. Some Domain Decomposition Algorithms for Elliptic Problems, pp. 273–291.

[30] T. Mathew, Domain decomposition methods for the numerical solution of partial differential equations, Vol. 61, Springer Science & Business Media, 2008.

[31] M. A. Olshanskii, E. E. Tyrtyshnikov, Iterative methods for linear systems: theory and applications, SIAM, 2014.

[32] M. Dryja, O. B. Widlund, On the optimality of an additive iterative refinement method, in: Proceedings of the Fourth Copper Mountain Conference on Multigrid Methods, J. Mandel, SF McCormick, JE Dendy, C. Farhat, G. Lansdale, SV Porter, JW Ruge, and K. Stüben, eds., SIAM, Philadelphia, 1989, pp. 161–170.

[33] V. Heuveline, F. Schieweck, An Interpolation Operator for H^1 Functions on General Quadrilateral and Hexahedral Meshes with Hanging Nodes, Universität Heidelberg. Interdisziplinäres Zentrum für Wissenschaftliches Rechnen [IWR], 2004.