Optimal Multilevel Methods for $H(\mathrm{grad})$, $H(\mathrm{curl})$, and $H(\mathrm{div})$ Systems on Graded and Unstructured Grids

Jinchao Xu, Long Chen, and Ricardo H. Nochetto

Abstract We give an overview of multilevel methods, such as V-cycle multigrid and BPX preconditioner, for solving various partial differential equations (including $H(\mathrm{grad})$, $H(\mathrm{curl})$ and $H(\mathrm{div})$ systems) on quasi-uniform meshes and extend them to graded meshes and completely unstructured grids. We first discuss the classical multigrid theory on the basis of the method of subspace correction of Xu and a key identity of Xu and Zikatanov. We next extend the classical multilevel methods in $H(\mathrm{grad})$ to graded bisection grids upon employing the decomposition of bisection grids of Chen, Nochetto, and Xu. We finally discuss a class of multilevel preconditioners developed by Hiptmair and Xu for problems discretized on unstructured grids and extend them to $H(\mathrm{curl})$ and $H(\mathrm{div})$ systems over graded bisection grids.

1 Introduction

How to effectively solve the large scale algebraic systems arising from the discretization of partial differential equations is a fundamental problem in scientific and engineering computing. In this paper, we give an overview of a special class of methods for solving such systems: multilevel iterative methods based on the method of subspace corrections [18, 91] and the method of auxiliary spaces [92, 52].

Jinchao Xu
Department of Mathematics, Pennsylvania State University, University Park, PA 16802 and LMAM, The School of Mathematical Sciences, Peking University. e-mail: xu@math.psu.edu.

Long Chen
Department of Mathematics, University of California at Irvine, Irvine, CA 92697.
e-mail: chenlong@math.uci.edu.

Ricardo H. Nochetto
Department of Mathematics and Institute for Physical Science and Technology, University of Maryland, College Park, MD 20742. e-mail: rhn@math.umd.edu.

The method of subspace corrections proves to be a very useful general framework for the design and analysis of various iterative methods. We give a rather detailed description of this method in Section §2 and apply it to additive and multiplicative multilevel methods. Of special interest is the sharp convergence identity of Xu and Zikatanov [94], which we also prove.

Most of the multilevel methods are dictated by the underlying mesh structure. In this paper, roughly speaking, we consider the following three types of grids:

- Quasi-uniform (and structured) grids with a hierarchy of nested sub-grids.
- Graded grids obtained by bisection with a hierarchy of nested sub-grids.
- Unstructured grids without a hierarchy of sub-grids.

Multilevel Methods on Quasi-Uniform Grids

The theoretical and algorithmic development of most traditional multilevel methods are devoted to quasi-uniform structured grids; see Brandt [21], Hackbusch [44], Xu [91, 16], and Yserentant [96]. In Section §3, using the method of subspace correction framework [18, 91], we discuss the classical V-cycle multigrid method and the BPX preconditioner. We also include a recent result by Xu and Zhu [93] that demonstrates that the conjugate gradient method with classical V-cycle multigrid or BPX-preconditioner as preconditioners provides a robust method with respect to jump discontinuities of coefficients.

Multilevel Methods on Graded Bisection Grids

Multilevel algorithms for graded grids generated by adaptive finite element methods (AFEM) is one main topic to be discussed in this paper. AFEM are now widely used in scientific and engineering computation to optimize the relation between accuracy and computational labor (degrees of freedom). We refer to the survey to [63] for an introduction to the theory of AFEM.

Of all possible refinement strategies, we are interested in *bisection*, the most popular and effective procedure for refinement in any dimension; see [63] and the references therein. Our goal is to design optimal multilevel solvers and analyze them within the framework of highly graded meshes created by bisection, from now on called *bisection meshes*.

In Section §4, we present multilevel methods and analysis for $H(\mathrm{grad})$ based on the novel decomposition of bisection grids of Chen, Nochetto, and Xu [27], which is conceptually simple and dimension and polynomial degree independent. Roughly speaking, for any triangulation $\mathscr{T}_N$ constructed from $\mathscr{T}_0$ by bisection, we can write

$$\mathscr{T}_N = \mathscr{T}_0 + \mathscr{B}, \quad \mathscr{B} = (b_1, b_2, \cdots, b_N),$$

where $\mathscr{B}$ denotes a sequence of N elementary bisections b_i. Each such b_i is restricted to a local region and the corresponding local grid is quasi-uniform. This decomposition serves as a general bridge to transfer results from quasi-uniform grids to graded bisection grids. We exploit this flexibility to design and analyze local multigrid methods for the $H(\text{curl})$ and $H(\text{div})$ systems in three dimensions in Section §5; we explicitly follow Chen, Nochetto, and Xu [28], which in turn build on Hiptmair and Xu [52].

Multilevel Methods on Unstructured Grids

In practical applications, finite element grids are often unstructured, namely, they have no natural geometric hierarchy that can be extracted from the mesh data structure and used for designing optimal multilevel algorithms. For such problems we turn to algebraic multigrid methods (AMG). What makes AMG attractive in practice is that they generate coarse-level equations without using any (or much) geometric information or re-discretization on the coarse levels. Despite the lack of rigorous theoretical justification, AMG methods are very successful in practice for various Poisson-like equations; see [73, 81] and reference therein.

Even though we do not describe AMG in any detail, in Section §6 we present a technique developed by Hiptmair and Xu [52] for quasi-uniform meshes that converts the solution of both $H(\text{curl})$ and $H(\text{div})$ systems into that of a number of Poisson-like equations, which can be efficiently solved by AMG.

2 The Method of Subspace Corrections

Most partial differential equations, after discretization, are reduced to solve some linear algebraic equations in the form

$$Au = f, \tag{1}$$

where $A \in \mathbb{R}^{N\times N}$ is a sparse matrix and $f \in \mathbb{R}^N$. How to solve (1) efficiently remains a basic question in numerical PDEs (and in all scientific computing). The Gaussian elimination still remains the most commonly used method in practice. It is a black-box as it can be applied to any problem in principle. But it is expensive: for a general $N \times N$ matrix, it required $\mathscr{O}(N^3)$ operations. For a sparse matrix, it may require less operations but still too expensive for large scale problems. Multigrid methods, on the other hand, are examples of problem-oriented algorithms, which, for some problems, only require $\mathscr{O}(N|\log N|^{\sigma}), \sigma > 0$, operations. In this section, we will give some general and basic results that will be used in later sections to construct efficient iterative methods (such as multigrid methods) for discretized partial differential equations.

Following [91], we shall use notation $x \lesssim y$ to stand for $x \le Cy$. We also use $x \eqsim y$ to mean $x \lesssim y$ and $y \lesssim x$.

2.1 Iterative Methods

2.1.1 Basic Iterative Method

In general, a basic linear iterative method for $Au = f$ can be written in the following form:

$$u^{k+1} = u^k + B(f - Au^k),$$

starting from an initial guess $u^0 \in \mathscr{V}$. It can be interpreted as a result of the following three steps:

1. form the residual $r = f - Au^k$;
2. solve the residual equation $Ae = r$ approximately by $\hat{e} = Br$ with $B \approx A^{-1}$;
3. correct the solution $u^{k+1} = u^k + \hat{e}$.

Here B is called *iterator*. As simple examples, if $A = (a_{ij}) \in \mathbb{R}^{N\times N}$ and $A = D+L+U$, we may take $B = D^{-1}$ to obtain the Jacobi method and $B = (D+L)^{-1}$ to obtain the Gauss-Seidel method.

The art of constructing *efficient* iterative methods lies on the design of B which captures the essential information of A^{-1} and its action is easily computable. In this context the notion of "efficient" implies two essential requirements:

1. One iteration requires a computational effort proportional to the number of unknowns.
2. The rate of convergence is well below 1 and independent with the number of unknowns.

2.1.2 Preconditioned Krylov Space Methods

The approximate inverse B, when it is SPD, can be used as a preconditioner for Conjugate Gradient (CG) method. The resulting method, known as preconditioned conjugate gradient method (PCG), admits the following error estimate:

$$\frac{\|u-u^k\|_A}{\|u-u^0\|_A} \le 2\left(\frac{\sqrt{\kappa(BA)}-1}{\sqrt{\kappa(BA)}+1}\right)^k (k\ge 1), \quad \left(\kappa(BA) = \frac{\lambda_{\max}(BA)}{\lambda_{\min}(BA)}\right).$$

Here B is called *preconditioner*. A good preconditioner should have the properties that the action of B is easy to compute and that $\kappa(BA)$ is significantly smaller than $\kappa(A)$.

An interesting fact is that the linear iterative method using iterator B may not be convergent at all whereas B can always be a preconditioner. For example, the Jacobi method is not convergent for all SPD systems, but $B = D^{-1}$ can always be used as a preconditioner which is often known as the diagonal preconditioner.

2.1.3 Convergence Analysis

Let $e^k = u - u^k$. The error equation of the basic iterative method is

$$e^{k+1} = (I - BA)e^k = (I - BA)^k e^0.$$

Thus the basic iterative method converges if and only if the spectral radius of the error operator $I - BA$ is less than one, i.e., $\rho(I - BA) < 1$.

Given an iterator B, we define the iteration operator $\Phi_B u = u + B(f - Au)$ and introduce a symmetric scheme $\Phi_{\overline{B}} = \Phi_{B^t}\Phi_B$. The convergence of the iteration scheme Φ_B and its symmetrization $\Phi_{\overline{B}}$ is connected by the following inequality:

$$\rho(I - BA) \le \sqrt{\rho(I - \overline{B}A)},$$

and the equality holds if $B = B^t$. Hence we shall focus on the analysis of the symmetric scheme.

By definition, we have the following formula for the error operator $I - \overline{B}A$

$$I - \overline{B}A = (I - B^t A)(I - BA), \text{ and thus } \overline{B} = B^t(B^{-t} + B^{-1} - A)B. \tag{2}$$

Since $\overline{B}$ is symmetric, $I - \overline{B}A$ is symmetric with respect to the inner product $(u,v)_A := (Au, v)$. Indeed, let $(\cdot)^*$ be the adjoint operator with respect to $(\cdot,\cdot)_A$, it is easy to show

$$I - \overline{B}A = (I - BA)^*(I - BA). \tag{3}$$

Consequently, $I - \overline{B}A$ is SPD with respect to $(\cdot,\cdot)_A$ and $\lambda_{\max}(\overline{B}A) < 1$. Therefore

$$\rho(I - \overline{B}A) = \max\{|1 - \lambda_{\min}(\overline{B}A)|, |1 - \lambda_{\max}(\overline{B}A)|\} = 1 - \lambda_{\min}(\overline{B}A). \tag{4}$$

A more quantitative information on $\lambda_{\min}(\overline{B}A)$ is given in the following lemma.

Lemma 1 (Least Eigenvalue). *When B is symmetric and nonsingular,*

$$\lambda_{\min}(BA) = \inf_{u \in \mathscr{V}\setminus\{0\}} \frac{(ABAu, u)}{(Au, u)} = \inf_{u \in \mathscr{V}\setminus\{0\}} \frac{(Au, u)}{(B^{-1}u, u)} = \left(\sup_{u \in \mathscr{V}\setminus\{0\}} \frac{(B^{-1}u, u)}{(Au, u)} \right)^{-1}.$$

Proof. The first two identities comes from the fact BA is symmetric with respect to $(\cdot,\cdot)_A$ and $(\cdot,\cdot)_{B^{-1}}$. The third identity comes from

$$\lambda_{\min}^{-1}(BA) = \lambda_{\max}((BA)^{-1}) = \sup_{u \in \mathscr{V}\setminus\{0\}} \frac{((BA)^{-1}u, u)_A}{(u, u)_A} = \sup_{u \in \mathscr{V}\setminus\{0\}} \frac{(B^{-1}u, u)}{(Au, u)}.$$

This completes the proof. □

2.2 Space Decomposition and Method of Subspace Correction

In the spirit of dividing and conquering, we shall decompose the space $\mathscr{V}$ as the summation of subspaces. Then the original problem (1) can be split into sub-problems with smaller sizes which are relatively easier to solve.

Let $\mathscr{V}_i \subset \mathscr{V}$, $i = 0, \ldots, J$, be subspaces of $\mathscr{V}$. If $\mathscr{V} = \sum_{i=0}^{J} \mathscr{V}_i$, then $\{\mathscr{V}_i\}_{i=0}^{J}$ is called a *space decomposition* of $\mathscr{V}$, and we can write $u = \sum_{i=0}^{J} u_i$. Since $\sum_{i=0}^{J} \mathscr{V}_i$ is not necessarily a direct sum, decompositions of u are in general not unique.

Throughout this paper, we use the following operators, for $i = 0, 1, \ldots, J$:

- $Q_i : \mathscr{V} \mapsto \mathscr{V}_i$ the projection with the inner product $(\cdot,\cdot)$;
- $I_i : \mathscr{V}_i \mapsto \mathscr{V}$ the natural inclusion which is often called prolongation;
- $P_i : \mathscr{V} \mapsto \mathscr{V}_i$ the projection with the inner product $(\cdot,\cdot)_A$;
- $A_i : \mathscr{V}_i \mapsto \mathscr{V}_i$ the restriction of A to the subspace $\mathscr{V}_i$;
- $R_i : \mathscr{V}_i \mapsto \mathscr{V}_i$ an approximation of A_i^{-1} (often known as smoother).

It is easy to verify the relation $Q_i A = A_i P_i$ and $Q_i = I_i^t$. The operator I_i^t is often called restriction. If $R_i = A_I^{-1}$, then we have an exact local solver and $R_i Q_i A = P_i$.

For a given residual $r \in \mathscr{V}$, we let $r_i = I_i^t r$ denote the restriction of the residual to the subspace and solve the residual equation in the subspaces

$$A_i e_i = r_i \quad \text{by} \quad \hat{e}_i = R_i r_i.$$

Subspace corrections $\hat{e}_i$ are assembled to give a correction in the space $\mathscr{V}$ and therefore is called the method of subspace correction. There are two basic ways to assemble subspace corrections.

Parallel Subspace Correction (PSC)

This method performs the correction on each subspace in parallel. In operator form, it reads

$$u^{k+1} = u^k + B(f - Au^k), \tag{5}$$

where

$$B = \sum_{i=0}^{J} I_i R_i I_i^t. \tag{6}$$

The subspace correction is $\hat{e}_i = I_i R_i I_i^t (f - Au^k)$, and the correction in $\mathscr{V}$ is $\hat{e} = \sum_{i=0}^{J} \hat{e}_i$.

Successive Subspace Correction (SSC)

This method performs the correction in a successive way. In operator form, it reads

$$v^0 = u^k, \quad v^{i+1} = v^i + I_i R_i I_i^t (f - Av^i), i = 0, \ldots, J, \quad u^{k+1} = v^{J+1}. \tag{7}$$

We have the following error formulae for PSC and SSC:

- Parallel Subspace Correction (PSC):

$$u - u^{k+1} = \left[I - \Big(\sum_{i=0}^{J} I_i R_i I_i^t \Big) A \right] (u - u^k);$$

- Successive Subspace Correction (SSC):

$$u - u^{k+1} = \left[\prod_{i=0}^{J} (I - I_i R_i I_i^t A) \right] (u - u^k).$$

Thus PSC is also called additive method while SSC is called multiplicative method. In the notation $\prod_{i=0}^{J} a_i$, we assume there is a build-in ordering from $i = 0$ to J, i.e., $\prod_{i=0}^{J} a_i = a_0 a_1 \ldots a_J$.

As a trivial example, we consider the space decomposition $\mathbb{R}^J = \sum_{i=1}^{J} \operatorname{span}\{e_i\}$. In this case, if we use exact (one dimensional) subspace solvers, the resulting SSC is just the Gauss-Seidel method and the PSC is just the Jacobi method. More complicated examples, including multigrid methods and multilevel preconditioners, will be discussed later on.

PSC or SSC can be also understood as Jacobi or Gauss-Seidel methods for a bigger equation in the product space [43, 94], respectively. The analysis of classical iterative methods can then be applied to more advanced PSC or SSC methods.

Given a decomposition $\mathscr{V} = \sum_{i=0}^{J} \mathscr{V}_i$, we can construct a product space $\tilde{\mathscr{V}} = \mathscr{V}_0 \times \mathscr{V}_1 \times \ldots \times \mathscr{V}_J$, with an inner product $(\tilde{u}, \tilde{v})_{\tilde{\mathbb{V}}} = \sum_{i=0}^{J} (u_i, v_i)$. We will reformulate the linear operator equation $Au = f$ to an equation posed on $\tilde{\mathscr{V}} : \tilde{A}\tilde{u} = \tilde{f}$.

Let us introduce the operator $\mathscr{R} : \tilde{\mathscr{V}} \to \mathscr{V}$ by $\mathscr{R}\tilde{u} = \sum_{i=0}^{J} u_i$. Because of the decomposition $\mathscr{V} = \sum_{i=0}^{J} \mathscr{V}_i$, $\mathscr{R}$ is surjective. In general$\mathscr{R}$ is not injective but it will be in the quotient space $\overline{\mathscr{V}} = \tilde{\mathscr{V}} / \ker(\mathscr{R})$. We define $\mathscr{R}^* : \mathscr{V} \mapsto \tilde{\mathscr{V}}$, the adjoint of $\mathscr{R}$ with respect to $(\cdot,\cdot)_A$, to be

$$(\mathscr{R}^* u, \tilde{v})_{\tilde{\mathscr{V}}} := (u, \mathscr{R}\tilde{v})_A = \sum_{i=0}^{J} (u, v_i)_A = \sum_{i=0}^{J} (Q_i A u, v_i), \quad \text{for all } \tilde{v} = (v_i)_{i=0}^{J} \in \tilde{\mathscr{V}}.$$

Therefore

$$\mathscr{R}^* = (Q_0 A, Q_1 A, \cdots, Q_J A)^t.$$

Similarly, the transpose $\mathscr{R}^t : \mathscr{V} \mapsto \tilde{\mathscr{V}}$ of $\mathscr{R}$ with respect to $(\cdot,\cdot)$ is

$$\mathscr{R}^t = (Q_0, Q_1, \cdots, Q_J)^t.$$

Since $\mathscr{R}$ is surjective, we conclude that $\mathscr{R}^t$ is injective. Let $\tilde{A} = \mathscr{R}^* \mathscr{R}$ and $\tilde{f} = \mathscr{R}^t f$. If $\tilde{u}$ is a solution of $\tilde{A}\tilde{u} = \tilde{f}$, it is straightforward to verify that then $u = \mathscr{R}\tilde{u}$ is the solution of $Au = f$.

SSC as Gauss-Seidel Method

The new formulation of the problem is used to characterize SSC for solving $Au = f$ as a Gauss-Seidel method for $\tilde{A}\tilde{u} = \tilde{f}$. In the sequel, we consider the SSC applied to the space decomposition $\mathscr{V} = \sum_{k=0}^{J} \mathscr{V}_j$ with $R_i = A_i^{-1}$, namely we solve the problem posed on the subspaces exactly.

Let $\tilde{A} = \tilde{D} + \tilde{L} + \tilde{U}$ and $\tilde{B} = (\tilde{D} + \tilde{L})^{-1}$. Then SSC for $Au = f$ with exact local solvers $R_i = A_i^{-1}$ is equivalent to the Gauss-Seidel method for solving $\tilde{A}\tilde{u} = \tilde{f}$:

$$\tilde{u}^{k+1} = \tilde{u}^k + \tilde{B}(\tilde{f} - \tilde{A}\tilde{u}^k). \tag{8}$$

The verification of the equivalence is as follows. We first compute the entries for $\tilde{A} = (\tilde{a}_{ij})_{(J+1)\times(J+1)}$. By definition,

$$\tilde{a}_{ij} = Q_i A I_j = A_i P_i I_j : \mathscr{V}_j \mapsto \mathscr{V}_i.$$

In particular $\tilde{a}_{ii} = A_i : \mathscr{V}_i \mapsto \mathscr{V}_i$ is SPD on $\mathscr{V}_i$.

We can write the standard Gauss-Seidel method using iterator $\tilde{B} = (\tilde{D} + \tilde{L})^{-1}$ as

$$\tilde{u}^{k+1} = \tilde{u}^k + \tilde{D}^{-1}(\tilde{f} - \tilde{L}\tilde{u}^{k+1} - (\tilde{D} + \tilde{U})\tilde{u}^k).$$

The component-wise formula is

$$\begin{aligned} u_i^{k+1} &= u_i^k + A_i^{-1}(f_i - \sum_{j=0}^{i-1} \tilde{a}_{ij} u_j^{k+1} - \sum_{j=i}^{J} \tilde{a}_{ij} u_j^k) \\ &= u_i^k + A_i^{-1} Q_i (f - A\sum_{j=0}^{i-1} u_j^{k+1} - A\sum_{j=i}^{J} u_j^k). \end{aligned}$$

Let

$$v^i = \sum_{j=0}^{i-1} u_j^{k+1} + \sum_{j=i}^{J} u_j^k.$$

Noting that $v^i - v^{i-1} = u_i^{k+1} - u_i^k$, we then get

$$v^i = v^{i-1} + A_i^{-1} Q_i (f - A v^{i-1}),$$

which is the correction on $\mathscr{V}_i$.

Similarly one can easily verify that PSC using exact local solvers $R_i = A_i^{-1}$ is equivalent to the Jacobi method for solving the large system $\tilde{A}\tilde{u} = \tilde{f}$.

2.3 Sharp Convergence Identities

The analysis of additive multilevel operator relies on the following identity which is well known in the literature [87, 91, 42, 94]. For completeness, we include a concise proof taken from [94].

Theorem 1 (Identity for PSC). *If R_i is SPD on $\mathscr{V}_i$ for $i = 0,\ldots,J$, then B defined by (6) is also SPD on $\mathscr{V}$. Furthermore*

$$(B^{-1}v,v) = \inf_{\sum_{i=0}^J v_i = v} \sum_{i=0}^{J} (R_i^{-1} v_i, v_i), \tag{9}$$

and

$$\lambda_{\min}(BA)^{-1} = \sup_{\|v\|_A=1} \inf_{\sum_{i=0}^J v_i = v} (R_i^{-1} v_i, v_i). \tag{10}$$

Proof. Note that B is symmetric, and

$$(Bv,v) = (\sum_{i=0}^{J} I_i R_i I_i^t v, v) = \sum_{i=0}^{J} (R_i Q_i v, Q_i v),$$

whence B is invertible and thus SPD. We now prove (9) by constructing a decomposition achieving the infimum. Let $v_i^* = R_i Q_i B^{-1} v$, $i = 0,\ldots,J$. By definition of B, we get a special decomposition $\sum_i v_i^* = v$, and

$$\begin{aligned}\inf_{\sum v_i = v} \sum_{i=0}^{J} (R_i^{-1} v_i, v_i) &= \inf_{\sum w_i = 0} \sum_{i=0}^{J} (R_i^{-1}(v_i^* + w_i), v_i^* + w_i) \\ &= \sum_{i=0}^{J} (R_i^{-1} v_i^*, v_i^*) + \inf_{\sum w_i = 0} \Big[\sum_{i=0}^{J} 2(R_i^{-1} v_i^*, w_i) + \sum_{i=0}^{J} (R_i^{-1} w_i, w_i) \Big]\end{aligned}$$

Since

$$\sum_{i=0}^{J} (R_i^{-1} v_i^*, u_i) = \sum_{i=0}^{J} (B^{-1} v, u_i) = (B^{-1} v, \sum_{i=0}^{J} u_i)$$

for all $(u_i)_{i=0}^J \in \mathscr{V}$, we deduce

$$\begin{aligned}\inf_{\sum v_i = v} \sum_{i=0}^{J} (R_i^{-1} v_i, v_i) &= (B^{-1} v, \sum_{i=0}^{J} v_i^*) \\ &\quad + \inf_{\sum w_i = 0} \Big[2(B^{-1} v, \sum_{i=0}^{J} w_i) + \sum_{i=0}^{J} (R_i^{-1} w_i, w_i) \Big] = (B^{-1} v, v).\end{aligned}$$

The proof of the equality (10) is a simple consequence of Lemma 1. □

As for additive methods, we now present an identity developed by Xu and Zikatanov [94] for multiplicative methods. For simplicity, we focus on the case $R_i = A_i^{-1}$, $i = 0,\ldots,J$, i.e., the subspace solvers are exact. In this case $I - I_i R_i I_i^t A = I - P_i$.

Theorem 2 (X-Z Identity for SSC). *The following identity is valid*

$$\Big\| \prod_{i=0}^{J} (I - P_i) \Big\|_A^2 = 1 - \frac{1}{1+c_0}, \tag{11}$$

with

$$c_0 = \sup_{\|v\|_A=1} \inf_{\sum_{i=0}^J v_i = v} \sum_{i=0}^{J} \Big\| P_i \sum_{j=i+1}^{J} v_j \Big\|_A^2. \tag{12}$$

Proof. Recall that SSC for solving $Au = f$ with exact local solvers $R_i = A_i^{-1}$ is equivalent to the Gauss-Seidel method for solving $\tilde{A}\tilde{u} = \tilde{f}$ using iterator $\tilde{B} = (\tilde{D} + \tilde{L})^{-1}$. Let $\overline{B}$ be the symmetrization of $\tilde{B}$ from (2). Direct computation yields

$$\overline{B}^{-1} = \tilde{A} + \tilde{L}\tilde{D}^{-1}\tilde{U}. \tag{13}$$

On the quotient space $\overline{\mathscr{V}} = \tilde{V}/\ker(\mathscr{R})$, $\tilde{A}$ is SPD and thus defines an inner produce $(\cdot,\cdot)_{\tilde{A}}$. Using Lemma 1 and (13), we have

$$\begin{aligned} \|\tilde{I} - \tilde{B}\tilde{A}\|_{\tilde{A}}^2 = \|\tilde{I} - \overline{B}\tilde{A}\|_{\tilde{A}} &= 1 - \left[\sup_{\tilde{v}\in\overline{\mathscr{V}}\setminus\{0\}} \frac{(\overline{B}^{-1}\tilde{v},\tilde{v})_{\tilde{\mathscr{V}}}}{(\tilde{A}\tilde{v},\tilde{v})_{\tilde{\mathscr{V}}}} \right]^{-1} \\ &= 1 - \left[1 + \sup_{\tilde{v}\in\overline{\mathscr{V}}\setminus\{0\}} \frac{(\tilde{D}^{-1}\tilde{U}\tilde{v},\tilde{U}\tilde{v})}{(\tilde{A}\tilde{v},\tilde{v})} \right]^{-1}. \end{aligned}$$

To finish the proof, we verify that

$$\sup_{\tilde{v}\in\overline{\mathscr{V}},\tilde{v}\neq 0} \frac{(\tilde{D}^{-1}\tilde{U}\tilde{v},\tilde{U}\tilde{v})}{(\tilde{A}\tilde{v},\tilde{v})} = \sup_{v\in\mathscr{V},\|v\|_A=1} \inf_{\sum v_i = v} \sum_{i=0}^{J} \|P_i \sum_{j=i+1}^{J} v_j\|_A^2.$$

For any $\tilde{v} \in \overline{\mathscr{V}}$, and corresponding $v = \mathscr{R}\tilde{v}$, we have

$$(\tilde{A}\tilde{v},\tilde{v})_{\tilde{\mathscr{V}}} = (\mathscr{R}^*\mathscr{R}\tilde{v},\tilde{v})_{\tilde{\mathscr{V}}} = (\mathscr{R}\tilde{v},\mathscr{R}\tilde{v})_A = (v,v)_A,$$

and

$$(\tilde{D}^{-1}\tilde{U}\tilde{v},\tilde{U}\tilde{v})_{\tilde{\mathscr{V}}} = \sum_{i=0}^{J} (A_i^{-1} \sum_{j=i+1}^{J} A_i P_i v_j, \sum_{j=i+1}^{J} A_i P_i v_j)$$

because $Q_i A = A_i P_i$ and $\sum_{j=i+1}^{J} Q_j A v_j = A_i P_i \sum_{j=i+1}^{J} v_j$. Consequently,

$$(\tilde{D}^{-1}\tilde{U}\tilde{v},\tilde{U}\tilde{v})_{\tilde{\mathscr{V}}} = \sum_{i=0}^{J} (\sum_{j=i+1}^{J} P_i v_j, A_i \sum_{j=i+1}^{J} P_i v_j) = \sum_{i=0}^{J} \| \sum_{j=i+1}^{J} P_i v_j\|_A^2.$$

Since $\tilde{v} \in \overline{\mathscr{V}}$, we should use the quotient norm (which gives the inf) to finish the proof. $\Box$

For SSC method with general smoothers, we present the following sharp estimate of Xu and Zikatanov [94] (see also [30]). We refer to [94, 30] for a proof.

Theorem 3 (X-Z General Identity for SSC). *The SSC is convergent if each subspace solver $T_i = R_i Q_i A$ is convergent. Furthermore*

$$\left\| \prod_{i=1}^{J} (I - T_i) \right\|_A^2 = 1 - \frac{1}{K}, \quad K = 1 + \sup_{\|v\|=1} \inf_{\sum_i v_i = v} \sum_{i=1}^{J} \|T_i^* w_i\|_{\overline{T}_i^{-1}}^2 \tag{14}$$

where $w_i = \sum_{j=i}^{J} v_i - T_i^{-1} v_i$ and $\overline{T}_i := T_i^* + T_i - T_i^* T_i$.

3 Multilevel Methods on Quasi-Uniform Grids

In this section, we apply PSC and SSC to the finite element discretization of second order elliptic equations. We use theory developed in the previous section to give a convergence analysis of multilevel iteration methods.

3.1 Finite Element Methods

For simplicity we illustrate the technique by considering the linear finite element method for the Poisson equation.

$$-\Delta u = f \quad \text{in } \Omega, \quad \text{and} \quad u = 0 \quad \text{on } \partial\Omega, \tag{15}$$

where $\Omega \subset \mathbb{R}^d$ is a polyhedral domain.

3.1.1 Weak formulation

The weak formulation of (15) reads: given an $f \in H^{-1}(\Omega)$ find $u \in H_0^1(\Omega)$ so that

$$a(u,v) = \langle f, v \rangle \quad \text{for all } v \in H_0^1(\Omega), \tag{16}$$

where

$$a(u,v) = (\nabla u, \nabla v) = \int_\Omega \nabla u \cdot \nabla v \, \mathrm{d}x,$$

and $\langle \cdot, \cdot \rangle$ is the duality pair between $H^{-1}(\Omega)$ and $H_0^1(\Omega)$.

By the Poincaré inequality, $a(\cdot,\cdot)$ defines an inner product on $H_0^1(\Omega)$. Thus by the Riesz representation theorem, for any $f \in H^{-1}(\Omega)$, there exists a unique $u \in H_0^1(\Omega)$ such that (16) holds. Furthermore, we have the following regularity result. There exists $\alpha \in (0,1]$ which depends on the smoothness of $\partial\Omega$ such that

$$\|u\|_{1+\alpha} \lesssim \|f\|_{\alpha-1}. \tag{17}$$

This inequality is valid if Ω is convex or $\partial\Omega$ is $C^{1,1}$.

3.1.2 Triangulation and Properties

Let Ω be a polyhedral domain in $\mathbb{R}^d$. A triangulation $\mathscr{T}$ (also called mesh or grid) of Ω is a partition of $\overline{\Omega}$ into a set of d-simplexes.

We impose two conditions on a triangulation $\mathscr{T}$ which are important in finite element construction. First, a triangulation $\mathscr{T}$ is called *conforming* or *compatible* if the intersection of any two simplexes τ and τ' in $\mathscr{T}$ is either empty or a common lower dimensional simplex.

The second important condition is shape regularity. A set of triangulations $\mathbb{T}$ is called *shape regular* if there exists a constant σ_1 such that

$$\max_{\tau\in\mathscr{T}} \frac{\operatorname{diam}(\tau)^d}{|\tau|} \le \sigma_1, \quad \text{for all } \mathscr{T}\in\mathbb{T}, \tag{18}$$

where $\operatorname{diam}(\tau)$ is the diameter of τ and $|\tau|$ is the measure of τ in $\mathbb{R}^d$. For shape regular triangulations, $\operatorname{diam}(\tau) \eqsim h_\tau := |\tau|^{1/d}$ which will be used to represent the size of τ.

Furthermore, a shape regular class of triangulations $\mathbb{T}$ is called *quasi-uniform* if there exists a constant σ_2 such that

$$\frac{\max_{\tau\in\mathscr{T}} h_\tau}{\min_{\tau\in\mathscr{T}} h_\tau} \le \sigma_2, \quad \text{for all } \mathscr{T}\in\mathbb{T}.$$

For a quasi-uniform triangulation $\mathscr{T}$, we simply call $h = \max_{\tau\in\mathscr{T}} h_\tau$ the mesh size and denote $\mathscr{T}$ by $\mathscr{T}_h$.

3.1.3 Finite Element Approximation

The standard finite element method is to solve problem (16) in a piecewise polynomial finite dimensional subspace. For simplicity we consider the piecewise linear finite element space $\mathscr{V}_h \subset H_0^1(\Omega)$ on quasi-uniform triangulations $\mathscr{T}_h$ of Ω:

$$\mathscr{V}_h := \{v \in H_0^1(\Omega) : v|_\tau \in \mathscr{P}_1(\tau) \text{ for all } \tau\in\mathscr{T}\}.$$

We now solve (16) in the finite element space $\mathscr{V}_h$: find $u_h \in \mathscr{V}_h$ such that

$$a(u_h, v_h) = \langle f, v_h\rangle, \quad \text{for all } v_h \in \mathscr{V}_h. \tag{19}$$

The existence and uniqueness of the solution to (19) follows again from the Riesz representation theorem since $\mathscr{V}_h \subset H_0^1(\Omega)$. By approximation and regularity theory, we can easily get an error estimate on quasi-uniform grids

$$\|u-u_h\|_1 \lesssim h^\alpha \|u\|_{1+\alpha} \lesssim h^\alpha \|f\|_{\alpha-1},$$

where $\alpha > 0$ is determined by the regularity result (17). Thus u_h converges to u when $h \to 0$. When the solution u is rough, e.g., $\alpha \ll 1$, the convergence rate can be improved using adaptive grids [12, 79, 23, 63]. We will assume $\mathscr{V}_h$ is given, and the main objective of this paper is to discuss how to compute u_h efficiently. We focus on quasi-uniform grids in this section and on graded grids in the next section.

In this application, the SPD operator A is $(Au,v) = (\nabla u, \nabla v)$ and $\|\cdot\|_A$ is $|\cdot|_1$. For quasi-uniform mesh $\mathscr{T}_h$, let A_h be the restriction of A on the finite element space $\mathscr{V}_h$ over $\mathscr{T}_h$. We then end up with a linear operator equation $A_h : \mathscr{V}_h \mapsto \mathscr{V}_h$ that is

$$A_h u_h = f_h. \tag{20}$$

It is easy to see A_h is a self-adjoint operator in the Hilbert space $\mathscr{V}_h$ using L^2 inner product. To simplify notation in the sequel, we remove the subscript h when it is clear from the context and leads to no confusion.

It can be easily shown that $\kappa(A_h) \eqsim h^{-2}$ and the convergence rate of classical iteration methods, including Richardson, Jacobi, and Gauss-Seidel methods, for solving (19) is like

$$\rho \le 1 - Ch^2.$$

Thus when $h \to 0$, we observe slow convergence of those classical iterative methods. We will construct efficient iterative methods using multilevel space decompositions.

3.2 Multilevel Space Decomposition and Multigrid Method

We first present a multilevel space decomposition. Let us assume that we have an initial quasi-uniform triangulation $\mathscr{T}_0$ and a nested sequence of triangulations $\{\mathscr{T}_k\}_{k=0}^J$ where $\mathscr{T}_k$ is obtained by the uniform refinement of $\mathscr{T}_{k-1}$ for $k > 0$. We then get a nested sequence (in the sense of trees [63]) of quasi-uniform triangulations

$$\mathscr{T}_0 \le \mathscr{T}_1 \le \cdots \le \mathscr{T}_J = \mathscr{T}_h.$$

Note that h_k, the mesh size of $\mathscr{T}_k$, satisfies $h_k \eqsim \gamma^{2k}$ for some $\gamma \in (0,1)$, and thus $J \eqsim |\log h|$. Let $\mathscr{V}_k$ denote the corresponding linear finite element space of $H_0^1(\Omega)$ based on $\mathscr{T}_k$. We thus get a sequence of multilevel nested spaces

$$\mathscr{V}_0 \subset \mathscr{V}_1 \ldots \subset \mathscr{V}_J = \mathscr{V},$$

and a macro space decomposition

$$\mathscr{V} = \sum_{k=0}^{J} \mathscr{V}_k. \tag{21}$$

There is redundant overlapping in this multilevel decomposition, so the sum is not direct. The subspace solvers need only to take care of the "non-overlapping" components of the error (high frequencies in $\mathscr{V}_k$). For each subspace problem $A_k e_k = r_k$ posed on $\mathscr{V}_k$, we use a simple Richardson method

$$R_k = h_k^2 I_k,$$

where $I_k : \mathscr{V}_k \to \mathscr{V}_k$ is the identity and $h_k \approx \lambda_{\max}(A_k)$.

Let N_k be the dimension of $\mathscr{V}_k$, i.e., the number of interior vertices of $\mathscr{T}_k$. The standard nodal basis in $\mathscr{V}_k$ will be denoted by $\phi_{(k,i)}, i = 1, \cdots, N_k$. By our characterization of Richardson method, it is PSC method on the micro decomposition $\mathscr{V}_k = \sum_{i=1}^{N_k} \mathscr{V}_{(k,i)}$ with $\mathscr{V}_{(k,i)} = \operatorname{span}\{\phi_{(k,i)}\}$. In summary we choose the space decomposition:

$$\mathscr{V} = \sum_{k=0}^{J} \mathscr{V}_k = \sum_{k=0}^{J} \sum_{i=1}^{N_k} \mathscr{V}_{(k,i)}. \tag{22}$$

If we apply PSC to the decomposition (22) with $R_{(k,i)} = h_k^2 I_{(K,i)}$, we obtain $I_{(k,i)} R_{(k,i)} I_{(k,i)}^t u = h^{2-d}(u, \phi_{(k,i)})\phi_{(k,i)}$. The resulting operator B, according to (6), is the so-called BPX preconditioner [19]

$$Bu = \sum_{k=0}^{J} \sum_{i=1}^{N_k} h_k^{2-d}(u, \phi_{(k,i)})\phi_{(k,i)}. \tag{23}$$

If we apply SSC to the decomposition (22) with exact subspace solvers $R_i = A_i^{-1}$, we obtain a V-cycle multigrid method with Gauss-Seidel smoothers.

3.3 Stable Decomposition and Optimality of BPX Preconditioner

For the optimality of the BPX preconditioner, we are to prove that the condition number $\kappa(BA)$ is uniformly bounded and thus PCG using BPX preconditioner converges in a fixed number of steps for a given tolerance regardless of the mesh size.

The estimate $\lambda_{\min}(BA) \gtrsim 1$ follows from the stability of the subspace decomposition. The first result is on the macro decomposition $\mathscr{V} = \sum_{k=0}^{J} \mathscr{V}_k$.

Lemma 2 (Stability of Macro Decomposition). *For any $v \in \mathscr{V}$, there exists a decomposition $v = \sum_{k=0}^{J} v_k$ with $v_k \in \mathscr{V}_k, k = 0, \ldots, J$ such that*

$$\sum_{k=0}^{J} h_k^{-2} \|v_k\|^2 \lesssim |v|_1^2. \tag{24}$$

Proof. Following the chronological development, we present two proofs. The first one uses full regularity and the second one minimal regularity.

$\boxed{1}$ *Full regularity* H^2: We assume $\alpha = 1$ in (17), which holds for convex polygons or polyhedrons. Recall that $P_k : \mathscr{V} \to \mathscr{V}_k$ is the projection onto $\mathscr{V}_k$ with the inner product $(u,v)_A = (\nabla u, \nabla v)$, and let $P_{-1} = 0$. We prove that the following decomposition

$$v = \sum_{k=0}^{J} (P_k - P_{k-1})v \tag{25}$$

satisfies (24). The full regularity assumption leads to the L^2 error estimate of P_k via a standard duality argument:

$$\|(I-P_k)v\| \lesssim h_k |(I-P_k)v|_1, \quad \text{for all } v \in H_0^1(\Omega). \tag{26}$$

Since $\mathscr{V}_{k-1} \subset \mathscr{V}_k$, we have $P_{k-1}P_k = P_{k-1}$ and

$$P_k - P_{k-1} = (I - P_{k-1})(P_k - P_{k-1}). \tag{27}$$

In view of (26) and (27), we have

$$\sum_{k=0}^{J} h_k^{-2} \|(P_k - P_{k-1})v\|^2 = \sum_{k=0}^{J} h_k^{-2} \|(I-P_{k-1})(P_k - P_{k-1})v\|^2$$
$$\lesssim \sum_{k=0}^{J} |(P_k - P_{k-1})v|_{1,\Omega}^2 = |v|_{1,\Omega}^2.$$

In the last step, we have used the fact $(P_k - P_{k-1})v$ is the orthogonal decomposition in the A-inner product.

$\boxed{2}$ *Minimal regularity* H^1: We relax the H^2-regularity upon using the decomposition

$$v = \sum_{k=0}^{J} (Q_k - Q_{k-1})v, \tag{28}$$

where $Q_k : \mathscr{V} \to \mathscr{V}_k$ is the L^2-projection onto $\mathscr{V}_k$. A simple proof of nearly optimal stability of (28) proceeds as follows. Invoking approximability and H^1-stability of the L^2-projection Q_k on quasi-uniform grids, we infer that

$$\|(Q_k - Q_{k-1})u\| = \|(I - Q_{k-1})Q_k u\| \lesssim h_k |Q_k u|_1 \lesssim h_k |u|_1.$$

Therefore

$$\sum_{k=0}^{J} h_k^2 \|(Q_k - Q_{k-1})u\|^2 \lesssim J|u|_1^2 \lesssim |\log h| |u|_1^2.$$

The factor $|\log h|$ in the estimate can be removed by a more careful analysis based on the theory of Besov spaces and interpolation spaces. The following crucial inequality can be found, for example, in [91, 31, 64, 15, 65]:

$$\sum_{k=0}^{J} h_k^2 \|(Q_k - Q_{k-1})u\|^2 \lesssim |u|_1^2. \tag{29}$$

This completes the proof. □

We next state the stability of the micro decomposition. For a finite element space $\mathscr{V}$ with nodal basis $\{\phi_i\}_{i=1}^N$, let Q_{ϕ_i} be the L^2-projection to the one dimensional subspace spanned by ϕ_i. We have the following norm equivalence which says the nodal decomposition is stable in L^2. The proof is classical in the finite element analysis and thus omitted here.

Lemma 3 (Stability of Micro Decomposition). *For any $u \in \mathscr{V}$ over a quasi-uniform mesh $\mathscr{T}$, we have the norm equivalence*

$$\|u\|^2 \eqsim \sum_{i=1}^{N} \|Q_{\phi_i} u\|^2.$$

Theorem 4 (Stable Space Decomposition). *For any $v \in \mathscr{V}$, there exists a decomposition of v of the form*

$$v = \sum_{k=0}^{J} \sum_{i=1}^{N_k} v_{(k,i)}, \quad v_{(k,i)} \in \mathscr{V}_{(k,i)}, i = 1, \ldots, N_k, k = 0, \ldots, J,$$

such that

$$\sum_{k=0}^{J} \sum_{i=1}^{N_k} h_k^{-2} \|v_{(k,i)}\|^2 \lesssim |v|_1^2.$$

Consequently $\lambda_{\min}(BA) \gtrsim 1$ for the BPX preconditioner B defined in (23).

Proof. In light of Lemma 1, it suffices to combine Lemmas 2 and 3, and use (23). □

To estimate $\lambda_{\max}(BA)$, we first present a *strengthened Cauchy-Schwarz* (SCS) inequality for the macro decomposition.

Lemma 4 (Strengthened Cauchy-Schwarz Inequality (SCS)). *For any $u_i \in \mathscr{V}_i, v_j \in \mathscr{V}_j$, $j \geq i$, we have*

$$(u_i, v_j)_A \lesssim \gamma^{j-i} |u_i|_1 h_j^{-1} \|v_j\|_0,$$

where $\gamma < 1$ is a constant such that $h_i \eqsim \gamma^{2i}$.

Proof. Let us first prove the inequality on one element $\tau \in \mathscr{T}_i$. Using integration by parts, Cauchy-Schwarz inequality, trace theorem, and inverse inequality, we have

$$\begin{aligned}\int_\tau \nabla u_i \cdot \nabla v_j \, \mathrm{d}x &= \int_{\partial \tau} \frac{\partial u_i}{\partial n} v_j \, ds \lesssim \|\nabla u_i\|_{0,\partial\tau} \|v_j\|_{0,\partial\tau} \lesssim h_i^{-1/2} \|\nabla u_i\|_{0,\tau} h_j^{-1/2} \|v_j\|_{0,\tau} \\ &\lesssim \left(\frac{h_j}{h_i}\right)^{1/2} |u_i|_{1,\tau} h_j^{-1} \|v_j\|_{0,\tau} \approx \gamma^{j-i} |u_i|_{1,\tau} h_j^{-1} \|v_j\|_{0,\tau}.\end{aligned}$$

Adding over $\tau \in \mathscr{T}_i$, and using Cauchy-Schwarz again, yields

$$\begin{aligned}(\nabla u_i, \nabla v_j) &= \sum_{\tau \in \mathscr{T}_i} (\nabla u_i, \nabla v_j)_\tau \lesssim \gamma^{j-i} h_j^{-1} \sum_{\tau \in \mathscr{T}_i} |u_i|_{1,\tau} \|v_j\|_{0,\tau} \\ &\lesssim \gamma^{j-i} h_j^{-1} \Big(\sum_{\tau \in \mathscr{T}_i} |u_i|_{1,\tau}^2 \Big)^{1/2} \Big(\sum_{\tau \in \mathscr{T}_i} \|v_j\|_{0,\tau}^2 \Big)^{1/2} = \gamma^{j-i} |u_i|_1 h_j^{-1} \|v_j\|_0,\end{aligned}$$

which is the asserted estimate. □

Before we prove the main consequence of SCS, we need an elementary estimate.

Lemma 5 (Auxiliary Estimate). *Given $\gamma < 1$, we have*

$$\sum_{i,j=1}^{n} \gamma^{|j-i|} x_i y_j \le \frac{2}{1-\gamma} \Big(\sum_{i=1}^{n} x_i^2 \Big)^{1/2} \Big(\sum_{i=1}^{n} y_i^2 \Big)^{1/2} \quad \forall (x_i)_{i=1}^n, (y_i)_{i=1}^n \in \mathbb{R}^n.$$

Proof. Let $\Gamma \in \mathbb{R}^{n\times n}$ be the matrix $\Gamma = (\gamma^{|j-i|})_{i,j=1}^n$. The spectral radius $\rho(\Gamma)$ of Γ satisfies

$$\rho(\Gamma) \le \|\Gamma\|_1 = \max_{1\le j\le n} \sum_{i=1}^{n} \gamma^{|j-i|} \le \frac{2}{1-\gamma}.$$

Consequently, utilizing the Cauchy-Schwarz inequality yields

$$\sum_{i,j=1}^{n} \gamma^{|j-i|} x_i y_j = (\Gamma \mathbf{x}, \mathbf{y}) \le \rho(\Gamma) \|\mathbf{x}\|_2 \|\mathbf{y}\|_2 \quad \forall \mathbf{x} = (x_i)_{i=1}^n, \mathbf{y} = (y_i)_{i=1}^n \in \mathbb{R}^n,$$

which is the desired estimate. □

Theorem 5 (Largest Eigenvalue of *BA*). *For any $v \in \mathscr{V}$, we have*

$$(Av, v) \lesssim \inf_{\sum_{k=0}^J v_k = v} \sum_{k=0}^{J} h_k^{-2} \|v_k\|^2.$$

Consequently $\lambda_{\max}(BA) \lesssim 1$ for the BPX preconditioner B defined in (23).

Proof. For $v \in \mathscr{V}$, let $v = \sum_{k=0}^J v_k, v_k \in \mathscr{V}_k$, $k = 0, \ldots, J$, be an arbitrary decomposition. By the SCS inequality of Lemma 4, we have

$$(\nabla v, \nabla v) = 2 \sum_{k=0}^{J} \sum_{j=k+1}^{J} (\nabla v_k, \nabla v_j) + \sum_{k=0}^{J} (\nabla v_k, \nabla v_k) \lesssim \sum_{k=0}^{J} \sum_{j=k}^{J} \gamma^{j-k} |v_k|_1 h_j^{-1} \|v_j\|.$$

Combining Lemma 5 with the inverse estimate $|v_k|_1 \lesssim h_k^{-1} \|v_k\|$, we obtain

$$(\nabla v, \nabla v) \lesssim (\sum_{k=0}^{J} |v_k|_1^2)^{1/2} (\sum_{k=0}^{J} h_k^{-2} \|v_k\|^2)^{1/2} \lesssim \sum_{k=0}^{J} h_k^{-2} \|v_k\|^2.$$

which is the assertion. □

We finally prove the optimality of the BPX preconditioner.

Corollary 1 (Optimality of BPX Preconditioner). *For the preconditioner B defined in (23), we have*

$$\kappa(BA) \lesssim 1$$

Proof. Simply combine Theorems 4 and 5. □

3.4 Uniform Convergence of V-cycle Multigrid

In this section, we prove the uniform convergence of V-cycle multigrid, namely SSC applied to the decomposition (22) with exact subspace solvers.

Lemma 6 (Nodal Decomposition). *Let $\mathscr{T}$ be a quasi-uniform triangulation with N nodal basis ϕ_i. For the nodal decomposition*

$$v = \sum_{i=1}^{N} v_i, \quad v_i = v(x_i)\phi_i,$$

we have

$$\sum_{i=1}^{N} \left| P_i \sum_{j>i}^{N} v_j \right|_1^2 \lesssim h^{-2} \|v\|^2.$$

Proof. For every $1 \le i \le N$, we define the index set $L_i := \{j \in \mathbb{N} : i < j \le N, \operatorname{supp}\phi_j \cap \operatorname{supp}\phi_i \neq \emptyset\}$ and $\Omega_i = \cup_{j\in L_i} \operatorname{supp}\phi_j$. Since $\mathscr{T}$ is shape-regular, the numbers of integers in each L_i is uniformly bounded. So we have

$$\sum_{i=1}^{N} \left| P_i \sum_{j>i}^{N} v_j \right|_{1,\Omega}^2 = \sum_{i=1}^{N} \left| P_i \sum_{j\in L_i} v_j \right|_{1,\Omega}^2 \lesssim \sum_{i=1}^{N} \sum_{j\in L_i} |v_j|_{1,\Omega_i}^2 \lesssim \sum_{i=1}^{N} |v_i|_{1,\Omega_i}^2 \lesssim \sum_{i=1}^{N} h_i^{-2} \|v_i\|_{0,\Omega_i}^2,$$

where we have used an inverse inequality in the last step. Since $\mathscr{T}$ is quasi-uniform, and the nodal basis decomposition is stable in the L^2 inner product (Lemma 3), i.e. $\sum_{i=1}^{N} \|v_i\|_{0,\Omega_i}^2 \approx \|v\|_{0,\Omega}^2$, we deduce

$$\sum_{i=1}^{N} \left| P_i \sum_{j>i}^{N} v_j \right|_{1,\Omega}^2 \lesssim h^{-2} \|v\|_{0,\Omega}^2,$$

which is the desired estimate. □.

Lemma 7 (H^1 vs L^2 Stability). *The following inequality holds for all $v \in \mathscr{V}$*

$$\sum_{k=0}^{J} |(P_k - Q_k)v|_{1,\Omega}^2 \lesssim \sum_{k=0}^{J} h_k^{-2} \|(Q_k - Q_{k-1})v\|^2. \tag{30}$$

Proof. We first use the definition of P_k, together with $(I-Q_k)v = \sum_{j=k+1}^{J}(Q_j - Q_{j-1})v$, to write

$$\sum_{k=0}^{J} |(P_k - Q_k)v|_1^2 = \sum_{k=0}^{J} ((P_k - Q_k)v, (I - Q_k)v)_A$$
$$= \sum_{k=0}^{J} \sum_{j=k+1}^{J} ((P_k - Q_k)v, (Q_j - Q_{j-1})v)_A.$$

Applying now Lemma 4 yields

$$\sum_{k=0}^{J} |(P_k - Q_k)v|_1^2 \lesssim \left(\sum_{k=1}^{J} |(P_k - Q_k)v|_1^2\right)^{1/2} \left(\sum_{k=0}^{J} h_k^{-2} \|(Q_k - Q_{k-1})v\|^2\right)^{1/2}.$$

The desired result then follows. □

Theorem 6 (Optimality of V-cycle Multigrid). *The V-cycle multigrid method, using SSC applied to the decomposition (22) with exact subspace solvers $R_i = A_i^{-1}$, converges uniformly.*

Proof. We use the telescopic multilevel decomposition

$$v = \sum_{k=0}^{J} v_k, \quad v_k = (Q_k - Q_{k-1})v,$$

along with the nodal decomposition

$$v_k = \sum_{i=1}^{N_k} v_{(k,i)}, \quad v_{(k,i)} = v_k(x_i)\phi_{(k,i)},$$

for each level k. By the X-Z identity of Theorem 2, it suffices to prove the inequality

$$\sum_{k=0}^{J} \sum_{i=1}^{N_k} |P_{(k,i)} \sum_{(l,j)>(k,i)} v_{(l,j)}|_1^2 \lesssim |v|_1^2, \tag{31}$$

where the inner sum is understood in lexicographical order. We first simplify the left hand side of (31) upon writing

$$\sum_{(l,j)>(k,i)} v_{(l,j)} = \sum_{j>i}^{N_k} v_{(k,j)} + \sum_{l>k} v_l = \sum_{j>i}^{N_k} v_{(k,j)} + (v - Q_k v).$$

We apply Lemma 6 and the stable decomposition (29) to get

$$\sum_{k=0}^{J} \sum_{i=1}^{N_k} |P_{(k,i)} \sum_{j>i} v_{(k,j)}|_1^2 \lesssim \sum_{k=0}^{J} h_k^{-2} \|v_k\|^2 \lesssim |v|_1^2.$$

We now estimate the remaining terms $|P_0(v-Q_0)v|^2$ and $\sum_{k=1}^J |P_{(k,i)}(v-Q_k v)|_{1,\Omega}^2$. For any function $u \in \mathscr{V}$,

$$\sum_{i=1}^{N_k} |P_{(k,i)}u|_1^2 = \sum_{i=1}^{N_k} |P_{(k,i)}P_k u|_{1,\Omega_{(k,i)}}^2 \le \sum_{i=1}^{N_k} |P_k u|_{1,\Omega_{(k,i)}}^2 \lesssim |P_k u|_1^2.$$

Thus, by (30) and (29), we get

$$\begin{aligned} &|P_0(v-Q_0)v|^2 + \sum_{k=1}^{J} |P_k(v-Q_k v)|_1^2 \\ &\qquad \lesssim \sum_{k=0}^{J} |(P_k - Q_k)v|_1^2 \lesssim \sum_{k=0}^{J} h_k^{-2}\|(Q_k - Q_{k-1})v\|^2 \lesssim |v|_1^2. \end{aligned}$$

This completes the proof. □

The proof of Theorem 6 hinges on Theorem 2 (X-Z identity), which in turn requires exact solvers $R_i = A_i^{-1}$ and makes $P_i = A_i^{-1}Q_iA$ the key operator to appear in (11). If the smoothers R_i are not exact, namely $R_i \neq A_i^{-1}$, then the key operator becomes $T_i = R_iQ_iA$ and Theorem 2 must be replaced by Theorem 3. We refer to [30] for details.

3.5 Systems with Strongly Discontinuous Coefficients

Elliptic problems with strongly discontinuous coefficients arise often in practical applications and are notoriously difficult to solve for iterative methods such as multigrid and domain decomposition. We are interested in the performance of these methods with respect to jumps. Consider the following model problem

$$\begin{cases} -\nabla \cdot (\omega \nabla u) = f & \text{in } \Omega, \\ u = g_D & \text{on } \Gamma_D, \\ -\omega \frac{\partial u}{\partial n} = g_N & \text{on } \Gamma_N \end{cases} \tag{32}$$

where $\Omega \in \mathbb{R}^d (d = 1,\ 2 \text{ or } 3)$ is a polygonal or polyhedral domain with Dirichlet boundary Γ_D and Neumann boundary Γ_N. We assume that the coefficient function $\omega = \omega(x)$ is positive and piecewise constant with respect to given subdomains Ω_m $(m = 1, \cdots, M)$ with $\overline{\Omega} = \cup_{m=1}^M \overline{\Omega}_m$, i.e., $\omega|_{\Omega_m} = \omega_m$ and

$$\mathscr{J}(\omega) \equiv \frac{\omega_{\max}}{\omega_{\min}} \gg 1.$$

These subdomains Ω_m are matched by the initial grid $\mathscr{T}_0$.

The question is how to make multigrid and domain decomposition methods converge (nearly) uniformly, not only with respect to the mesh size, but also with respect to the jump $\mathscr{J}(\omega)$. There has been a lot of interest in the development of iterative methods with robust convergence rates with respect to the size of both jumps and mesh; see [17, 25, 77, 85, 86, 95] and the references cited therein. Domain decomposition (DD) methods have been developed for this purpose with special coarse spaces [95]. We refer to the monograph [82] and the survey [24] for a summary on DD methods. However, in general, the convergence rates of multigrid and domain decomposition methods are known to deteriorate with respect to $\mathscr{J}(\omega)$, especially in three dimensions.

The BPX and overlapping domain decomposition preconditioners are proven to be robust for some special cases: interface has no cross points [20, 66]; every subdomain touches part of the Dirichlet boundary [93]; and quasi-monotone coefficients [33, 34]. If the number of levels is fixed, multigrid converges uniformly with the convergence rate $\rho_k \le 1-\delta^k$ where $\delta \in (0,1)$ is a constant and k is the number of levels. In general, the worst convergence rate is $1-Ch$ and, for BPX preconditioned system, $\sup_\omega \kappa(BA) \ge Ch^{-1}$ (see [66, 90]).

An interesting open problem is how to make multigrid method work uniformly with respect to jumps without introducing "expensive" coarse spaces. Recently, Xu and Zhu [93] proved that BPX and multigrid V-cycle lead to a nearly uniform convergent preconditioned conjugate gradient method (see [97] for a similar result on DD preconditioners). We now report this result.

Theorem 7 (Nearly Optimal PCG). *For BPX and multigrid V-cycle preconditioners (without using any special coarse spaces), PCG converges uniformly with respect to jumps in the sense that there exist c_0, c_1 and m_0 so that*

$$\|u-u_k\|_A \le 2(c_0/h-1)^{m_0}(1-c_1/|\log h|)^{k-m_0}\|u-u_0\|_A \quad (k \ge m_0), \tag{33}$$

where m_0 is a fixed number depending only on the distribution of the coefficients.

This result is motivated by [41, 84] where PCG with diagonal scaling or overlapping DD is considered, and the following convergence result is proved by using pure algebraic methods:

$$\|u-u_k\|_A \le C(h, \mathscr{J}(\omega))(1-ch)^{k-m_0}\|u-u_0\|_A.$$

Unfortunately, this estimate deteriorates severely with respect to mesh size. The improved estimate (33) implies that after m_0 steps, the convergent rate of the PCG is nearly uniform with respect to the mesh size and uniform with respect to jumps. The first m_0 steps are necessary for PCG to deal with small eigenvalues created by the jumps. To account for the effect of a finite cluster of eigenvalues in the convergence rate of PCG, the following estimate from [45] will be instrumental. Suppose that we can split the spectrum $\sigma(BA)$ of BA into two sets $\sigma_0(BA)$ and $\sigma_1(BA)$, where σ_0 consists of all "bad" eigenvalues and the remaining eigenvalues in σ_1 are bounded above and below.

Theorem 8 (CG for Clusters of Eigenvalues). *If $\sigma(BA) = \sigma_0(BA) \cup \sigma_1(BA)$ is such that $\sigma_0(BA)$ contains m eigenvalues and $\lambda \in [a,b]$ for each $\lambda \in \sigma_1(BA)$, then*

$$\|u-u_k\|_A \le 2(\kappa(BA)-1)^m \left(\frac{\sqrt{b/a}-1}{\sqrt{b/a}+1}\right)^{k-m} \|u-u_0\|_A.$$

Proof of Theorem 7. We introduce the weighted L^2 and H^1 inner products and corresponding norms

$$(u,v)_{0,\omega} = \int_\Omega uv\omega\,\mathrm{d}x = \sum_{m=1}^{M} \omega_m (u,v)_{\Omega_m},\ \|u\|_{0,\omega} = (u,u)_\omega^{1/2},$$

$$(u,v)_{1,\omega} = \int_\Omega \nabla u\cdot\nabla v\omega\,\mathrm{d}x = \sum_{m=1}^{M} \omega_m (u,v)_{1,\Omega_m},\ \|u\|_{1,\omega} = (u,u)_{1,\omega}^{1/2}.$$

The SPD operator A and corresponding inner product of finite element discretization of (32) is $(Au,v) = (u,v)_{1,\omega}$. Let $\mathscr{V}_h$ be the linear finite element space based on a shape regular triangulation $\mathscr{T}_h$. The weighted L^2-projection to $\mathscr{V}_h$ with respect to $(\cdot,\cdot)_{0,\omega}$ will be denoted by Q_h^ω.

We now introduce the following auxiliary subspace:

$$\widetilde{\mathscr{V}_h} = \left\{ v\in \mathscr{V}_h : \int_{\Omega_m} v\,\mathrm{d}x = 0,\ |\partial\Omega_m \cap \Gamma_D| = 0 \right\}.$$

Note that this subspace satisfies $\dim(\widetilde{\mathscr{V}_h}) = n - m_0$ where $m_0 < M$ is a fixed number, and more importantly,

$$\|v\|_{0,\omega} \lesssim |v|_{1,\omega} \quad \text{for all } v\in \widetilde{\mathscr{V}_h}.$$

As a consequence, we obtain the approximation and stability of the weighted L^2-projection Q_h^ω (see [20, 93, 97]),

$$\|(I-Q_h^\omega)v\|_{0,\omega} \lesssim h\,|\log h|^{\frac{1}{2}}\,|v|_{1,\omega}, \quad |Q_h^\omega v|_{1,\omega} \lesssim |\log h|^{\frac{1}{2}}\,|v|_{1,\omega}, \quad \text{for all } v\in \widetilde{\mathscr{V}_h}.$$

Using the arguments in Lemma 2-step 2, we can prove that the decomposition using weighted L^2 projection is almost stable, i.e.,

$$\sum_{k=0}^{J} h_k^{-2}\|(Q_k^\omega - Q_{k-1}^\omega)u\|^2 \lesssim |\log h|^2 |u|_{1,\omega}^2. \tag{34}$$

Repeating the argument of Theorem 4, we obtain the estimate $\lambda_{\min}(BA) \gtrsim |\log h|^{-2}$.

On the other hand, the strengthened Cauchy Schwarz inequality (SCS) of Lemma 4 is valid for weighted inner products because its proof can be carried out element-wise when ω is piecewise constant. Consequently Theorem 5 holds for weighted L^2-norm and implies $\lambda_{\max}(BA) \lesssim 1$. We thus infer that the condition number of BA restricted to $\widetilde{\mathscr{V}_h}$ is nearly uniformly bounded, namely $\kappa(BA) \lesssim |\log h|^2$.

To estimate the convergent rate of PCG in the space $\mathscr{V}_h$, we introduce the mth effective condition number by $\kappa_{m+1}(A) = \lambda_{\max}(A)/\lambda_{m+1}(A)$, where $\lambda_{m+1}(A)$ is the $(m+1)$th minimal eigenvalue of A. By the Courant "minmax" principle (see e.g., [40])

$$\lambda_{m+1}(A) = \max_{S, \dim(S)=m} \min_{0 \neq v \in S^{\perp}} \frac{(Av, v)_{0,\omega}}{(v, v)_{0,\omega}}.$$

In particular, the fact $\dim(\widetilde{\mathscr{V}}_h) = n - m_0$, together with the nearly stable decomposition (34), implies that $\lambda_{m_0+1}(BA) \geq |\log h|^{-2}$.

The asserted estimate finally follows from Theorem 8 □.

Results such as Theorem 8 provide convincing evidence of a general rule of thumb: an iterative method, whenever possible, should be used together with certain preconditioned Krylov space (such as conjugate gradient) method.

4 Multilevel Methods on Graded Grids

Adaptive methods are now widely used in scientific and engineering computation to optimize the relation between accuracy and computational labor (degrees of freedom). Let $\mathscr{V}_0 \subseteq \mathscr{V}_1 \subseteq \cdots \subseteq \mathscr{V}_J = \mathscr{V}$ be nested finite element spaces obtained by local mesh refinement. A standard multilevel method contains a smoothing step on the spaces $\mathscr{V}_j, j = 0, \ldots, J$. For graded grids obtained by adaptive procedure, it is possible that $\mathscr{V}_j$ results from $\mathscr{V}_{j-1}$ by just adding few, say one, basis function. Thus smoothing on both $\mathscr{V}_j$ and $\mathscr{V}_{j-1}$ leads to a lot of redundancy. If we let N be the number of unknowns in the finest space $\mathscr{V}$, then the complexity of smoothing can be as bad as $\mathscr{O}(N^2)$ [62]. To achieve optimal complexity $\mathscr{O}(N)$, the smoothing in each space $\mathscr{V}_j$ must be restricted to the new unknowns and their neighbors. Such methods are referred to as *adaptive multilevel methods* or *local multilevel methods*.

Of all possible refinement strategies, we are interested in *bisection*, the most popular and effective procedure for refinement in any dimension [6, 9, 56, 59, 68, 69, 70, 71, 72, 74, 80, 83]. We refer to [31] for the optimality of BPX preconditioner for regular refinement (one triangle is divided into four similar triangles) in 2-D and [1] for similar results in 3-D (one tetrahedron is divided into eight tetrahedrons).

We still consider the finite element approximation of Poisson equation (15); see Section §3.1 for the problem setting. The additional difficulty is that the mesh is no longer quasi-uniform. We present a decomposition of bisection grids and transfer results from quasi-uniform grids to bisection grids. As an example, we present a stable decomposition of finite element spaces and SCS inequality. The optimality of BPX preconditioner and uniform convergence of multigrid can then be established upon applying the general theory of Section §3; we refer to [27].

4.1 Bisection Methods

In this section, we introduce bisection methods for simplicial grids and present a novel decomposition of conforming triangulations obtained by bisection methods.

Given a simplex τ, we assign one of its edges as the *refinement edge* of τ. Starting from an initial triangulation $\mathscr{T}_0$, a bisection method consists of the following rules:

R1. assign refinement edges for each element $\tau \in \mathscr{T}_0$;
R2. divide a simplex with a refinement edge into two simplexes;
R3. assign refinement edges to the two children of a bisected simplex.

We now give a mathematical description. Let τ be a simplex that bisects into simplexes τ_1 and τ_2. R2 can be described by a mapping $b_\tau : \{\tau\} \to \{\tau_1, \tau_2\}$. If we denote a simplex τ with a refinement edge e by a pair (τ, e), then R2 and R3 can be described by a mapping $\{(\tau,e)\} \to \{(\tau_1,e_1),(\tau_2,e_2)\}$. The pair (τ,e) is called a *labeled simplex* and the set $(\mathscr{T},L) := \{(\tau,e) : \tau \in \mathscr{T}\}$ is called a *labeled triangulation*. Then R1 can be described by a mapping $\mathscr{T}_0 \to (\mathscr{T}_0, L)$ and called *initial labeling*. The first rule is an essential ingredient of bisection methods. Once the initial labeling is done, the subsequent grids inherit labels according to R2-R3 such that the bisection process can proceed. We refer to [63, Section 4] for details.

For a labeled triangulation $(\mathscr{T},L)$, and a bisection $b_\tau : \{(\tau,e)\} \to \{(\tau_1,e_1),(\tau_2,e_2)\}$ for $\tau \in \mathscr{T}$, we define a formal addition

$$\mathscr{T} + b_\tau := (\mathscr{T},L)\backslash\{(\tau,e)\} \cup \{(\tau_1,e_1),(\tau_2,e_2)\}.$$

For a sequence of bisections $\mathscr{B} = (b_{\tau_1}, b_{\tau_2}, \cdots, b_{\tau_N})$, we define

$$\mathscr{T} + \mathscr{B} := ((\mathscr{T} + b_{\tau_1}) + b_{\tau_2}) + \cdots + b_{\tau_N},$$

whenever the addition is well defined (i.e. τ_i should exists in the previous labeled triangulation). These additions are a convenient mathematical description of bisection on triangulations.

Given a labeled initial grid $\mathscr{T}_0$ of Ω and a bisection method, we define

$$\mathbb{F}(\mathscr{T}_0) = \{\mathscr{T} : \text{ there exists a bisection sequence } \mathscr{B} \text{ such that } \mathscr{T} = \mathscr{T}_0 + \mathscr{B}\},$$
$$\mathbb{T}(\mathscr{T}_0) = \{\mathscr{T} \in \mathbb{F}(\mathscr{T}_0) : \mathscr{T} \text{ is conforming}\}.$$

Therefore $\mathbb{F}(\mathscr{T}_0)$ contains all triangulations obtained from $\mathscr{T}_0$ using the bisection method, and is unique once the rules R1-3 have been set. But a triangulation $\mathscr{T} \in \mathbb{F}(\mathscr{T}_0)$ could be non-conforming and thus we define $\mathbb{T}(\mathscr{T}_0)$ as a subset of $\mathbb{F}(\mathscr{T}_0)$ containing only conforming triangulations.

We also define the sequence of uniformly refined meshes $\{\overline{\mathscr{T}}_k\}_{k=0}^\infty$ by:

$$\overline{\mathscr{T}}_0 = \mathscr{T}_0, \text{ and } \overline{\mathscr{T}}_k = \overline{\mathscr{T}}_{k-1} + \{b_\tau : \tau \in \overline{\mathscr{T}}_{k-1}\}, \text{ for } k \geq 1.$$

This means that $\overline{\mathscr{T}}_k$ is obtained by bisecting all elements in $\overline{\mathscr{T}}_{k-1}$ only once. Note that $\overline{\mathscr{T}}_k \in \mathbb{F}(\mathscr{T}_0)$ but not necessarily in the set $\mathbb{T}(\mathscr{T}_0)$.

We consider bisection methods which satisfy the following two assumptions:

(B1) Shape Regularity: $\mathbb{F}(\mathscr{T}_0)$ is shape regular.

(B2) Conformity of Uniform Refinement: $\overline{\mathscr{T}}_k \in \mathbb{T}(\mathscr{T}_0)$, i.e., $\overline{\mathscr{T}}_k$ is conforming for all $k \geq 0$.

All existing bisection methods share the same rule R2 described now. Given a simplex τ with refinement edge e, the two children of τ are defined by bisecting e and connecting the midpoint of e to the other vertices of τ. More precisely, let $\{x_1, x_2, \cdots, x_{d+1}\}$ be vertices of τ and let $e = \overline{x_1x_2}$ be the refinement edge. Let x_m denote the midpoint of e. The children of τ are two simplexes τ_1 with vertices $\{x_1, x_m, x_3, \cdots, x_{d+1}\}$ and τ_2 with $\{x_2, x_m, x_3, \cdots, x_{d+1}\}$; we refer to [63, Section 4] for a through discussion of the notion of vertex type order and type. There is another class of refinement method, called regular refinement, which divide one simplex into 2^d children; see [8, 58].

All existing bisection methods differ in R1 and R3. For the so-called *longest edge bisection* [68, 70, 71, 72, 69], the refinement edge of a simplex is always assigned as the longest edge of this simplex. It is also used in R1 to assign the longest edge for each element in the initial triangulation. This method is simple, but (B1) is only proved for two dimensional triangulations [72] and (B2) only holds for special cases.

Regarding R3, the *newest vertex bisection* method for two dimensional triangulations assigns the edge opposite to the newest vertex of each child as their refinement edge. Sewell [76] showed that all the descendants of a triangle in $\mathscr{T}_0$ fall into four similarity classes and hence (B1) holds. Note that (B2) may not hold for an arbitrary rule R1, namely the refinement edge for elements in the initial triangulation cannot be selected freely. Mitchell [60] came up with a rule R1 for which (B2) holds. He proved the existence of such initial labeling scheme (so-called *compatible initial labeling*), and Biedl, Bose, Demaine, and Lubiw [11] gave an optimal $\mathscr{O}(N)$ algorithm to find a compatible initial labeling for a triangulation with N elements. In summary, in two dimensions, newest vertex bisection with compatible initial labeling is a bisection method which satisfies (B1) and (B2).

There are several bisection methods proposed in three and higher dimensions which generalize the newest vertex bisection in two dimensions [9, 56, 67, 6, 59, 80]. We shall not give detailed description of these bisection methods since the description of rules R1 and R3 is very technical for three and higher dimensions; we refer to [63, Section 4]. In these methods, (B1) is relatively easy to prove by showing all descendants of a simplex in $\mathscr{T}_0$ fall into similarity classes. As in the two dimensional case, (B2) requires special initial labeling, i.e., R1. We refer to Kossaczký [56] for the discussion of such rule in three dimensions and Stevenson [80] for the generalization to d-dimensions. However the algorithms proposed in [56, 80] to enforce such initial labeling consist of modifying the initial triangulation by further refinement of each element, which deteriorates the shape regularity. Although (B2) imposes a severe restriction on the initial labeling, we emphasize that it is also used to prove the optimal complexity of adaptive finite element methods [23, 63].

4.2 Compatible Bisections

The set of vertices of the triangulation $\mathscr{T}$ will be denoted by $\mathscr{N}(\mathscr{T})$ and the set of all edges will be denoted by $\mathscr{E}(\mathscr{T})$. For a vertex $x \in \mathscr{N}(\mathscr{T})$ or an edge $e \in \mathscr{E}(\mathscr{T})$, we define the *first ring* of x or e to be

$$\mathscr{R}_x = \{\tau \in \mathscr{T} \,|\, x \in \tau\}, \quad \mathscr{R}_e = \{\tau \in \mathscr{T} \,|\, e \subset \tau\},$$

and the local patch of x or e as $\omega_x = \cup_{\tau\in\mathscr{R}_x}\tau$, and $\omega_e = \cup_{\tau\in\mathscr{R}_e}\tau$. Note that ω_x and ω_e are subsets of Ω, while $\mathscr{R}_x$ and $\mathscr{R}_e$ are subsets of $\mathscr{T}$ which can be thought of as triangulations of ω_x and ω_e, respectively. The cardinality of a set S will be denoted by $\#S$.

Given a labeled triangulation $(\mathscr{T}, L)$, an edge $e \in \mathscr{E}(\mathscr{T})$ is called a *compatible edge* if e is the refinement edge of τ for all $\tau \in \mathscr{R}_e$. For a compatible edge, the ring $\mathscr{R}_e$ is called a *compatible ring*, and the patch ω_e is called a *compatible patch*. Let x be the midpoint of e and $\mathscr{R}_x$ be the ring of x in $\mathscr{T} + \{b_\tau : \tau \in \mathscr{R}_e\}$. A *compatible bisection* is a mapping $b_e : \mathscr{R}_e \to \mathscr{R}_x$. We then define the addition

$$\mathscr{T} + b_e := \mathscr{T} + \{b_\tau : \tau \in \mathscr{R}_e\} = \mathscr{T} \backslash \mathscr{R}_e \cup \mathscr{R}_x.$$

For a compatible bisection sequence $\mathscr{B}$, the addition $\mathscr{T} + \mathscr{B}$ is defined as before.

Note that if $\mathscr{T}$ is conforming, then $\mathscr{T} + b_e$ is conforming for a compatible bisection b_e, whence compatible bisections preserve the conformity of triangulations. Hence, compatible bisection is a fundamental concept both in theory and practice.

In two dimensions, a compatible bisection b_e has only two possible configurations; see Fig. 1. One is bisecting an interior compatible edge, in which case the patch ω_e is a quadrilateral. Another case is bisecting a boundary edge, which is always compatible, and ω_e is a triangle. In three dimensions, the configuration of compatible bisections depends on the initial labeling; see Fig. 2 for a simple case.

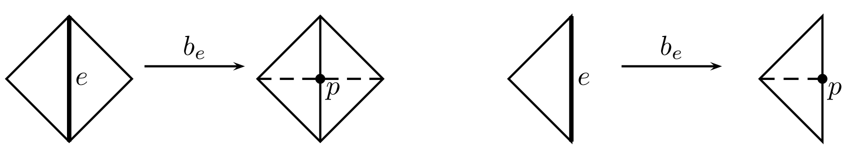
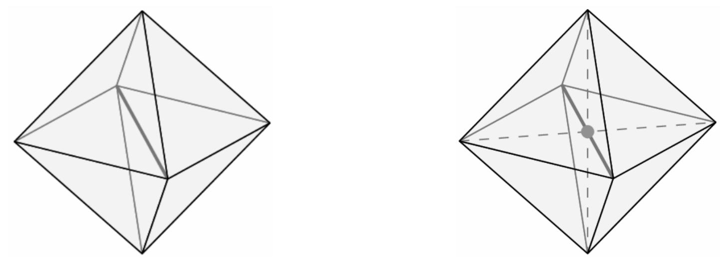


Fig. 1 Two compatible bisections for $d = 2$. Left: interior edge; right: boundary edge. The edge with boldface is the compatible refinement edge, and the dash-line represents the bisection.

The bisection of paired triangles was first introduced by Mitchell [60, 61]. The idea was generalized by Kossaczký [56] to three dimensions, and Maubach [59] and Stevenson [80] to any dimension. In the aforementioned references, efficient recursive completion procedures of bisection methods are introduced based on compatible bisections. We use them to characterize the conforming mesh obtained by bisection methods.



(a) A compatible patch ω_e (b) After a compatible bisection ω_p

Fig. 2 A compatible bisection for $d=3$: the edge e (in bold) is the refinement edge all elements in the patch ω_e. Connecting e to the other vertices bisects each element of the compatible ring $\mathscr{R}_e$ and keeps the mesh conforming without spreading refinement outside ω_e. This is an atomic operation.

4.3 Decomposition of Bisection Grids

We now present a decomposition of meshes in $\mathbb{T}(\mathscr{T}_0)$ using compatible bisections. This is due to Chen, Nochetto, and Xu [27] and will be instrumental later.

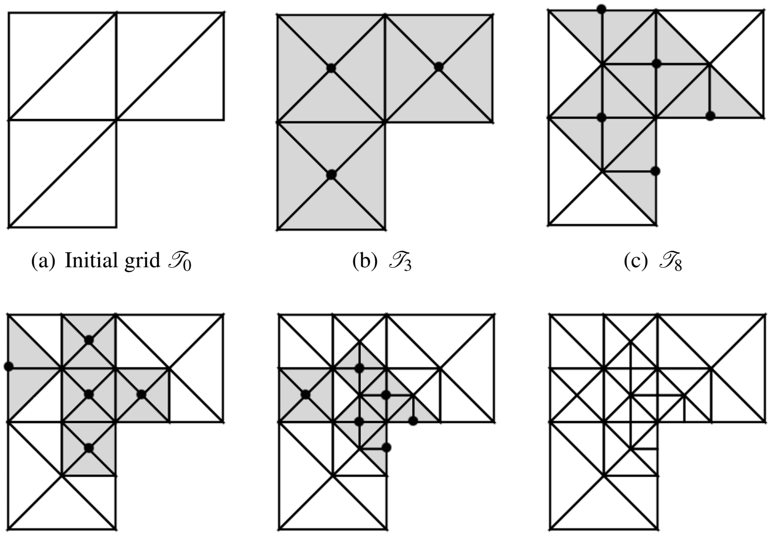
Theorem 9 (Decomposition of Bisection Grids). *Let $\mathscr{T}_0$ be a conforming triangulation. Suppose the bisection method satisfies assumptions (B2), i.e., for all $k \geq 0$ all uniform refinements $\overline{\mathscr{T}}_k$ of $\mathscr{T}_0$ are conforming. Then for any $\mathscr{T} \in \mathbb{T}(\mathscr{T}_0)$, there exists a compatible bisection sequence $\mathscr{B} = (b_1, b_2, \cdots, b_N)$ with $N = \#\mathscr{N}(\mathscr{T}) - \#\mathscr{N}(\mathscr{T}_0)$ such that*

$$\mathscr{T} = \mathscr{T}_0 + \mathscr{B}. \tag{35}$$

We use the example in Figure 3 to illustrate the decomposition of a bisection grid. In Figure 3 (a), we display the initial triangulation $\mathscr{T}_0$ which uses the longest edge as the refinement edge for each triangle. We display the fine grid $\mathscr{T} \in \mathbb{T}(\mathscr{T}_0)$ in Figure 3 (f). In Figure 3 (b)-(e), we give several intermediate triangulations during the refinement process: each triangulation is obtained by performing several compatible bisections on the previous one. Each compatible patch is indicated by a gray region and the new vertices introduced by bisections are marked by black dots. In these figures, we denoted by $\mathscr{T}_i := \mathscr{T}_0 + (b_1, b_2, \cdots, b_i)$ for $1 \leq i \leq 19$.

To prove Theorem 9, we introduce the *generation* of elements and vertices. The generation of each element in the initial grid $\mathscr{T}_0$ is defined to be 0, and the generation of a child is 1 plus that of the father. The generation of an element $\tau \in \mathscr{T} \in \mathbb{F}(\mathscr{T}_0)$ is denoted by g_τ and coincides with the number of bisections needed to create τ from $\mathscr{T}_0$. Consequently, the uniformly refined mesh $\overline{\mathscr{T}}_k$ can be characterized as the triangulation in $\mathbb{F}(\mathscr{T}_0)$ with all elements of $\overline{\mathscr{T}}_k$ of the same generation k. Vice versa, an element τ with generation k can only exist in $\overline{\mathscr{T}}_k$.

Let $\mathbb{N}(\mathscr{T}_0) = \cup\{\mathscr{N}(\mathscr{T}) : \mathscr{T} \in \mathbb{F}(\mathscr{T}_0)\}$ denote the set of all possible vertices. For any vertex $p \in \mathbb{N}(\mathscr{T}_0)$, the generation of p is defined as the minimal integer k such that $p \in \mathscr{N}(\overline{\mathscr{T}}_k)$ and is denoted by g_p. For convenience of notation, we regard g as



(a) Initial grid $\mathscr{T}_0$ (b) $\mathscr{T}_3$ (c) $\mathscr{T}_8$

(d) $\mathscr{T}_{13}$ (e) $\mathscr{T}_{19}$ (f) Fine grid $\mathscr{T} = \mathscr{T}_{19}$

Fig. 3 Decomposition of a bisection grid for $d = 2$: Each frame displays a mesh $\mathscr{T}_{i+k} = \mathscr{T}_i + \{b_{i+1}, \cdots, b_{i+k}\}$ obtained from $\mathscr{T}_i$ by a sequence of compatible bisections $\{b_j\}_{j=i+1}^{i+k}$ using the longest edge. The order of bisections is irrelevant within each frame, but matters otherwise.

either a piecewise linear function on $\mathscr{T}$ defined as $g(p) = g_p$ for $p \in \mathscr{N}(\mathscr{T})$ or a piecewise constant defined as $g(\tau) = g_\tau$ for $\tau \in \mathscr{T}$.

The following properties about the generation of elements or vertices for uniformly refined mesh $\overline{\mathscr{T}}_k$ are a consequence of the definition above:

$$\tau \in \overline{\mathscr{T}}_k \text{ if and only if } g_\tau = k; \tag{36}$$

$$p \in \mathscr{N}(\overline{\mathscr{T}}_k) \text{ if and only if } g_p \leq k; \tag{37}$$

$$\text{for } \tau \in \overline{\mathscr{T}}_k, \max_{q \in \mathscr{N}(\tau)} g_q = k = g_\tau. \tag{38}$$

Lemma 8. *Let $\mathscr{T}_0$ be a conforming triangulation. Let the bisection method satisfy assumption (B2). For any $\mathscr{T} \in \mathbb{T}(\mathscr{T}_0)$, let $p \in \mathscr{N}(\mathscr{T})$ be a vertex with maximal generation in the sense that $g_p = \max_{q \in \mathscr{N}(\mathscr{T})} g_q$. Then*

$$g_\tau = g_p \quad \textit{for all } \tau \in \mathscr{R}_p \tag{39}$$

and

$$\mathscr{R}_p = \overline{\mathscr{R}}_{k,p}, \tag{40}$$

where $k = g_p$ and $\overline{\mathscr{R}}_{k,p}$ is the first ring of p in the uniformly refined mesh $\overline{\mathscr{T}}_k$. Equivalently, all elements in $\mathscr{R}_p$ have the same generation g_p.

Proof. We prove (39) by showing $g_p \le g_\tau$ and $g_\tau \le g_p$. Since $\mathscr{T}$ is conforming, p is a vertex of each element $\tau \in \mathscr{R}_p$. This implies that $p \in \mathscr{N}(\overline{\mathscr{T}}_{g_\tau})$ and thus $g_\tau \ge g_p$ by (37). On the other hand, from (38), we have

$$g_\tau = \max_{q\in\mathscr{N}(\tau)} g_q \le \max_{q\in\mathscr{N}(\mathscr{T})} g_q = g_p, \quad \text{for all } \tau \in \mathscr{R}_p.$$

Now we prove (40). By (36), $\overline{\mathscr{R}}_{k,p}$ is made of all elements with generation k containing p. By (39), we conclude $\mathscr{R}_p \subseteq \overline{\mathscr{R}}_{k,p}$. On the other hand, p cannot belong to the domain of $\Omega\backslash\omega_p$, because of the topology of ω_p, whence $\overline{\mathscr{R}}_{k,p}\backslash\mathscr{R}_p = \varnothing$. This proves (40). □

Now we are in the position to prove Theorem 9.

Proof of Theorem 9. We prove the result by the induction on $N = \#\mathscr{N}(\mathscr{T}) - \#\mathscr{N}(\mathscr{T}_0)$. Nothing needs to be proved for $N = 0$. Assume that (35) holds for N.

Let $\mathscr{T} \in \mathbb{T}(\mathscr{T}_0)$ with $\#\mathscr{N}(\mathscr{T}) - \#\mathscr{N}(\mathscr{T}_0) = N+1$. Let $p \in \mathscr{N}(\mathscr{T})$ be a vertex with maximal generation, i.e., $g_p = \max_{q\in\mathscr{N}(\mathscr{T})} g_q$. Then by Lemma 8, we know that $\mathscr{R}_p = \overline{\mathscr{R}}_{k,p}$ for $k = g_p$. Now by assumption (B2), $\overline{\mathscr{R}}_{k,p}$ is created by a compatible bisection, say

$$b_e : \overline{\mathscr{R}}_e \to \overline{\mathscr{R}}_{k,p},$$

with $e \in \mathscr{E}(\overline{\mathscr{T}_{k-1}})$. Since the compatible bisection giving rise to p is unique within $\mathbb{F}(\mathscr{T}_0)$, it must thus be b_e. This means that if we undo the bisection operation, then we still have a conforming mesh $\mathscr{T}'$, or equivalently $\mathscr{T} = \mathscr{T}' + b_e$. We can now apply the induction assumption to $\mathscr{T}' \in \mathbb{T}(\mathscr{T}_0)$ with $\#\mathscr{N}(\mathscr{T}') - \#\mathscr{N}(\mathscr{T}_0) = N$ to finish the proof. □

4.4 Generation of Compatible Bisections

For a compatible bisection $b_i \in \mathscr{B}$, we use the same subscript i to denote related quantities such as:

- e_i: the refinement edge;
- p_i: the midpoint of e_i;
- $\widetilde{\omega}_i = \omega_{p_i} \cup \omega_{p_{l_i}} \cup \omega_{p_{r_i}}$;
- $\mathscr{T}_i = \mathscr{T}_0 + \{b_1, \cdots, b_i\}$;
- ω_i: the patch of p_i i.e. ω_{p_i};
- p_{l_i}, p_{r_i}: two end points of e_i;
- h_i: the local mesh size of ω_i;
- $\mathscr{R}_i$: the first ring of p_i in $\mathscr{T}_i$.

We understand $h \in L^\infty(\Omega)$ as a piecewise constant mesh-size function, i.e., $h_\tau = \operatorname{diam}(\tau)$ in each simplex $\tau \in \mathscr{T}$.

Lemma 9. *If $b_i \in \mathscr{B}$ is a compatible bisection, then all elements of $\mathscr{R}_i$ have the same generation g_i.*

Proof. Let $p_i \in \mathscr{N}(\mathscr{T}_0)$ be the vertex associated with b_i. Let $\overline{\mathscr{T}_k}$ be the coarsest uniformly refined mesh containing p_i, so $k = g_{p_i}$. In view of assumption (B2), p_i

arises from uniform refinement of $\overline{\mathscr{T}}_{k-1}$. Since the bisection giving rise to p_i is unique within $\mathbb{F}(\mathscr{T}_0)$, we realize that all elements in $\mathscr{R}_{e_i}$ are bisected and have generation $k-1$ because they belong to $\overline{\mathscr{T}}_{k-1}$. This implies that all elements of $\mathscr{R}_{p_i}$ have generation k, as asserted. □

This lemma allows us to introduce the concept of generation of compatible bisections. For a compatible bisection $b_i : \mathscr{R}_{e_i} \to \mathscr{R}_{p_i}$, we define $g_i = g(\tau), \tau \in \mathscr{R}_{p_i}$. Throughout this paper we always assume $h(\tau) \eqsim 1$ for $\tau \in \mathscr{T}_0$. We have the following important relation between generation and mesh size

$$h_i \eqsim \gamma^{g_i}, \text{ with } \gamma = \left(\frac{1}{2}\right)^{1/d} \in (0,1). \tag{41}$$

Beside this relation, we give now two more important properties on the generation of compatible bisections. The first property says that different bisections with the same generation have weakly disjoint local patches.

Lemma 10. *Let $\mathscr{T}_N \in \mathbb{T}(\mathscr{T}_0)$ be $\mathscr{T}_N = \mathscr{T}_0 + \mathscr{B}$, where $\mathscr{B}$ is a compatible bisection sequence $\mathscr{B} = (b_1, \cdots, b_N)$. For any $i \neq j$ and $g_j = g_i$, we have*

$$\mathring{\omega}_i \cap \mathring{\omega}_j = \varnothing. \tag{42}$$

Proof. Since $g_i = g_j = g$, both bisection patches $\mathscr{R}_i$ and $\mathscr{R}_j$ belong to the uniformly refined mesh $\overline{\mathscr{T}}_q$. If (42) were not true, then there would exists $\tau \in \mathscr{R}_i \cap \mathscr{R}_j \subset \overline{\mathscr{T}}_q$ containing distinct refinement edges e_i and e_j because $i \neq j$. This contradicts rules R2 and R3 which assign a unique refinement edge to each element. □

A simple consequence of (42) is that, for all $u \in L^2(\Omega)$ and $k \geq 1$,

$$\sum_{g_i = k} \|u\|^2_{\omega_i} \leq \|u\|^2_{\Omega}, \tag{43}$$

$$\sum_{g_i = k} \|u\|^2_{\widetilde{\omega}_i} \lesssim \|u\|^2_{\Omega}. \tag{44}$$

The second property is on the ordering of generations. For a given bisection sequence $\mathscr{B}$, we define $b_i < b_j$ if $i < j$, which means bisection b_i is performed before b_j. The generation sequence $(g_1, \cdots, g_N)$, however, is not necessary monotone increasing; there could exist $b_i < b_j$ but $g_i > g_j$. This happens for bisections driven by *a posteriori* error estimators in practice. Adaptive algorithms usually refine elements around a singularity region first, thereby creating many elements with large generations, and later they refine coarse elements away from the singularity. This mixture of generations is the main difficulty for the analysis of multilevel methods on adaptive grids. We now prove the following quasi-monotonicity property of generations restricted to a fixed bisection patch.

Lemma 11. *Let $\mathscr{T}_N \in \mathbb{T}(\mathscr{T}_0)$ be $\mathscr{T}_N = \mathscr{T}_0 + \mathscr{B}$, where $\mathscr{B}$ is a compatible bisection sequence $\mathscr{B} = (b_1, \cdots, b_N)$. For any $j > i$ and $\mathring{\widetilde{\omega}}_j \cap \mathring{\widetilde{\omega}}_i \neq \varnothing$, we have*

$$g_j \geq g_i - g_0, \tag{45}$$

where $g_0 > 0$ *is an integer depending only the shape regularity of* $\mathscr{T}_0$.

Proof. Since $\overset{\circ}{\tilde{\omega}}_j \cap \overset{\circ}{\tilde{\omega}}_i \neq \varnothing$, there must be elements $\tau_j \in \mathscr{R}_{p_j} \cup \mathscr{R}_{p_{l_j}} \cup \mathscr{R}_{p_{r_j}}$ and $\tau_i \in \mathscr{R}_{p_i} \cup \mathscr{R}_{p_{l_i}} \cup \mathscr{R}_{p_{r_i}}$ such that $\overset{\circ}{\tau}_j \cap \overset{\circ}{\tau}_i \neq \varnothing$. Since we consider triangulations in $\mathbb{T}(\mathscr{T}_0)$, the intersection $\tau_j \cap \tau_i$ is still a simplex. When b_j is performed, only τ_j exists in the current mesh. Thus $\tau_j = \tau_j \cap \tau_i \subseteq \tau_i$ and $g_{\tau_j} \geq g_{\tau_i}$.

Shape regularity implies the existence of a constant g_0 only depending on $\mathscr{T}_0$ such that

$$g_j + g_0/2 \geq g_{\tau_j} \geq g_{\tau_i} \geq g_i - g_0/2,$$

and (45) follows. □

4.5 Node-Oriented Coarsening Algorithm

A key practical issue is to find a decomposition of a bisection grid. We present a node-oriented coarsening algorithm recently developed by Chen and Zhang [29].

A crucial observation is that the inverse of a compatible bisection can be thought as a coarsening process. It is restricted to a compatible star and thus no conformity issue arises; See Figure 1. For a triangulation $\mathscr{T} \in \mathbb{T}(\mathscr{T}_0)$, a vertex p is called a *good-for-coarsening vertex*, or a *good vertex* in short, if there exist a compatible bisection b_e such that p is the middle point of e. The set of all good vertices in the grid $\mathscr{T}$ will be denoted by $G(\mathscr{T})$. By the decomposition of bisection grids (Theorem 9), the existence of good vertices is evident. Moreover, for bisection grids in 2-D, we have the following characterization of good vertices due to Chen and Zhang [29].

Theorem 10 (Coarsening). *Let* $\mathscr{T}_0$ *be a conforming triangulation. Suppose the bisection method satisfies assumptions (B2), i.e., for all* $k \geq 0$ *all uniform refinements* $\overline{\mathscr{T}}_k$ *of* $\mathscr{T}_0$ *are conforming. Then for any* $\mathscr{T} \in \mathbb{T}(\mathscr{T}_0)$ *and* $\mathscr{T} \neq \mathscr{T}_0$*, the set of good vertices* $G(\mathscr{T})$ *is not empty. Furthermore* $x \in G(\mathscr{T})$ *if and only if*

1. *it is not a vertex of the initial grid* $\mathscr{T}_0$*;*
2. *it is the newest vertex of all elements in the ring of* $\mathscr{R}_p$*.*
3. $\#\mathscr{R}_p = 4$ *for an interior vertex* x *or* $\#\mathscr{R}_p = 2$ *for a boundary vertex* p*.*

Remark 1. The assumption that $\mathscr{T}_0$ is compatible labeled could be further relaxed by using the longest edge of each triangle as its refinement edge for the initial triangulation $\mathscr{T}_0$; see Kossaczký [56].

The coarsening algorithm is simply read as the following:

```
ALGORITHM COARSEN (𝒯)
  Find all good nodes G(𝒯) of 𝒯.
  For each good node p ∈ G(𝒯)
```

```
    Replace the star R_p by b_e^{-1}(R_p).
END
```

Chen and Zhang [29] prove that one can finally obtain the initial grid back by applying the coarsening algorithm `coarsen` repeatedly. It is possible that `coarsen(T)` applied to the current grid $\mathscr{T}$ gives a coarse grid which is not in the adaptive history. Indeed our coarsening algorithm may remove vertices added in several different stages of the adaptive procedure.

For details on the implementation of this coarsening algorithm and the application to multilevel preconditioners and multigrid methods, we refer to [29] and [26].

4.6 Space Decomposition on Bisection Grids

We give a space decomposition for Lagrange finite element spaces on bisection grids. Given a conforming triangulation $\mathscr{T}$ of the domain $\Omega \subset \mathbb{R}^d$ and an integer $m \geq 1$, the mth order finite element space on $\mathscr{T}$ is defined as follows:

$$\mathscr{V}(\mathscr{P}_m, \mathscr{T}) := \{v \in H^1(\Omega) : v|_\tau \in \mathscr{P}_m(\tau) \text{ for all } \tau \in \mathscr{T}\}.$$

We restrict ourselves to bisection grids in $\mathbb{T}(\mathscr{T}_0)$ satisfying (B1) and (B2). Therefore by Theorem 9, for any $\mathscr{T}_N \in \mathbb{T}(\mathscr{T}_0)$, there exists a compatible bisection sequence $\mathscr{B} = (b_1, \cdots, b_N)$ such that

$$\mathscr{T}_N = \mathscr{T}_0 + \mathscr{B}.$$

We give a decomposition of the finite element space $\mathscr{V} := \mathscr{V}(\mathscr{P}_m, \mathscr{T}_N)$ using this decomposition of $\mathscr{T}_N$. If $\mathscr{T}_i$ is the triangulation $\mathscr{T}_0 + (b_1, \cdots, b_i)$, let $\phi_{i,p} \in \mathscr{V}(\mathscr{P}_1, \mathscr{T}_i)$ denote the linear nodal basis at a vertex $p \in \mathscr{N}(\mathscr{T}_i)$. Motivated by the stable three-point wavelet constructed by Stevenson [78], we define the sub-spaces

$$\mathscr{V}_0 = \mathscr{V}(\mathscr{P}_1, \mathscr{T}_0), \text{ and } \mathscr{V}_i = \text{span}\{\phi_{i,p_i}, \phi_{i,p_{l_i}}, \phi_{i,p_{r_i}}\}. \tag{46}$$

Since the basis functions of $\mathscr{V}_i, i = 0, \ldots, N$, are piecewise linear polynomials on $\mathscr{T}_N$, we know $\mathscr{V}_i \subseteq \mathscr{V}$. Let $\{\phi_p, p \in \Lambda\}$ be a basis of $\mathscr{V}(\mathscr{P}_m, \mathscr{T}_N)$ such that $v = \sum_{p\in\Lambda} v(p)\phi_p$ for all $v \in \mathscr{V}(\mathscr{P}_m, \mathscr{T}_N)$, where Λ is the index set of basis. For example, for quadratic element spaces, Λ consists of vertices and middle points of edges. We define $\mathscr{V}_p = \text{span}\{\phi_p\}$ and end up with the following space decomposition:

$$\mathscr{V} = \sum_{p\in\Lambda} \mathscr{V}_p + \sum_{i=0}^{N} \mathscr{V}_i. \tag{47}$$

Since $\dim \mathscr{V}_i = 3$, we have a three-point local smoother and the total computational cost for subspace correction methods based on (47) is CN. This is optimal and the constant in front of N is relatively small. In addition, the three-point local smoother simplifies the implementation of multilevel methods especially in dimensions higher than 3. For example, we only need to maintain an ordered vertex array with two parent vertices and do not need tree structure to maintain a hierarchical structure of meshes. The following result is due to Chen, Nochetto, and Xu [27].

Theorem 11 (Space Decomposition over Graded Meshes). *For any $v \in \mathscr{V}$, there exist $v_p, p \in \Lambda, v_i \in \mathscr{V}_i, i = 0, \cdots, N$ such that $v = \sum_{p\in\Lambda} v_p + \sum_{i=0}^{N} v_i$ and*

$$\sum_{p\in\Lambda} h_p^{-2}\|v_p\|^2 + \sum_{i=0}^{N} h_i^{-2}\|v_i\|^2 \lesssim \|v\|_A^2. \tag{48}$$

The idea of the proof is to use Scott-Zhang quasi-interpolation operator [75]

$$\mathscr{I}_{\mathscr{T}} : H^1(\Omega) \mapsto \mathscr{V}(\mathscr{P}_1, \mathscr{T})$$

for a conforming triangulation $\mathscr{T}$; see also Oswald [65]. For any $p \in \mathscr{N}(\mathscr{T})$ and p is an interior point, we choose a $\tau_p \subset \mathscr{R}_p$. Let $\{\lambda_{\tau_p,i}, i = 1, \cdots, d+1\}$ be the barycentric coordinates of τ which span $\mathscr{P}_1(\tau_p)$. We construct the L^2-dual basis $\Theta(\tau_p) = \{\theta_{\tau_p,i} : i = 1, \cdots, d+1\}$ of $\{\lambda_{\tau_p,i} : i = 1, \cdots, d+1\}$. Suppose $\theta_p \in \Theta(\tau_p)$ is the dual basis such that $\int_{\tau_p} \theta_p v \,\mathrm{d}x = v(p)$, for all $v \in \mathscr{P}_1(\tau_p)$. We then define

$$\mathscr{I}_{\mathscr{T}} v = \sum_{p\in\mathscr{N}(\mathscr{T})} \Big(\int_{\tau_p} \theta_p v \,\mathrm{d}x \Big) \phi_p.$$

For boundary vertex p, we simply define $\mathscr{I}_{\mathscr{T}} v(p) = 0$ to reflect the vanishing boundary condition of v. By definition, $\mathscr{I}_{\mathscr{T}}$ preserves piecewise linear functions and satisfies the following estimate and stability [75, 65]

$$|\mathscr{I}_{\mathscr{T}} v|_1 + \|h^{-1}(v - \mathscr{I}_{\mathscr{T}} v)\| \lesssim |v|_1, \tag{49}$$

$$h_i^{d-2} |\mathscr{I}_{\mathscr{T}} v(p_i)|^2 \lesssim h_i^{-2} \|v\|_{\tau_{p_i}}, \tag{50}$$

where h_i is the size of τ_{p_i}.

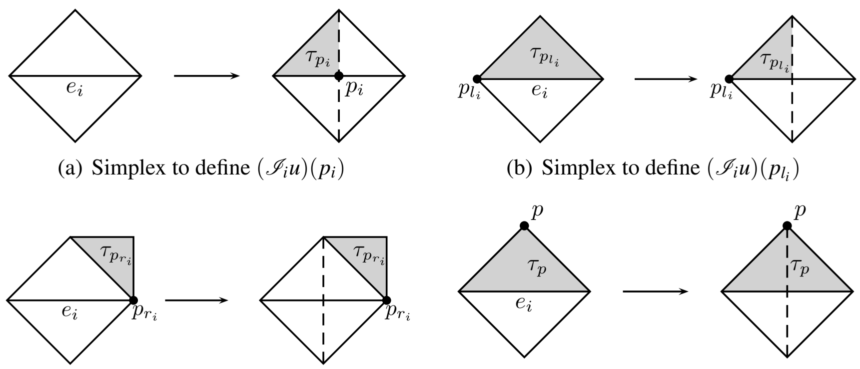
Given $v \in \mathscr{V}(\mathscr{P}_m, \mathscr{T})$, we define $u = \mathscr{I}_{\mathscr{T}} v$ and a decomposition $v = u + (v - u)$, where $\mathscr{I}_{\mathscr{T}} : \mathscr{V}(\mathscr{P}_m, \mathscr{T}) \to \mathscr{V}(\mathscr{P}_1, \mathscr{T})$. We first give a multilevel decomposition of u using quasi-interpolation. For a vertex p, we denote by τ_p the simplex used to define the nodal value at p. The following construction of a sequence of quasi-interpolations will update τ_p carefully.

Let $\mathscr{I}_0$ be a quasi-interpolation operator defined $\mathscr{V}(\mathscr{P}_1, \mathscr{T}) \to \mathscr{V}_0$. Suppose $\mathscr{I}_{i-1}$ is defined on $\mathscr{V}(\mathscr{P}_1, \mathscr{T}_{i-1})$. After the compatible bisection b_i, we define the nodal values at the new added vertex p_i using a simplex introduced by the bisection, i.e. $\tau_{p_i} \subset \omega_i$. For other vertices p, let $\tau_p \in \mathscr{T}_{i-1}$ be the simplex used to define $(\mathscr{I}_{i-1} u)(p)$, we define $(\mathscr{I}_i u)(p)$ according to the following two cases:

1. if $\tau_p \subset \omega_p(\mathscr{T}_i)$ we keep the nodal value, i.e., $(\mathscr{I}_i u)(p) = (\mathscr{I}_{i-1} u)(p)$;
2. otherwise we choose a new $\tau_p \subset \omega_p(\mathscr{T}_i) \cap \omega_p(\mathscr{T}_{i-1})$ to define $(\mathscr{I}_i u)(p)$.

In either case, we ensure that the simplex $\tau_p \subset \omega_p(\mathscr{T}_i)$.

An important property of the bisection is that b_i only changes the local patches of two end points of the refinement edge e_i going from $\mathscr{T}_{i-1}$ to $\mathscr{T}_i$. The construction in the second case is thus well defined. By construction $(\mathscr{I}_i - \mathscr{I}_{i-1})u(p) = 0$ for $p \in \mathscr{N}(\mathscr{T}_i), p \neq p_i, p_{l_i}$ or p_{r_i}, which implies $(\mathscr{I}_i - \mathscr{I}_{i-1})u \in \mathscr{V}_i$. Furthermore a close look reveals that if $(\mathscr{I}_i - \mathscr{I}_{i-1})u(p) \neq 0$, then the elements τ_p used to define $\mathscr{I}_i(p)$ or $\mathscr{I}_{i-1}(p)$ are inside the patch ω_i; see Figure 4.



(a) Simplex to define $(\mathscr{I}_i u)(p_i)$ (b) Simplex to define $(\mathscr{I}_i u)(p_{l_i})$

(c) Simplex to define $(\mathscr{I}_i u)(p_{r_i})$ (d) Simplex to define $(\mathscr{I}_i u)(p)$

Fig. 4 Update of nodal values $\mathscr{I}_i u$ to yield $\mathscr{I}_{i-1}u$: the element τ chosen to perform the averaging that gives $(\mathscr{I}_i u)(p)$ must belong to $\omega_p(\mathscr{T}_i)$. This implies $(\mathscr{I}_i - \mathscr{I}_{i-1})u(p) \neq 0$ possibly for $p = p_i, p_{l_i}, p_{r_i}$ and $= 0$ otherwise.

In this way, we obtain a sequence of quasi-interpolation operators

$$\mathscr{I}_i : \mathscr{V}(\mathscr{P}_1, \mathscr{T}_N) \to \mathscr{V}(\mathscr{P}_1, \mathscr{T}_i), i = 0 : N.$$

We define $v_i = (\mathscr{I}_i - \mathscr{I}_{i-1})u \in \mathbb{V}_i$ for $i = 1 : N$. In general $\mathscr{I}_N u \neq u$ since the simplex used to define nodal values of $\mathscr{I}_N u$ may not be in the finest mesh $\mathscr{T}_N$ but in $\mathscr{T}_{N-1}$. Nevertheless, the difference $v - \mathscr{I}_N u$ is of high frequency in the finest mesh.

Let $v - \mathscr{I}_N u = \sum_{p \in \Lambda} v_p$ be the basis decomposition. We then obtain a decomposition

$$v = \sum_{p \in \Lambda} v_p + \sum_{i=0}^{N} v_i, \quad v_i \in \mathbb{V}_i, \tag{51}$$

where for convenience we define $\mathscr{I}_{-1} := 0$.

To prove that the decomposition (51) is stable we first study $\sum_{p \in \Lambda} v_p$. Let τ_p be the simplex used to define $\mathscr{I}_N u(p)$ for $p \in \mathscr{N}(\mathscr{T}_N)$. By construction, although τ_p may not be a simplex in the triangulation $\mathscr{T}_N$, it is still in the patch $\omega_p(\mathscr{T}_N)$. Then by (49)

$$\sum_{p \in \Lambda} h_p^{-2} \|v_p\|^2 \lesssim \|h^{-1}(v - Q_N v)\|^2 \lesssim |v|_1^2. \tag{52}$$

We next prove that the decomposition $\mathscr{I}_N u = \sum_{i=0}^{N}(\mathscr{I}_i - \mathscr{I}_{i-1})u$ is stable. For this purpose, we need the auxiliary decomposition on the uniform refinement. We choose minimal L such that $\mathscr{V} \subseteq \overline{\mathscr{V}}_L$. By Lemma 2, we have a stable decomposition $u = \sum_{k=0}^{L} \bar{v}_k$, with $\bar{v}_k = (\overline{Q}_k - \overline{Q}_{k-1})u, k = 0, \cdots, L$.

We apply the slicing operator $\mathscr{I}_i - \mathscr{I}_{i-1}$ to this decomposition. When $k \le g_i - 1$, $\bar{v}_k$ is piecewise linear in ω_{e_i}, $(\mathscr{I}_i - \mathscr{I}_{i-1})\bar{v}_k = 0$ since $\mathscr{I}_i$ preserves piecewise linear functions. So the slicing operator detects frequencies higher than or equal to the generation of bisection, namely

$$v_i = (\mathscr{I}_i - \mathscr{I}_{i-1}) \sum_{l=g_i}^{L} \bar{v}_l. \tag{53}$$

By construction of v_i and the stability of quasi-interpolation, we conclude

$$\|v_i\|_{\widetilde{\omega}_i}^2 \lesssim h_i^{2+d}\left[v_i(p_i)^2 + v_i(p_{l_i})^2 + v_i(p_{r_i})^2\right] \lesssim \Big\| \sum_{l=g_i}^{L} \bar{v}_l \Big\|_{\omega_i}^2.$$

In the last step, the domain is changed to ω_i since the simplexes used to define nonzero values of $v_i(p_i), v_i(p_{l_i r})$ or $v_i(p_{l_i r})$ are inside ω_i.

Note that for different bisections with the same generation, their local patches are weakly disjoint (Lemma 10): for any $i \neq j$ and $g_j = g_i$, we have

$$\mathring{\omega}_i \cap \mathring{\omega}_j = \varnothing. \tag{54}$$

Consequently

$$\sum_{g_i=k} \|v_i\|^2 = \sum_{g_i=k} \|v_i\|_{\widetilde{\omega}_i}^2 \lesssim \sum_{g_i=k} \Big\| \sum_{l=g_i}^{L} \bar{v}_l \Big\|_{\omega_i}^2 \lesssim \Big\| \sum_{l=g_i}^{L} \bar{v}_l \Big\|_{\Omega}^2 = \sum_{l=k}^{L} \|\bar{v}_l\|^2.$$

In the last step, we use the fact $\bar{v}_k$ are L^2-orthogonal decomposition.

The following elementary result will be useful and can be found in [32].

Lemma 12 (Discrete Hardy Inequality). *If the sequences $\{a_k\}_{k=0}^{L}, \{b_k\}_{k=0}^{L}$ satisfy*

$$b_k \le \sum_{l=k}^{L} a_l, \quad \textit{for all } k \ge 0$$

and are non-negative, then for any $s \in (0,1)$, we have

$$\sum_{k=0}^{L} s^{-k} b_k \le \frac{1}{1-s} \sum_{k=0}^{L} s^{-k} a_k.$$

Proof. Since

$$\sum_{k=0}^{L} s^{-k} b_k \le \sum_{k=0}^{L}\sum_{l=k}^{L} s^{-k} a_l = \sum_{l=0}^{L}\sum_{k=0}^{l} s^{-k} a_l = \sum_{l=0}^{L} s^{-l} a_l \sum_{k=0}^{l} s^{l-k},$$

and $s < 1$, the geometric series is bounded by $1/(1-s)$ and concludes the proof. □

Applying Lemma 12 to $a_k = \|\bar{v}_k\|^2$ and $b_k = \sum_{g_i=k} \|v_i\|^2$, we obtain

$$\sum_{k=0}^{L} \bar{h}_k^{-2} \sum_{g_i=k} \|v_i\|^2 \lesssim \sum_{k=0}^{L} \bar{h}_k^{-2} \|\bar{v}_k\|^2,$$

and thus from the stable decomposition corresponding to uniform refinement, we conclude

$$\sum_{i=0}^{N} h_i^{-2} \|v_i\|^2 = \sum_{k=0}^{L} \bar{h}_k^{-2} \sum_{g_i=k} \|v_i\|^2 \lesssim \sum_{k=0}^{L} \bar{h}_k^{-2} \|\bar{v}_k\|^2 \lesssim |\mathscr{I}_{\mathscr{T}} v|_1^2 \lesssim |v|_1^2. \qquad (55)$$

4.7 Strengthened Cauchy-Schwarz Inequality

In this section we establish the SCS inequality for the space decomposition $\sum_{i=0}^{N} \mathscr{V}_i$.

Theorem 12. *For any $u_i, v_i \in \mathscr{V}_i, i = 0, \cdots, N$, we have*

$$\Big|\sum_{i=0}^{N}\sum_{j=i+1}^{N} (u_i, v_j)_A\Big| \lesssim \left(\sum_{i=0}^{N} \|u_i\|_A^2\right)^{1/2} \left(\sum_{i=0}^{N} h_i^{-2} \|v_i\|^2\right)^{1/2}. \qquad (56)$$

Proof. The proof consists of several careful summations using the concept of generation to relate with uniform refinements. The proof is divided into four steps.

[1] For a fixed index $i \in [1, N]$, we denote by

$$n(i) = \{j > i : \widetilde{\omega}_j \cap \widetilde{\omega}_i \neq \varnothing\} \text{ and } w_k^i = \sum_{j \in n(i), g_j = k} v_j.$$

Shape regularity implies that $w_k^i \in \overline{\mathscr{V}}_{k+g_0}$ and $k = g_j \ge g_i - g_0$ (Lemma 11). For any $\tau \in \tilde{\omega}_i$, we apply the SCS inequality of Lemma 4 over τ to u_i and w_k^i and obtain

$$(u_i, w_k^i)_{A,\tau} \lesssim \gamma^{k+g_0-g_i} \|u_i\|_{A,\tau} \bar{h}_{k+g_0}^{-1} \|w_k^i\|_\tau \lesssim \gamma^{k-g_i} \|u_i\|_{A,\tau} \bar{h}_k^{-1} \|w_k^i\|_\tau.$$

Then

$$
\begin{aligned}
(u_i, w_k^i)_{A,\tilde{\omega}_i} &= \sum_{\tau \subset \tilde{\omega}_i} (u_i, w_k^i)_{A,\tau} \\
&\lesssim \gamma^{k-g_i} \sum_{\tau \subset \tilde{\omega}_i} \|u_i\|_{A,\tau} \bar{h}_k^{-1} \|w_k^i\|_\tau \\
&\lesssim \gamma^{k-g_i} \|u_i\|_{A,\tilde{\omega}_i} \bar{h}_k^{-1} \Big(\sum_{\tau \subset \tilde{\omega}_i} \|w_k^i\|_\tau^2 \Big)^{1/2}.
\end{aligned}
$$

Since v_j's with the same generation $g_j = k$ have supports with finite overlap, we infer that $\|w_k^i\|_\tau^2 \lesssim \sum_{j \in n(i), g_j = k} \|v_j\|_\tau^2 \le \sum_{g_j = k} \|v_j\|_\tau^2$ and

$$
(u_i, w_k^i)_{A,\tilde{\omega}_i} \lesssim \gamma^{k-g_i} \|u_i\|_{A,\tilde{\omega}_i} \bar{h}_k^{-1} \Big(\sum_{g_j = k} \|v_j\|_{0,\tilde{\omega}_i}^2 \Big)^{1/2}.
$$

$\boxed{2}$ We fix u_i and consider

$$
|(u_i, \sum_{j=i+1}^{N} v_j)_A| = |(u_i, \sum_{j \in n(i)} v_j)_{A,\tilde{\omega}_i}| = |(u_i, \sum_{k=(g_i-g_0)^+}^{L} \sum_{j \in n(i), g_j = k} v_j)_{A,\tilde{\omega}_i}|,
$$

because $w_k^j = 0$ for $k < g_i - g_0$ (Lemma 11). Since $k \ge 0$, this is equivalent to $k \ge (g_i - g_0)^+ := \max\{g_i - g_0, 0\}$, whence

$$
\begin{aligned}
|(u_i, \sum_{j=i+1}^{N} v_j)_A| &\lesssim \sum_{k=(g_i-g_0)^+}^{L} |(u_i, w_k^i)_{A,\tilde{\omega}_i}| \\
&\lesssim \sum_{k=(g_i-g_0)^+}^{L} \gamma^{k-g_i} \|u_i\|_{A,\tilde{\omega}_i} \bar{h}_k^{-1} \Big(\sum_{g_j = k} \|v_j\|_{0,\tilde{\omega}_i}^2 \Big)^{1/2}.
\end{aligned}
$$

$\boxed{3}$ We now sum over i but keeping the generation $g_i = l \ge 0$ fixed:

$$
\begin{aligned}
\sum_{g_i = l} |(u_i, \sum_{j=i+1}^{N} v_j)_A| &\lesssim \sum_{k=(l-g_0)^+}^{L} \gamma^{k-l} \left\{ \sum_{g_i = l} \left[\|u_i\|_{A,\tilde{\omega}_i} \left(\bar{h}_k^{-2} \sum_{g_j = k} \|v_j\|_{\tilde{\omega}_i}^2 \right)^{1/2} \right] \right\} \\
&\lesssim \sum_{k=(l-g_0)^+}^{L} \gamma^{k-l} \left(\sum_{g_i = l} \|u_i\|_{A,\tilde{\omega}_i}^2 \right)^{1/2} \left(\bar{h}_k^{-2} \sum_{g_i = l} \sum_{g_j = k} \|v_j\|_{\tilde{\omega}_i}^2 \right)^{1/2}.
\end{aligned}
$$

In view of the finite overlap of patches $\tilde{\omega}_i$ for generation $g_i = l$ (see (44)), we deduce

$$
\sum_{g_i = l} |(u_i, \sum_{j=i+1}^{N} v_j)_A| \lesssim \sum_{k=(l-g_0)^+}^{L} \gamma^{k-l} \left(\sum_{g_i = l} \|u_i\|_{A,\tilde{\omega}_i}^2 \right)^{1/2} \left(\bar{h}_k^{-2} \sum_{g_j = k} \|v_j\|^2 \right)^{1/2}.
$$

$\boxed{4}$. We finally sum over all generations $0 \le l \le L$ to get

$$\sum_{l=0}^{L}\sum_{g_i=l}|(u_i,\sum_{j=i+1}^{N}v_j)_A| \lesssim \sum_{l=0}^{L}\sum_{k=(l-g_0)^+}^{L}\gamma^{k-l}\left(\sum_{g_i=l}\|u_i\|_{A,\tilde{\omega}_i}^2\right)^{1/2}\left(\bar{h}_k^{-2}\sum_{g_j=k}\|v_j\|^2\right)^{1/2}$$
$$\lesssim \left(\sum_{l=0}^{L}\sum_{g_i=l}\|u_i\|_{A,\tilde{\omega}_i}^2\right)^{1/2}\left(\sum_{k=0}^{L}\bar{h}_k^{-2}\sum_{g_j=k}\|v_j\|^2\right)^{1/2}.$$

where we have applied Lemma 5. Therefore, since $\sum_{i=0}^{N}=\sum_{l=0}^{L}\sum_{g_i=l}$ and $\bar{h}_k=h_j$ for $k=g_j$, we end up with the desired estimate (56). □

4.8 BPX Preconditioner and Multigrid on Graded Bisection Grids

Proceeding as in Section §3, with quasi-uniform grids created by uniform refinement, we can obtain the optimality of BPX preconditioner and optimal convergent rate of V-cycle multigrid. We state the results below and refer to [27] for proofs.

Theorem 13 (Optimality of BPX on Graded Bisection Grids). *For the BPX preconditioner based on the space decomposition (47)*

$$Bu=\sum_{p\in\Lambda}h_p^{2-d}(u,\phi_p)\phi_p+\sum_{i=1}^{N}h_i^{2-d}[(u,\phi_{p_i})\phi_{p_i}+(u,\phi_{p_{l_i}})\phi_{p_{l_i}}+(u,\phi_{p_{r_i}})\phi_{p_{r_i}}],$$

we have

$$\kappa(BA)\lesssim 1.$$

A V-cycle type multigrid method can be obtained by applying SSC to the space decomposition (47). A symmetric V-cycle loop is like

1. pre-smoothing (forward Gauss-Seidel) in the finest space $\mathscr{V}(\mathscr{P}_m,\mathscr{T}_N)$;
2. multilevel smoothing in *linear* finite element spaces $\mathscr{V}_i$ for $i=N$ to 1;
3. exact solver in the coarsest linear finite element spaces $\mathscr{V}_0$;
4. multilevel smoothing in *linear* finite element spaces $\mathscr{V}_i$ for $i=1$ to N;
5. post-smoothing (backward Gauss-Seidel) in the finest space $\mathscr{V}(\mathscr{P}_m,\mathscr{T}_N)$.

Theorem 14 (Uniform Convergence of V-cycle Multigrid on Graded Bisection Grids). *The above V-cycle multigrid, namely SSC based on the space decomposition (47), is uniformly convergent.*

5 Multilevel Methods for H(curl) and H(div) Systems

In this section, we design and analyze multigrid methods for solving finite element discretization of $H(\mathrm{curl})$ and $H(\mathrm{div})$ systems

$$\operatorname{curl} \times \operatorname{curl} \times u + u = f, \qquad \text{in } \Omega, \tag{57}$$

$$-\operatorname{grad} \operatorname{div} u + u = f, \qquad \text{in } \Omega, \tag{58}$$

with homogeneous Neumann boundary condition. Here $\Omega \subset \mathbb{R}^3$ is a simply connected and bounded polyhedron. We study edge elements for (57) and face elements for (58) over shape regular tetrahedra triangulations $\mathscr{T}$ of Ω.

Standard multigrid methods developed for H^1 problem, i.e.,

$$-\Delta u + u = -\operatorname{div} \operatorname{grad} u + u = f$$

cannot be transferred to the $H(\text{curl})$ and $H(\text{div})$ systems directly. The reason is that for vector fields, the operators $\operatorname{curl} \times \operatorname{curl}$ and $-\operatorname{grad} \operatorname{div}$ are only part of the Laplace operator because

$$-\Delta := \operatorname{curl} \times \operatorname{curl} - \operatorname{grad} \operatorname{div}.$$

Therefore in the divergence free space, the operator $\operatorname{curl} \times \operatorname{curl} + I$ behaves like $-\Delta + I$, while in the kernel space of the curl operator, the space of gradients, it is like I. Similarly, the operator $-\operatorname{grad} \operatorname{div} + I$ behaves like $-\Delta + I$ on gradients and I on curls. Efficient solvers should account for the different behavior of curl and div in their kernel and orthogonal complement. In particular, the smoother in the kernel space is critical. We note that for the grad operator, the kernel space is a one dimensional (constant) space, while for the curl and div operators, the kernel space is infinite dimensional. The decomposition of spaces used in multigrid methods should satisfy certain properties (see [57] and [98]). One approach is to perform a smoothing in the kernel space which can be expressed explicitly using properties of exact sequences between finite element spaces of $H^1, H(\text{curl})$ and $H(\text{div})$ systems. This is used by Hiptmair to obtain the first results for multigrid of $H(\text{div})$ [47] and $H(\text{curl})$ [49] systems in three dimensions. See also Hiptmair and Toselli [51] for a unified and simplified treatment. Another important approach taken by Arnold, Falk and Winther in [3, 4] is to perform the smoothing on patches of vertices which contain a basis of the kernel space of curl and div operator. In [3, 4], the analysis hinges on the following two assumptions:

- Ω is a bounded and *convex* polyhedron in $\mathbb{R}^3$;
- $\mathscr{T}$ is a shape regular and quasi-uniform mesh of Ω.

The first assumption is used in duality arguments which require full regularity of the elliptic equations, whereas the second one is used to prove certain approximation properties. We regard both items as regularity assumptions, first on the solutions of the elliptic equation and second on the underlying mesh.

In practice, most problems are posed on non-convex domains and thus solutions exhibit singularities. Finite element approximations based on quasi-uniform grids cannot deliver optimal rates due to lack of regularity. Mesh refinements restore optimal convergence rates in terms of degree of freedoms, but make the above regularity assumptions inadequate for studying adaptive finite element methods for $H(\text{curl})$ and $H(\text{div})$ systems.

We will design multilevel methods for these systems on graded grids obtained by bisection. In the analysis, we relax the regularity assumptions used in the previous work [47, 49, 3, 4] by using two new techniques developed recently in [52] and [27]. More precisely, we employ

- Discrete regular decompositions of finite element spaces [52] to relax the regularity assumption on the solution;
- Decomposition of bisection grids and corresponding space decompositions [27], already discussed in section §4, to relax the regularity assumption on the grids.

We should mention that a local multigrid method similar to ours for $H(\mathrm{curl})$ system on adaptive grids has been independently developed by Hiptmair and Zheng [53]. We follow closely our recent work [28] to present a unified treatment for both $H(\mathrm{curl})$ and $H(\mathrm{div})$ systems.

To focus on the two aforementioned issues, we consider the simplest scenario, that is we do not include Dirichlet type boundary conditions for (57) or (58) nor variable coefficients. We note that results in [4] hold uniformly for variable coefficients and results in this paper extend to this case as well.

5.1 Preliminaries

5.1.1 Sobolev Spaces and Finite Element Spaces

Let $\Omega \subset \mathbb{R}^3$ be a bounded domain which is homeomorphic to a ball. We define the following Sobolev spaces

$$\begin{aligned} H(\mathrm{grad};\Omega) &= \{v \in L^2(\Omega) : \mathrm{grad}\, v \in L^2(\Omega)\} = H^1(\Omega),\\ H(\mathrm{curl};\Omega) &= \{v \in (L^2(\Omega))^3 : \mathrm{curl}\, v \in (L^2(\Omega))^3\},\\ H(\mathrm{div};\Omega) &= \{v \in (L^2(\Omega))^3 : \mathrm{div}\, v \in (L^2(\Omega))^3\}. \end{aligned}$$

We use a generic notation $H(\mathscr{D},\Omega)$ to refer to $H(\mathrm{grad};\Omega), H(\mathrm{curl};\Omega)$ or $H(\mathrm{div};\Omega)$, where $\mathscr{D} = \mathrm{grad}, \mathrm{curl}$ or div represents differential operators according to the context. Since $\mathrm{curl}\, v$ and $\mathrm{div}\, v$ are special combinations of components of $\mathrm{grad}\, v$, in general $H^1(\Omega) \subset H(\mathscr{D},\Omega)$.

Let $(\cdot,\cdot)$ denote the inner product for $L^2(\Omega)$ or $[L^2(\Omega)]^3$. As subspaces of $[L^2(\Omega)]^3$, $H(\mathrm{grad};\Omega)$, $H(\mathrm{curl};\Omega)$, and $H(\mathrm{div};\Omega)$ are endowed with $(\cdot,\cdot)$ as their default inner product. We assign new inner products using differential operator $\mathscr{D}$ to these spaces:

$$\begin{aligned} H(\mathrm{grad};\Omega): &\quad (u,v)_{A^g} := (u,v) + (\mathrm{grad}\, u, \mathrm{grad}\, v),\\ H(\mathrm{curl};\Omega): &\quad (u,v)_{A^c} := (u,v) + (\mathrm{curl}\, u, \mathrm{curl}\, v),\\ H(\mathrm{div};\Omega): &\quad (u,v)_{A^d} := (u,v) + (\mathrm{div}\, u, \mathrm{div}\, v). \end{aligned}$$

The corresponding norm are denoted by $\|\cdot\|_{A^g}$, $\|\cdot\|_{A^c}$ and $\|\cdot\|_{A^d}$, respectively.

These inner products introduce corresponding symmetric positive definite operators (with respect to the default $(\cdot,\cdot)$ inner product).

$$\begin{aligned} A^g: H(\text{grad};\Omega) \to H(\text{grad};\Omega)^* \qquad & (A^g u, v) := (u,v)_{A^g}, \\ A^c: H(\text{curl};\Omega) \to H(\text{curl};\Omega)^* \qquad & (A^c u, v) := (u,v)_{A^c}, \\ A^d: H(\text{div};\Omega) \to H(\text{div};\Omega)^* \qquad & (A^d u, v) := (u,v)_{A^d}. \end{aligned}$$

We focus on the $H(\text{curl})$ and $H(\text{div})$ systems, namely,

$$A^c u = \text{curl}\times\text{curl}\, u + u = f, \tag{59}$$

$$A^d u = -\,\text{grad}\,\text{div}\, u + u = f, \tag{60}$$

with homogeneous Neumann boundary condition. We build on the study of the H^1 problem, $A^g u = f$, in previous sections.

Given a shape regular triangulation $\mathscr{T}$ of Ω and integer $k \ge 1$, we define the following finite element spaces:

$$\begin{aligned} \mathscr{V}(\text{grad},\mathscr{P}_k,\mathscr{T}) &:= \{v \in H(\text{grad};\Omega) : v|_\tau \in \mathscr{P}_k(\tau),\ \forall \tau \in \mathscr{T}\}, \\ \mathscr{V}(\text{curl},\mathscr{P}_k^-,\mathscr{T}) &:= \{v \in H(\text{curl};\Omega) : v|_\tau \in \mathscr{P}_{k-1}^3(\tau) + \mathscr{P}_{k-1}^3(\tau)\times x,\ \forall \tau \in \mathscr{T}\}, \\ \mathscr{V}(\text{curl},\mathscr{P}_k,\mathscr{T}) &:= \{v \in H(\text{curl};\Omega) : v|_\tau \in \mathscr{P}_k^3(\tau),\ \forall \tau \in \mathscr{T}\}, \\ \mathscr{V}(\text{div},\mathscr{P}_k^-,\mathscr{T}) &:= \{v \in H(\text{div};\Omega) : v|_\tau \in \mathscr{P}_{k-1}^3(\tau) + \mathscr{P}_{k-1}(\tau)x,\ \forall \tau \in \mathscr{T}\} \\ \mathscr{V}(\text{div},\mathscr{P}_k,\mathscr{T}) &:= \{v \in H(\text{div};\Omega) : v|_\tau \in \mathscr{P}_k^3(\tau),\ \forall \tau \in \mathscr{T}\} \\ \mathscr{V}(L^2,\mathscr{P}_{k-1},\mathscr{T}) &:= \{v \in L^2(\Omega) : v|_\tau \in \mathscr{P}_{k-1}(\tau),\ \forall \tau \in \mathscr{T}\}. \end{aligned}$$

As in [5], the notation $\mathscr{P}_k^-$ indicates that the polynomial space is a proper subspace of $\mathscr{P}_k$. When we do not refer to a specific finite element space, we use the generic notation $\mathscr{V}(\mathscr{D},\mathscr{T})$. In particular, we simply denote by $\mathscr{V} = \mathscr{V}(\text{grad},\mathscr{P}_1,\mathscr{T})$ the continuous piecewise linear finite element space.

The degrees of freedom of these finite element spaces, and their unisolvency, are not easy to sketch here. We refer to [2, 5, 48, 50] for a unified presentation using differential forms.

Since $\mathscr{V}(\mathscr{D},\mathscr{T}) \subset H(\mathscr{D};\Omega)$, the operator equations (59) or (60) can be restricted to the finite element spaces $\mathscr{V}(\text{curl},\mathscr{T})$ or $\mathscr{V}(\text{div},\mathscr{T})$. Existence and uniqueness of the ensuing discrete problems follow from the Riesz representation theorem. Our task is to develop fast solvers for these linear algebraic systems over graded bisection grids as well as unstructured grids $\mathscr{T}$.

5.1.2 Exact Sequences and Commutative Diagram

The following exact sequence, called de Rham differential complex, plays an important role in the error analysis of finite element approximations as well as the iteration methods for solving the algebraic systems:

$$\mathbb{R} \hookrightarrow H^1(\Omega) \xrightarrow{\text{grad}} H(\text{curl};\Omega) \xrightarrow{\text{curl}} H(\text{div};\Omega) \xrightarrow{\text{div}} L^2(\Omega). \tag{61}$$

For a differential operator $\mathscr{D}$, we denote by $\mathscr{D}^-$ the previous one in the exact sequence: if $\mathscr{D} = \text{curl}$, then $\mathscr{D}^- = \text{grad}$, and if $\mathscr{D} = \text{div}$, then $\mathscr{D}^- = \text{curl}$. The following crucial properties of (61) are valid:

$$\ker(\text{grad}) = \mathbb{R}, \quad \ker(\text{curl}) = \text{img}(\text{grad}), \quad \ker(\text{div}) = \text{img}(\text{curl}). \tag{62}$$

We now state two results, Theorem 15 for $\mathscr{D} = \text{curl}$ and Theorem 16 for $\mathscr{D} = \text{div}$, which make this precise. We refer to Girault-Raviart [39] for Theorem 15.

Theorem 15 (Irrotational Fields). *Let Ω be a bounded, simply connected Lipschitz domain in $\mathbb{R}^3$ and suppose $u \in [L^2(\Omega)]^3$. Then $\text{curl}\, u = 0$ in Ω if and only if there exists a scalar potential $\phi \in H^1(\Omega)$ such that $u = \text{grad}\,\phi$ and*

$$\|\phi\|_1 \lesssim \|u\|. \tag{63}$$

To verify that $\ker(\text{div}) = \text{img}(\text{curl})$, we first present a result in $\mathbb{R}^3$.

Lemma 13. *Let $N(\text{div};\mathbb{R}^3) = \{v \in H(\text{div};\mathbb{R}^3) : \text{div}\, v = 0\}$ be the kernel of operator* div. *Then for any $u \in N(\text{div};\mathbb{R}^3)$ there exists $\phi \in [H^1_{\text{loc}}(\mathbb{R}^3)]^3$ such that*

$$\text{curl}\,\phi = u, \quad \text{div}\,\phi = 0, \quad \|\phi\|_{1,\text{loc},\mathbb{R}^3} \lesssim \|u\|_{0,\mathbb{R}^3}. \tag{64}$$

Proof. In terms of Fourier transform, the conditions $u = \text{curl}\,\phi$ and $\text{div}\,\phi = 0$ become

$$\hat{u} = 2\pi i\, \xi \times \hat{\phi} = 2\pi i(\xi_2\hat{\phi}_3 - \xi_3\hat{\phi}_2, \xi_3\hat{\phi}_1 - \xi_1\hat{\phi}_3, \xi_1\hat{\phi}_2 - \xi_2\hat{\phi}_1),$$
$$\xi \cdot \hat{\phi} = \sum_{j=1}^{3} \xi_j \hat{\phi}_j = 0,$$

respectively. We observe that the first relation implies

$$\xi \cdot \hat{u} = \sum_{j=1}^{3} \xi_j \hat{u}_j = 0,$$

or equivalently $\text{div}\, u = 0$. Computing $\hat{u} \times \xi$ and using the first two relations gives $\hat{\phi}$ uniquely as follows:

$$\hat{\phi} = \frac{1}{2\pi i|\xi|^2}\hat{u} \times \xi = \frac{1}{2\pi i|\xi|^2}(\xi_3\hat{u}_2 - \xi_2\hat{u}_3, \xi_1\hat{u}_3 - \xi_3\hat{u}_1, \xi_2\hat{u}_1 - \xi_1\hat{u}_2).$$

The desirable ϕ is the inverse Fourier transform of $\hat{\phi}$. In addition, we have

$$|\xi_j\phi_i| \le \sum_{i=1}^{3} |\hat{u}_i|.$$

Parseval's identity shows that $\phi \in H^1_{\text{loc}}(\mathbb{R}^3)$. $\square$

Theorem 16 (Solenoidal Fields). *Let Ω be a simply connected bounded domain. For any function $u \in H(\text{div};\Omega)$ such that $\operatorname{div} u = 0$, there exists a vector field $\phi \in [H^1(\Omega)]^3$ such that $u = \operatorname{curl}\phi$ and $\operatorname{div}\phi = 0$ in Ω and*

$$\|\phi\|_{H^1(\Omega)} \lesssim \|u\|_{L^2(\Omega)}.$$

Proof. We first construct an extension of u to $N(\text{div};\mathbb{R}^3)$. Let $\mathscr{O}$ be a smooth domain containing Ω. We let $p \in H^1(\mathscr{O}\backslash\Omega)/\mathbb{R}$ satisfy

$$-\Delta p = 0 \text{ in } \mathscr{O}\backslash\Omega,$$
$$\frac{\partial p}{\partial n} = u \cdot n \text{ on } \partial\Omega, \quad \frac{\partial p}{\partial n} = 0 \text{ on } \partial\mathscr{O}.$$

This solution exists since $\langle u \cdot n, 1\rangle_{\partial\Omega} = \int_\Omega \operatorname{div} u \, \mathrm{dx} = 0$. We define $\tilde{u} \in L^2(\mathbb{R}^3)$ by

$$\tilde{u} = \begin{cases} u & \text{in } \Omega, \\ \operatorname{grad} p & \text{in } \mathscr{O}\backslash\Omega, \\ 0 & \text{in } \mathbb{R}^3\backslash\mathscr{O}. \end{cases}$$

Since $\operatorname{div}\tilde{u} = 0$ in Ω and $\mathscr{O}\backslash\Omega$ and the normal component of $\tilde{u}$ is continuous across the common boundary $\partial\Omega$, we conclude $\tilde{u} \in H(\text{div};\mathbb{R}^3)$ and $\operatorname{div}\tilde{u} = 0$.

We then apply Lemma 13 to get a ϕ satisfying (64). Since $\|\tilde{u}\|_{L^2(\mathbb{R}^3)} = \|\tilde{u}\|_{H(\text{div};\mathbb{R}^3)} \lesssim \|u\|_{H(\text{div};\Omega)} = \|u\|_{L^2(\Omega)}$, restricting ϕ to Ω leads to a desirable ϕ. $\Box$

Exact Sequences (ES). The discrete counterpart of the de Rham differential complex (61) is also valid for the finite element spaces $\mathscr{V}(\mathscr{D},\mathscr{T})$:

$$\mathbb{R} \hookrightarrow \mathscr{V}(\text{grad},\mathscr{T}) \xrightarrow{\text{grad}} \mathscr{V}(\text{curl},\mathscr{T}) \xrightarrow{\text{curl}} \mathscr{V}(\text{div},\mathscr{T}) \xrightarrow{\text{div}} \mathscr{V}(L^2,\mathscr{T}). \tag{65}$$

The starting finite element space $\mathscr{V}(\text{grad},\mathscr{T})$ and the ending space $\mathscr{V}(L^2,\mathscr{T})$ are continuous and discontinuous complete polynomial spaces, respectively. For the two spaces in the middle, each one has two types. Therefore we have 4 exact sequences in $\mathbb{R}^3$ and these are all possible exact sequences in $\mathbb{R}^3$ [5]. For completeness we list these exact sequences below:

$$\mathbb{R} \hookrightarrow \mathscr{V}(\text{grad},\mathscr{P}_k,\mathscr{T}) \xrightarrow{\text{grad}} \mathscr{V}(\text{curl},\mathscr{P}_{k-1},\mathscr{T}) \xrightarrow{\text{curl}} \mathscr{V}(\text{div},\mathscr{P}_{k-2},\mathscr{T}) \xrightarrow{\text{div}} \mathscr{V}(L^2,\mathscr{P}_{k-3},\mathscr{T})$$
$$\mathbb{R} \hookrightarrow \mathscr{V}(\text{grad},\mathscr{P}_k,\mathscr{T}) \xrightarrow{\text{grad}} \mathscr{V}(\text{curl},\mathscr{P}_{k-1},\mathscr{T}) \xrightarrow{\text{curl}} \mathscr{V}(\text{div},\mathscr{P}^-_{k-1},\mathscr{T}) \xrightarrow{\text{div}} \mathscr{V}(L^2,\mathscr{P}_{k-2},\mathscr{T})$$
$$\mathbb{R} \hookrightarrow \mathscr{V}(\text{grad},\mathscr{P}_k,\mathscr{T}) \xrightarrow{\text{grad}} \mathscr{V}(\text{curl},\mathscr{P}^-_{k},\mathscr{T}) \xrightarrow{\text{curl}} \mathscr{V}(\text{div},\mathscr{P}_{k-1},\mathscr{T}) \xrightarrow{\text{div}} \mathscr{V}(L^2,\mathscr{P}_{k-2},\mathscr{T})$$
$$\mathbb{R} \hookrightarrow \mathscr{V}(\text{grad},\mathscr{P}_k,\mathscr{T}) \xrightarrow{\text{grad}} \mathscr{V}(\text{curl},\mathscr{P}^-_{k},\mathscr{T}) \xrightarrow{\text{curl}} \mathscr{V}(\text{div},\mathscr{P}^-_{k},\mathscr{T}) \xrightarrow{\text{div}} \mathscr{V}(L^2,\mathscr{P}_{k-1},\mathscr{T}).$$

There exist a sequence of interpolation operators

$$\Pi^{\mathscr{D}} : H(\mathscr{D},\Omega) \cap \operatorname{dom}(\Pi^{\mathscr{D}}) \to \mathscr{V}(\mathscr{D},\mathscr{T})$$

to connect the Sobolev spaces $H(\mathscr{D},\Omega)$ with corresponding finite element spaces $\mathscr{V}(\mathscr{D},\mathscr{T})$. These operators enjoy the following commutative diagram:

$$\begin{array}{ccccccccc}
\mathbb{R} & \longrightarrow & C^\infty(\Omega) & \xrightarrow{\text{grad}} & C^\infty(\Omega) & \xrightarrow{\text{curl}} & C^\infty(\Omega) & \xrightarrow{\text{div}} & C^\infty(\Omega) \\
\big\downarrow & & \Pi^{\text{grad}}\big\downarrow & & \Pi^{\text{curl}}\big\downarrow & & \Pi^{\text{div}}\big\downarrow & & \Pi^{L^2}\big\downarrow \\
\mathbb{R} & \longrightarrow & \mathscr{V}(\text{grad},\mathscr{T}) & \xrightarrow{\text{grad}} & \mathscr{V}(\text{curl},\mathscr{T}) & \xrightarrow{\text{curl}} & \mathscr{V}(\text{div},\mathscr{T}) & \xrightarrow{\text{div}} & \mathscr{V}(L^2,\mathscr{T}),
\end{array}$$

where for simplicity, we replace $H(\mathscr{D},\Omega)\cap \text{dom}(\Pi^{\mathscr{D}})$ by its subspace $C^\infty(\Omega)$.

The sequence in the bottom should be one of the 4 exact sequences in (**ES**). The operator $\Pi^{\mathscr{D}}$, of course, also depends on the specific choice of $\mathscr{V}(\mathscr{D},\mathscr{T})$. Operator $\Pi^{\mathscr{D}}$ is the identity restricted to $\mathscr{V}(\mathscr{D},\mathscr{T})$, namely

$$\Pi^{\mathscr{D}} v = v, \quad \text{for all } v\in \mathscr{V}(\mathscr{D},\mathscr{T}). \tag{66}$$

We refer to [5, 48, 50] for the construction of such canonical interpolation operators. Here we list properties used later and refer to [50, Section 3.6 and Lemma 4.6] for proofs.

Lemma 14 (Operator Π^{curl}). *The interpolation operator Π^{curl} is bounded on $V = \{v\in H^1(\Omega) : \text{curl}\, v \in \mathscr{V}(\text{div},\mathscr{T})\}$ and, with constants only depending on the shape regularity of $\mathscr{T}$, it satisfies*

$$\|h^{-1}(I-\Pi^{\text{curl}})v\| \lesssim \|v\|_1, \quad \textit{for all } v\in V. \tag{67}$$

Lemma 15 (Operator Π^{div}). *The interpolation operator Π^{div} is bounded on $H^1(\Omega)$ and, with constants only depending on the shape regularity of $\mathscr{T}$, it satisfies*

$$\|h^{-1}(I-\Pi^{\text{div}})v\| \lesssim \|v\|_1, \quad \textit{for all } v\in H^1(\Omega). \tag{68}$$

5.1.3 Regular Decomposition

The Helmholtz (or Hodge) decomposition states that a vector field can be written as the sum of a gradient plus a curl. This decomposition is orthogonal in $L^2(\Omega)$ but requires regularity of Ω to be useful to us. Upon sacrificing L^2 orthogonality, we can decompose the space $H(\mathscr{D},\Omega)$ into a regular part $H^1(\Omega)$ plus the kernel of $\mathscr{D}$.

Theorem 17 (Regular Decomposition of $H(\text{curl};\Omega)$). *For any $v\in H(\text{curl};\Omega)$, there exists $\phi\in [H^1(\Omega)]^3$ and $u\in H^1(\Omega)$ such that*

$$v = \phi + \text{grad}\, u.$$

This decomposition is stable in the sense that

$$\|\phi\|_1 + \|u\|_1 \lesssim \|v\|_{A^c}.$$

Proof. For $v \in H(\text{curl};\Omega)$, let $u = \operatorname{curl} v \in [L^2(\Omega)]^3$. Since $\operatorname{div}\operatorname{curl} v = 0$, we can apply Theorem 16 to obtain $\phi \in [H^1(\Omega)]^3$ such that

$$\operatorname{curl}\phi = u = \operatorname{curl} v, \text{ in } \Omega,$$

and

$$\|\phi\|_1 \lesssim \|u\| \le \|v\|_{A^c}.$$

Since $\operatorname{curl}(v-\phi) = 0$, by Theorem 15, there exists $u \in H(\text{grad};\Omega)$ such that

$$\operatorname{grad} u = v - \phi,$$

and

$$\|u\|_1 \lesssim \|v\| + \|\phi\| \lesssim \|v\|_{A^c}.$$

This completes the proof. □

The following lemma concerns the regular inversion of div operator.

Lemma 16 (Regular Inverse of div**).** *For any $v \in H(\text{div};\Omega)$, there exists $\phi \in [H^1(\Omega)]^3$ such that*

$$\operatorname{div}\phi = \operatorname{div} v, \quad \|\phi\|_1 \lesssim \|\operatorname{div} v\|.$$

Proof. Given $v \in H(\text{div};\Omega)$, let f be the zero extension of $\operatorname{div} v$ to a smooth domain $\mathscr{O} \subset \mathbb{R}^3$ containing Ω; obviously $f \in L^2(\mathscr{O})$. We then solve the Poisson equation

$$-\Delta u = f \text{ in } \mathscr{O}, \quad u|_{\partial\mathscr{O}} = 0.$$

If $\phi = -\operatorname{grad} u$, then $\operatorname{div}\phi = -\Delta u = \operatorname{div} v$ in $L^2(\mathscr{O})$. Since $u \in H^2(\mathscr{O})$ and $\|u\|_{2,\mathscr{O}} \lesssim \|f\|_{0,\mathscr{O}}$ because $\mathscr{O}$ is smooth, we deduce that $\phi \in [H^1(\Omega)]^3$ and

$$\|\phi\|_{1,\Omega} \le \|\phi\|_{1,\mathscr{O}} \le \|\operatorname{grad} u\|_{2,\mathscr{O}} \lesssim \|f\|_{0,\mathscr{O}} = \|\operatorname{div} v\|_{0,\Omega},$$

which proves the assertion. □

Similar results can even be established for functions with appropriate traces on the boundary $\partial\Omega$. We refer to [35, 7] for specific constructions.

Theorem 18 (Regular Decomposition of $H(\text{div};\Omega)$). *For any $v \in H(\text{div};\Omega)$, there exist $\phi, u \in [H^1(\Omega)]^3$ such that*

$$v = \phi + \operatorname{curl} u.$$

This decomposition is stable in the sense that

$$\|\phi\|_1 + \|u\|_1 \lesssim \|v\|_{A^d}.$$

Proof. We first apply Lemma 16 to v to find $\phi \in [H^1(\Omega)]^3$ such that

$$\operatorname{div}\phi = \operatorname{div} v, \quad \|\phi\|_1 \lesssim \|\operatorname{div} v\|.$$

Now since $\operatorname{div}(v-\phi)=0$, we apply Theorem 16 to find $u \in [H^1(\Omega)]^3$ such that

$$\operatorname{curl} u = v-\phi, \quad \|u\|_1 \lesssim \|v-\phi\| \le \|v\| + \|\phi\| \lesssim \|v\|_{A^d}.$$

This is the asserted estimate. □

5.1.4 Discrete Regular Decomposition

We now present discrete regular decompositions for finite element spaces $\mathscr{V}(\mathrm{curl},\mathscr{T})$ and $\mathscr{V}(\mathrm{div},\mathscr{T})$, Theorem 19 and 20, following Hiptmair and Xu [52].

Theorem 19 (Discrete Regular Decomposition of $\mathscr{V}(\mathrm{curl},\mathscr{T})$). *Let $\mathscr{V}(\mathrm{grad},\mathscr{T}_h)$ and $\mathscr{V}(\mathrm{curl},\mathscr{T}_h)$ be a pair in the four exact sequences. For any $v \in \mathscr{V}(\mathrm{curl},\mathscr{T}_h)$, there exist $\tilde{v} \in \mathscr{V}(\mathrm{curl},\mathscr{T}_h), \phi \in \mathscr{V}^3$, and $u \in \mathscr{V}(\mathrm{grad},\mathscr{T}_h)$ such that*

$$v = \tilde{v} + \Pi^{\mathrm{curl}}\phi + \operatorname{grad} u, \tag{69}$$

$$\|h^{-1}\tilde{v}\| + \|\phi\|_1 + \|u\|_1 \lesssim \|v\|_{A^c}. \tag{70}$$

Proof. For $v \in H(\mathrm{curl};\Omega)$, we apply the regular decomposition of Theorem 17 to obtain $v = \Psi + \operatorname{grad} U$ with

$$\Psi \in [H^1(\Omega)]^3, U \in H^1(\Omega), \quad \|\Psi\|_1 + \|U\|_1 \lesssim \|v\|_{A^c}.$$

We then split Ψ as $\Psi = (I-\mathscr{I}_{\mathscr{T}})\Psi + \mathscr{I}_{\mathscr{T}}\Psi$, where $\mathscr{I}_{\mathscr{T}} : [H^1(\Omega)]^3 \to \mathscr{V}^3$ is the vector version of the Scott-Zhang quasi-interpolation operator.

Since $\operatorname{curl}\Psi = \operatorname{curl} v \in \mathscr{V}(\mathrm{div},\mathscr{T}_h)$, by Lemma 14, $\Pi^{\mathrm{curl}}\Psi$ is well defined. We apply the interpolation operator Π^{curl} to the decomposition

$$v = (I-\mathscr{I}_{\mathscr{T}})\Psi + \mathscr{I}_{\mathscr{T}}\Psi + \operatorname{grad} U,$$

and use (66) to obtain the discrete decomposition

$$v = \Pi^{\mathrm{curl}}(I-\mathscr{I}_{\mathscr{T}})\Psi + \Pi^{\mathrm{curl}}\mathscr{I}_{\mathscr{T}}\Psi + \operatorname{grad}\Pi^{\mathrm{grad}}U.$$

This implies (69) with

$$\begin{aligned}
&\tilde{v} = \Pi^{\mathrm{curl}}(I-\mathscr{I}_{\mathscr{T}})\Psi \in \mathscr{V}(\mathrm{curl},\mathscr{T}_h),\\
&\phi = \mathscr{I}_{\mathscr{T}}\Psi \in \mathscr{V}^3, \text{ and}\\
&u = \Pi^{\mathrm{grad}}U - \frac{1}{\Omega}\int_{\Omega}\Pi^{\mathrm{grad}}U\,\mathrm{d}x \in \mathscr{V}(\mathrm{grad},\mathscr{T}_h).
\end{aligned}$$

We then prove this decomposition satisfies (70). First, by (67) and (49), we get

$$\begin{aligned}
\|h^{-1}\tilde{v}\| &\le \|h^{-1}(I-\Pi^{\mathrm{curl}})(I-\mathscr{I}_{\mathscr{T}})\Psi\| + \|h^{-1}(I-\mathscr{I}_{\mathscr{T}})\Psi\|\\
&\lesssim \|(I-\mathscr{I}_{\mathscr{T}})\Psi\|_1 + \|\Psi\|_1 \lesssim \|\Psi\|_1 \lesssim \|v\|_{A^c}.
\end{aligned}$$

Second, by the stability of $\mathscr{I}_T$ we obtain

$$\|\phi\|_1 = \|\mathscr{I}_{\mathscr{T}}\Psi\|_1 \lesssim \|\Psi\|_1 \lesssim \|v\|_{A^c},$$

and by that of Π^{grad} we have

$$\|u\|_1 \lesssim \|U\|_1 \lesssim \|v\|_{A^c}.$$

This finishes the proof. □

The following regular decomposition is taken from Hiptmair and Xu [52]; see also Cascón, Nochetto, and Siebert [22].

Theorem 20 (Discrete Regular Decomposition of $\mathscr{V}(\text{div},\mathscr{T})$). *Let $\mathscr{V}(\text{curl},\mathscr{T}_h)$ and $\mathscr{V}(\text{div},\mathscr{T}_h)$ be a pair in the four exact sequences. For any $v \in \mathscr{V}(\text{div},\mathscr{T}_h)$, there exist $\tilde{v} \in \mathscr{V}(\text{div},\mathscr{T}_h), \phi \in \mathscr{V}^3$, and $u \in \mathscr{V}(\text{curl},\mathscr{T}_h)$ such that*

$$v = \tilde{v} + \Pi^{\text{div}}\phi + \operatorname{curl} u, \tag{71}$$

$$\|h^{-1}\tilde{v}\| + \|\phi\|_1 + \|u\|_{A^c} \lesssim \|v\|_{A^d}. \tag{72}$$

Proof. The proof is similar to that of Theorem 19 but a bit trickier. We first obtain

$$v = \Psi + \operatorname{curl} U, \quad \|\Psi\|_1 + \|U\|_1 \lesssim \|v\|_{A^d}.$$

But we cannot apply the interpolation operator Π^{div} directly and use the commutative diagram relation $\Pi^{\text{div}}\operatorname{curl} U = \operatorname{curl}\Pi^{\text{curl}}U$ because $U \in [H^1(\Omega)]^3$ only and the interpolation Π^{curl} is not well defined on $[H^1(\Omega)]^3$.

To overcome this difficulty, we further split v as follows:

$$v = (I - \Pi^{\text{div}})\Psi + \Pi^{\text{div}}(I - \mathscr{I}_{\mathscr{T}})\Psi + \Pi^{\text{div}}\mathscr{I}_{\mathscr{T}}\Psi + \operatorname{curl} U. \tag{73}$$

Invoking the commutative diagram property $\operatorname{div}\Pi^{\text{div}}\Psi = \Pi^{L^2}\operatorname{div}\Psi$, and the fact $\operatorname{div}\Psi = \operatorname{div} v \in \mathscr{V}(L^2,\mathscr{T})$, we have $\operatorname{div}(I - \Pi^{\text{div}})\Psi = 0$. Applying the regular inversion of curl operator (Lemma 17), there exists $Q \in H^1(\Omega)$ such that $\operatorname{curl} Q = (I - \Pi^{\text{div}})\Psi$.

If $\tilde{U} = U + Q$, then $\tilde{U} \in H^1(\Omega)$ and $\operatorname{curl}\tilde{U} \in \mathscr{V}(\text{div},\mathscr{T}_h)$. By Lemma 14, $\Pi^{\text{curl}}\tilde{U}$ is well defined. The decomposition (73) thus becomes

$$v = \Pi^{\text{div}}(I - \mathscr{I}_{\mathscr{T}})\Psi + \Pi^{\text{div}}\mathscr{I}_{\mathscr{T}}\Psi + \operatorname{curl}\tilde{U}.$$

We then apply Π^{div} operator to both sides and use property $\Pi^{\text{div}}\operatorname{curl}\tilde{U} = \operatorname{curl}\Pi^{\text{curl}}\tilde{U}$ to obtain

$$v = \Pi^{\text{div}}(I - \mathscr{I}_{\mathscr{T}})\Psi + \Pi^{\text{div}}\mathscr{I}_{\mathscr{T}}\Psi + \operatorname{curl}\Pi^{\text{curl}}\tilde{U},$$

which implies (71).

The stability (72) of this decomposition is similar to that of Theorem 19. □

5.2 Space Decomposition and Multigrid Methods

In this section, we first recall the space decomposition of $\mathscr{V}(\text{grad},\mathscr{T})$ discussed in Section §4, following [27], and then present space decompositions for $\mathscr{V}(\text{curl},\mathscr{T})$ and $\mathscr{V}(\text{div},\mathscr{T})$. On the basis of these space decompositions, we develop multigrid methods for solving $H(\text{curl})$ and $H(\text{div})$ systems. We consider bisection grids $\mathscr{T}_N$ which admits a decomposition $\mathscr{T}_N = \mathscr{T}_0 + \mathscr{B}$.

Let $\{\phi_p : p \in \mathscr{P}\}$, $\{\phi_e : e \in \mathscr{E}\}$, and $\{\phi_f : f \in \mathscr{F}\}$ be "nodal" basis functions. Namely $\mathscr{V}(\text{grad},\mathscr{T}) = \text{span}\{\phi_p : p \in \mathscr{P}\}, \mathscr{V}(\text{curl},\mathscr{T}) = \text{span}\{\phi_e : e \in \mathscr{E}\}$, and $\mathscr{V}(\text{div},\mathscr{T}) = \text{span}\{\phi_f : f \in \mathscr{F}\}$, where $\mathscr{P}$ (nodes), $\mathscr{E}$ (edges), and $\mathscr{F}$ (faces) are the degrees of freedom of the three spaces under consideration.

If $\mathscr{V}_p = \text{span}\{\phi_p\}$, $\mathscr{V}_e = \text{span}\{\phi_e\}$, and $\mathscr{V}_f = \text{span}\{\phi_f\}$ denote one dimensional subspaces, we then have the standard basis decompositions:

$$\mathscr{V}(\text{grad},\mathscr{T}) = \sum_{p\in\mathscr{P}} \mathscr{V}_p, \quad \mathscr{V}(\text{curl},\mathscr{T}) = \sum_{e\in\mathscr{E}} \mathscr{V}_e, \quad \mathscr{V}(\text{div},\mathscr{T}) = \sum_{f\in\mathscr{F}} \mathscr{V}_f.$$

Moreover, if $v = \sum_{p\in\mathscr{P}} v_p$, $v = \sum_{e\in\mathscr{E}} v_e$ and $v = \sum_{f\in\mathscr{F}} v_f$, then mesh shape regularity implies

$$\begin{aligned} \sum_{p\in\mathscr{P}} \|h^{-1}v_p\|^2 &\lesssim \|h^{-1}v\|^2, \\ \sum_{e\in\mathscr{E}} \|h^{-1}v_e\|^2 &\lesssim \|h^{-1}v\|^2, \\ \sum_{f\in\mathscr{F}} \|h^{-1}v_f\|^2 &\lesssim \|h^{-1}v\|^2. \end{aligned} \tag{74}$$

Let $\mathscr{T}_i = \mathscr{T}_0 + (b_1,\cdots,b_i)$ be the i-th mesh and $\phi_{i,p_i} \in \mathscr{V}(\mathscr{T}_i;\mathscr{P}_1)$ denote the linear nodal basis associated with vertex $p_i \in \mathscr{N}(\mathscr{T}_i)$. We define the sub-spaces

$$\mathscr{V}_0 = \mathscr{V}(\mathscr{T}_0;\mathscr{P}_1), \quad \mathscr{V}_i = \text{span}\{\phi_{i,p_i}, \phi_{i,p_{l_i}}, \phi_{i,p_{r_i}}\},\ p_i \in \mathscr{N}(\mathscr{T}_i), \tag{75}$$

where recall that p_{l_i} and p_{r_i} are two end points of the edge and p_i is the middle point of that edge.

Space Decompositions. We now present space decompositions of $\mathscr{V}(\text{curl},\mathscr{T})$ and $\mathscr{V}(\text{div},\mathscr{T})$ in the same vein of that for $\mathscr{V}(\text{grad},\mathscr{T})$ of Section §4.6:

$$\mathscr{V}(\text{grad},\mathscr{T}) = \sum_{p\in\mathscr{P}} \mathscr{V}_p + \sum_{i=1}^{N} \mathscr{V}_i. \tag{76}$$

If $\mathscr{R}_i$ is the ring of vertex p_i, which consists of all simplexes of $\mathscr{T}_i$ containing the vertex p_i, we define $\mathscr{V}_i(\mathscr{D},\mathscr{R}_i)$ as follows:

$$\mathscr{V}_i(\text{curl},\mathscr{R}_i) = \Pi_i^{\text{curl}} \mathscr{V}_i^3 + \text{grad}\,\mathscr{V}_i. \tag{77}$$

$$\mathscr{V}_i(\text{div},\mathscr{R}_i) = \Pi_i^{\text{div}} \mathscr{V}_i^3 + \text{curl}\,\mathscr{V}_i. \tag{78}$$

If $\mathscr{V}_i^3 \subset V(\mathscr{D},\mathscr{T})$, then the interpolation operator $\Pi_i^{\mathscr{D}}$ is the identity and we can ignore it. The macro space decompositions of $\mathscr{V}(\mathscr{D},\mathscr{T})$ are as follows:

$$\mathscr{V}(\text{curl},\mathscr{T}) = \sum_{e\in\mathscr{E}} \mathscr{V}_e + \sum_{p\in\mathscr{P}} \text{grad}\,\mathscr{V}_p + \sum_{i=0}^{N} \mathscr{V}_i(\text{curl},\mathscr{R}_i), \tag{79}$$

$$\mathscr{V}(\text{div},\mathscr{T}) = \sum_{F\in\mathscr{F}} \mathscr{V}_f + \sum_{e\in\mathscr{E}} \text{curl}\,\mathscr{V}_e + \sum_{i=0}^{N} \mathscr{V}_i(\text{div},\mathscr{R}_i). \tag{80}$$

Here for the convenience of notation, we include the coarsest space by defining $\mathscr{R}_0 = \mathscr{T}_0$ and $\mathscr{V}_0(\mathscr{D},\mathscr{R}_0) = \mathscr{V}(\mathscr{D},\mathscr{T}_0)$.

We will apply the Successive Subspace Correction (SSC) method to the space decompositions (79) and (80). The common feature is to apply smoothing in the finest space first and then the multilevel iteration to $\mathscr{V}_i(\mathscr{D},\mathscr{R}_i)$. For completeness, we also list the algorithm for $H(\text{grad})$ problem.

$H(\text{grad})$ **Problem**

$$u \leftarrow u + B^{\text{grad}}(f - A^g u).$$

The operation of B^{grad} consists of two steps:

1. Smoothing in the finest space: $u \leftarrow u + S^{\text{grad}}(f - A^g u)$
2. SSC for $H(\text{grad})$ system on $\sum_i \mathscr{V}_i$:

$$u \leftarrow u + R_i Q_i (f - A^g u), \quad i = 0:N.$$

$H(\text{curl})$ **System**

$$u \leftarrow u + B^{\text{curl}}(f - A^c u).$$

The operation of B^{curl} consists of three steps:

1. Smoothing in the finest space: $u \leftarrow u + S^{\text{curl}}(f - A^c u)$
2. Smoothing in the kernel space for the finest space $\sum_p \mathscr{V}_p$:

$$u \leftarrow u + \text{grad}\, S^{\text{grad}}(f - A^c u),$$

3. SSC for $H(\text{curl})$ system on $\sum_i \mathscr{V}_i(\text{curl},\mathscr{R}_i)$:

$$u \leftarrow u + R_i Q_i (f - A^c u), \quad i = 0:N.$$

$H(\text{div})$ **System**

$$u \leftarrow u + B^{\text{div}}(f - A^d u).$$

The operation of B^{div} consists of three steps:

1. Smoothing in the finest space: $u \leftarrow u + S^{\text{div}}(f - A^d u)$
2. Smoothing in the kernel space for the finest space $\sum_e \mathscr{V}_e$:

$$u \leftarrow u + \text{curl}\, S^{\text{curl}}(f - A^d u).$$

3. SSC for $H(\mathrm{div})$ system on $\sum_i \mathscr{V}_i(\mathrm{curl}, \mathscr{R}_i)$:

$$u \leftarrow u + R_i Q_i (f - A^d u), \quad i = 0:N.$$

5.3 Stable Decomposition

We now prove that the multilevel space decompositions (79) and (80) are stable. Our approach is based on the stable decomposition for $\mathscr{V}(\mathrm{grad};\mathscr{T})$ discussed in Section §4 (Theorem 11): for any $v \in \mathscr{V}(\mathrm{grad},\mathscr{T})$, there exist $v_p \in \mathscr{V}_p, v_i \in \mathscr{V}_i$ such that

$$v = \sum_{p\in\mathscr{P}} v_p + \sum_{i=0}^{N} v_i, \tag{81}$$

$$\sum_{p\in\mathscr{P}} \|h^{-1} v_p\|^2 + \sum_{i=0}^{N} \|h_i^{-1} v_i\|^2 \lesssim \|v\|_{A^g}^2. \tag{82}$$

We first use the stable decomposition (81) and the discrete regular decomposition to give a space decomposition for $\mathscr{V}(\mathrm{curl},\mathscr{T})$.We next employ the results of $\mathscr{V}(\mathrm{curl},\mathscr{T})$ to give a stable decomposition of $\mathscr{V}(\mathrm{div},\mathscr{T})$.

Theorem 21 (Stable Decomposition of $\mathscr{V}(\mathrm{curl};\mathscr{T})$). *Let $\mathscr{T}_N = \mathscr{T}_0 + \mathscr{B}$ be a bisection grid. For every $v \in \mathscr{V}(\mathrm{curl},\mathscr{T}_N)$, there exist $\tilde{v}_e \in \mathscr{V}_e, \tilde{u}_p \in \mathscr{V}_p$ and $w_i = \Pi_i^{\mathrm{curl}} \phi_i + \mathrm{grad}\, u_i \in \mathscr{V}_i(\mathrm{curl},\mathscr{R}_i)$ for all $e \in \mathscr{E}, p \in \mathscr{P}, i = 1:N$, such that*

$$v = \sum_{e\in\mathscr{E}} \tilde{v}_e + \sum_{p\in\mathscr{P}} \mathrm{grad}\, \tilde{u}_p + \sum_{i=0}^{N} w_i, \tag{83}$$

$$\sum_{e\in\mathscr{E}} \|h^{-1}\tilde{v}_e\|^2 + \sum_{p\in\mathscr{P}} \|h^{-1}\tilde{u}_p\|^2 + \sum_{i=0}^{N} \left(\|h^{-1}\phi_i\|^2 + \|h^{-1}u_i\|^2 \right) \lesssim \|v\|_{A^c}^2. \tag{84}$$

Proof. [1] We first consider the case $\mathscr{V}^3 \subset \mathscr{V}(\mathrm{curl},\mathscr{T}_N)$ which excludes only the lowest order space $\mathscr{V}(\mathrm{curl},\mathscr{P}_1^-,\mathscr{T}_N)$.

For any $v \in \mathscr{V}(\mathrm{curl},\mathscr{T}_N)$, we can apply Theorem 19 to obtain a discrete regular decomposition $\tilde{v} \in \mathscr{V}(\mathrm{curl},\mathscr{T}_N), \phi \in \mathscr{V}^3$ and $u \in \mathscr{V}(\mathrm{grad},\mathscr{T}_N)$ such that

$$v = \tilde{v} + \phi + \mathrm{grad}\, u$$
$$\|h^{-1}\tilde{v}\|^2 + \|\phi\|_1^2 + \|u\|_1^2 \lesssim \|v\|_1^2.$$

For $\mathscr{T}_N$, we can choose ϕ so that $\phi = \sum_{i=0}^{N} \phi_i$ using the quasi-interpolation operator $\mathscr{I}_{\mathscr{T}}$ adapted to bisection grids; see Section §4.6 for the construction of $\mathscr{I}_{\mathscr{T}}$.

We apply the basis and multilevel decompositions of H^1 finite element spaces to obtain the desirable decomposition

$$\tilde{v} = \sum_{e\in\mathscr{E}} \tilde{v}_e,\ \phi = \sum_{i=0}^{N} \phi_i, \quad u = \sum_{p\in\mathscr{P}} \tilde{u}_p + \sum_{i=0}^{N} u_i.$$

The stability (84) of the decomposition results from the following inequalities:

1. $\sum_{e\in\mathscr{E}} \|h^{-1}\tilde{v}_e\|^2 \lesssim \|h^{-1}\tilde{v}\|^2$ by (74);
2. $\sum_{i=0}^{N} \|h^{-1}\phi_i\|^2 \lesssim \|\phi\|_1^2$ by the stable decomposition (55);
3. $\sum_{p\in\mathscr{P}} \|h^{-1}\tilde{u}_p\|^2 + \sum_{i=0}^{N} \|h^{-1}u_i\|^2 \lesssim \|u\|_1^2$ by the stable decomposition (82).

[2] Now we consider the case $\mathscr{V}^3 \not\subseteq \mathscr{V}(\text{curl}, \mathscr{T}_N)$, i.e., the space $\mathscr{V}(\text{curl}, \mathscr{P}_1^-, \mathscr{T}_N)$. By Theorem 19, we have the discrete regular decomposition

$$v = \tilde{v} + \Pi^{\text{curl}}\phi + \operatorname{grad} u. \tag{85}$$

The key is a multilevel decomposition of the middle term. If $\phi = \sum_{i=0}^{N} \phi_i$ is the stable decomposition of ϕ, then

$$\Pi^{\text{curl}}\phi = \sum_{i=0}^{N} \Pi_i^{\text{curl}}\phi_i + \Pi^{\text{curl}} \sum_{i=0}^{N} (\phi_i - \Pi_i^{\text{curl}}\phi_i), \tag{86}$$

because $\mathscr{V}(\text{curl}, \mathscr{R}_i) \subset \mathscr{V}(\text{curl}, \mathscr{T}_N)$ and $\Pi_i^{\text{curl}} = \Pi^{\text{curl}}\Pi_i^{\text{curl}}$. We now show $\operatorname{curl}(\phi_i - \Pi_i^{\text{curl}}\phi_i) = 0$. For any face $f \in \mathscr{F}(\mathscr{R}_i)$, using integration by parts and the definition of Π_i^{curl}, we conclude

$$\int_f \operatorname{curl}(\phi_i - \Pi_i^{\text{curl}}\phi_i)\cdot n\, dS = \int_{\partial f} (\phi_i - \Pi_i^{\text{curl}}\phi_i)\cdot t\, ds = 0.$$

Since $\operatorname{curl}(\phi_i - \Pi_i^{\text{curl}}\phi_i)$ is piecewise constant, we deduce $\operatorname{curl}(\phi_i - \Pi_i^{\text{curl}}\phi_i) = 0$.

From the exact sequence

$$\mathscr{V}(\text{grad}, \mathscr{P}_2, \mathscr{R}_i) \to \mathscr{V}(\text{curl}, \mathscr{P}_1, \mathscr{R}_i) \to \mathscr{V}(\text{div}, \mathscr{P}_1^-, \mathscr{R}_i),$$

there exists $q_i \in \mathscr{V}(\text{grad}, \mathscr{P}_2, \mathscr{R}_i)$ such that $\phi_i - \Pi_i^{\text{curl}}\phi_i = \operatorname{grad} q_i$ and $\|q_i\| \lesssim \|\operatorname{grad} q_i\|$. Let $q = \sum q_i$ and $\int_\Omega q\,\mathrm{d}x = 0$. Using the commutative diagram, we have

$$\Pi^{\text{curl}} \sum_{i=0}^{N} (\phi_i - \Pi_i^{\text{curl}}\phi_i) = \Pi^{\text{curl}} \operatorname{grad} \sum_{i=0}^{N} q_i = \operatorname{grad} \Pi^{\text{grad}} q,$$

where $\Pi^{\text{grad}} : \mathscr{V}(\text{grad}, \mathscr{P}_2, \mathscr{T}_N) \to \mathscr{V}(\text{grad}, \mathscr{P}_1, \mathscr{T}_N)$. Let $\hat{u} = u + \Pi^{\text{grad}} q$. Then $\hat{u} \in \mathscr{V}(\text{grad}, \mathscr{P}_1, \mathscr{T}_N)$ and the decomposition (85) becomes

$$v = \tilde{v} + \sum_i \Pi_i^{\text{curl}}\phi_i + \operatorname{grad}\hat{u}. \tag{87}$$

We then apply the decomposition (81) to $\hat{u}$ as in the previous case, i.e.

$$\hat{u} = \sum_{p\in\mathscr{P}} \tilde{u}_p + \sum_{i=0}^{N} u_i,$$

to obtain the desired decomposition (83).

To prove the stability (84) of the decomposition, it suffices to prove

$$\|\operatorname{grad} q\| \lesssim \|v\|_{A^c}, \tag{88}$$

which can be obtained from the Strengthened Cauchy Schwarz inequality

$$\begin{aligned}\|\operatorname{grad} q\|^2 &= (\sum_{i=0}^{N}\operatorname{grad} q_i, \sum_{i=0}^{N}\operatorname{grad} q_j) \le \sum_{i=0}^{N}\|\operatorname{grad} q_i\|^2 + 2\sum_{i=0}^{N}\sum_{j>i}^{N}|(\operatorname{grad} q_i, \operatorname{grad} q_j)| \\ &\lesssim \sum_{i=0}^{N}\|\operatorname{grad} q_i\|^2 = \sum_{i=0}^{N}\|\phi_i - \Pi_i^{\mathrm{curl}}\phi_i\|^2 \lesssim \sum_{i=0}^{N}\|h^{-1}\phi_i\|^2 \lesssim \|\phi\|_1^2 \lesssim \|v\|_{A_c}^2.\end{aligned}$$

This completes the proof. □

We conclude with a similar result for $\mathscr{V}(\mathrm{div},\mathscr{T})$. Its proof follows along the same lines as those of Theorem 21. We refer to [28] for details.

Theorem 22 (Stable Decomposition of $H(\mathrm{div};\Omega)$). *Let $\mathscr{T}_N = \mathscr{T}_0 + \mathscr{B}$ be a bisection grid. For every $v \in \mathscr{V}(\mathrm{div},\mathscr{T}_N)$ with $\mathscr{V}^3 \subset \mathscr{V}(\mathrm{curl},\mathscr{T}_N)$, there exist $\tilde{v}_f \in \mathscr{V}_f, \tilde{u}_e \in \mathscr{V}_e$ and $w_i \in \mathscr{V}_i(\mathrm{div},\mathscr{R}_i)$ for all $f\in\mathscr{F}, e\in\mathscr{E}, i = 0:N$, such that*

$$v = \sum_{f\in\mathscr{F}} \tilde{v}_f + \sum_{e\in\mathscr{E}} \operatorname{curl}\tilde{u}_e + \sum_{i=0}^{N} w_i, \tag{89}$$

and

$$\sum_{F\in\mathscr{F}}\|h^{-1}v_f\|^2 + \sum_{e\in\mathscr{E}}\|h^{-1}\tilde{u}_e\|^2 + \sum_{i=0}^{N}\left(\|h^{-1}\phi_i\|^2 + \|h^{-1}u_i\|^2\right) \lesssim \|v\|_{A^d}^2. \tag{90}$$

A remaining important ingredient, the SCS inequality for the space decompositions (79) and (80), can be established as well. Consequently, we have uniform convergence of multigrid methods for $H(\mathrm{curl})$ or $H(\mathrm{div})$ systems. We state the result below and refer to [28] for details.

Theorem 23. *The multigrid methods (c.f. algorithms in §5.2) for $H(\mathrm{curl})$ or $H(\mathrm{div})$ systems based on the space decompositions (79) or (80)), respectively, are uniformly convergent.*

6 The Auxiliary Space Method and HX Preconditioner for Unstructured Grids

In previous sections, we study multilevel methods formulated over a hierarchy of quasi-uniform or graded meshes. The geometric structure of these meshes is essential for both the design and analysis of such methods. Unfortunately, many grids in practice are not hierarchical.

We use the term *unstructured grids* to refer those grids that do not possess much geometric or topological structure. The design and analysis of efficient multilevel solvers for unstructured grids is a topic of great theoretical and practical interest. In this section, we discuss a special class of optimal preconditioners developed by Hiptmair and Xu [52] that can be effectively applied to unstructured grids. This type of preconditioners have been developed in the theoretical framework of the *auxiliary space method*.

6.1 The Auxiliary Space Method

The method of subspace correction consists of solving a system of equations in a vector space by solving on appropriately chosen *subspaces* of the original space. Such subspaces are, however, not always available. The auxiliary space method (Xu 1996 [92]) is for designing preconditioners using auxiliary spaces which are not necessarily subspaces of the original subspace.

To solve the equation $a(u,v) = (f,v)$ in a Hilbert space $\mathscr{V}$, we consider

$$\overline{\mathscr{V}} = \mathscr{V} \times \mathscr{W}_1 \times \cdots \times \mathscr{W}_J \ , \tag{91}$$

where $\mathscr{W}_1, \ldots, \mathscr{W}_J$, $J \in \mathbb{N}$ are auxiliary (Hilbert) spaces endowed with inner products $\overline{a}_j(\cdot,\cdot)$, $j = 1, \ldots, J$.

A distinctive feature of the auxiliary space method is the presence of $\mathscr{V}$ in (91), but as a component of $\overline{\mathscr{V}}$. The space $\mathscr{V}$ is equipped with an inner product $d(\cdot,\cdot)$ different from $a(\cdot,\cdot)$. The operator $D : \mathscr{V} \mapsto \mathscr{V}$ induced by $d(\cdot,\cdot)$ on $\mathscr{V}$ leads to the smoother $S = D^{-1}$. For each $\mathscr{W}_j$ we need $\Pi_j : \mathscr{W}_j \mapsto \mathscr{V}$ which gives

$$\Pi := Id \times \Pi_1 \times \cdots \times \Pi_J : \overline{\mathscr{V}} \mapsto \mathscr{V}, \tag{92}$$

with properties

$$\|\Pi_j w_j\|_A \le c_j \overline{a}(w_j, w_j)^{1/2} \ , \qquad \text{for all } w_j \in \mathscr{W}_j, j = 1, \cdots, J \ , \tag{93}$$

$$\|v\|_A \le c_s d(v,v)^{1/2} \ , \qquad \text{for all } v \in \mathscr{V} \ , \tag{94}$$

and for every $v \in \mathscr{V}$, there exist $v_0 \in \mathscr{V}$ and $w_j \in \mathscr{W}_j$ such that $v = v_0 + \sum_{j=1}^{J} \Pi_j w_j$ and

$$d(v_0,v_0)^{1/2}+\sum_{j=1}^{J}\overline{a}_j(w_j,w_j)^{1/2}\le c_0\|v\|_A\,. \tag{95}$$

Let $\bar{A}_i$, for $i=1,\cdots,J$, be operators induced by $(\cdot,\cdot)_{A_i}$. Then the auxiliary space preconditioner is given by

$$B=S+\sum_{j=1}^{J}\Pi_j\bar{A}_j^{-1}\Pi_j^*. \tag{96}$$

The estimate of the condition number $\kappa(BA)$ is given below.

Theorem 24. *Let $\Pi = Id\times\Pi_1\times\cdots\times\Pi_J:\bar{\mathscr{V}}=\mathscr{V}\times\mathscr{W}_1\times\cdots\times\mathscr{W}_J\mapsto\mathscr{V}$ satisfy properties (93), (94), and (95). Then the auxiliary space preconditioner B given in (96) admits the following estimate:*

$$\kappa(BA)\le c_0^2(c_s^2+c_1^2+\cdots+c_J^2)\,. \tag{97}$$

Proof. [1] We first prove $(BAu,u)_A\le(c_s^2+c_1+\cdots+c_J^2)(u,u)_A$ and consequently $\lambda_{\max}(BA)\le(c_s^2+c_1+\cdots+c_J^2)$. By definition of B, we have:

$$(BAu,u)_A=(SAu,u)_A+\sum_{j=1}^{J}(\Pi_j\bar{A}_j^{-1}\Pi_j^*Au,u)_A.$$

We use Cauchy-Schwarz inequality and (94) to control the first term as

$$\begin{aligned}(SAu,u)_A&\le(SAu,SAu)_A^{1/2}(u,u)_A^{1/2}\le c_s d(SAu,SAu)^{1/2}(u,u)_A^{1/2}\\&=c_s(SAu,Au)^{1/2}(u,u)_A^{1/2}=c_s(SAu,u)_A^{1/2}(u,u)_A^{1/2},\end{aligned}$$

which leads to $(SAu,u)_A\le c_s^2(u,u)_A$.

Similarly we use Cauchy-Schwarz inequality and (93) to control the term as

$$\begin{aligned}(\Pi_j\bar{A}_j^{-1}\Pi_j^*Au,u)_A&\le(\Pi_j\bar{A}_j^{-1}\Pi_j^*Au,\,\Pi_j\bar{A}_j^{-1}\Pi_j^*Au)_A^{1/2}(u,u)_A^{1/2}\\&\le c_j(\bar{A}_j^{-1}\Pi_j^*Au,\,\bar{A}_j^{-1}\Pi_j^*Au)_{\bar{A}_j}^{1/2}(u,u)_A^{1/2}\\&=c_j(\bar{A}_j^{-1}\Pi_j^*Au,\,\Pi_j^*Au)^{1/2}(u,u)_A^{1/2}\\&=c_j(\Pi_j\bar{A}_j^{-1}\Pi_j^*Au,\,u)_A^{1/2}(u,u)_A^{1/2},\end{aligned}$$

which leads to $(\Pi_j\bar{A}_j^{-1}\Pi_j^*Au,u)_A\le c_j^2(u,u)_A$.

[2] We then prove there exists $u\in\mathscr{V}$ such that $(u,u)_A\le c_0^2(BAu,u)_A$ and consequently $\lambda_{\min}(BA)\ge c_0^{-2}$.

We choose $u=v_0+\sum_{j=1}^{J}\Pi_jw_j$ satisfying (95). Then

$$
\begin{aligned}
(\Pi_j w_j, u)_A &= (\Pi_j w_j, Au) = (w_j, \Pi_j^* Au) = (w_j, \bar{A}_j^{-1}\Pi_j^* Au)_{\bar{A}_j} \\
&\le \|w_j\|_{\bar{A}_j}(\bar{A}_j^{-1}\Pi_j^* Au, \bar{A}_j^{-1}\Pi_j^* Au)_{\bar{A}_j}^{1/2} = \|w_j\|_{\bar{A}_j}(BAu,u)_A^{1/2}.
\end{aligned}
$$

Similarly $(v_0, u)_A \le \|v_0\|_D (BAu,u)_A^{1/2}$. Therefore

$$
\begin{aligned}
(u,u)_A &= (v_0 + \sum_{j=1}^J w_j, u)_A \le (\|v_0\|_D + \sum_{j=1}^J \|w_j\|_{\bar{A}_j})(BAu,u)_A^{1/2} \\
&\le c_0 (u,u)_A^{1/2}(BAu,u)_A^{1/2},
\end{aligned}
$$

which leads to the desired result. □

6.2 HX Preconditioner

We present an *auxiliary space preconditioner* for $H(\text{curl})$ and $H(\text{div})$ systems developed in Hiptmair and Xu [52] (see also R. Beck [10] for a special case). The basic idea is to apply an auxiliary space preconditioner framework in [92], to the discrete regular decompositions of $\mathscr{V}(\text{curl},\mathscr{T})$ or $\mathscr{V}(\text{div},\mathscr{T})$. The resulting preconditioner for the $H(\text{curl})$ systems is

$$
B^{\text{curl}} = S^{\text{curl}} + \Pi^{\text{curl}} B^{\text{grad}} (\Pi^{\text{curl}})^t + \text{grad } B^{\text{grad}} (\text{grad})^t. \tag{98}
$$

The implementation makes use of the input data: the $H(\text{curl})$ stiffness matrix A, the coordinates of the grid points, along with the discrete gradient grad (for the lowest order Nédélec element case, it is simply the "vertex"-to-"edge" mapping with entries 1 or -1). Based on the coordinates, one can easily construct the interpolation operator Π_h^{curl}. Then the "Auxiliary space Maxwell solver" consists of the following three components:

1. The smoother S^{curl} of A (it could be the standard Jacobi or symmetric Gauss-Seidel methods).
2. An algebraic multigrid (AMG) solver B^{grad} for $\text{grad}^t A\, \text{grad}$
3. An (vector) AMG solver B^{grad} for $\left(\Pi^{\text{curl}}\right)^T A \Pi^{\text{curl}}$.

Similarly

$$
\begin{aligned}
B^{\text{div}} &= S^{\text{div}} + \Pi^{\text{div}} B^{\text{grad}} (\Pi_h^{\text{div}})^t + \text{curl } B^{\text{curl}} (\text{curl})^t \\
&= S^{\text{div}} + \Pi^{\text{div}} B^{\text{grad}} (\Pi^{\text{div}})^t + \text{curl } S^{\text{curl}} (\text{curl})^t + \text{curl}\, \Pi^{\text{curl}} B^{\text{grad}} (\Pi^{\text{curl}})^t (\text{curl})^t.
\end{aligned}
$$

This preconditioner consists of 4 Poisson solvers B^{grad} for $H(\text{curl})$ (and 6 for $H(\text{div})$) as well as 1 simple relaxation method (S^{curl}) such as point Jacobi for $H(\text{curl})$ (and 2 relaxation methods for $H(\text{div})$).

The point here is that we can use well-developed AMG for H^1 systems for the Poisson solver B^{grad} to obtain robust AMG methods for $H(\text{curl})$ and $H(\text{div})$ systems.

These classes of preconditioners are in some way a “grey-box” AMG as it makes use of information on geometric grids (and associated interpolation operators). But the overhead is minimal and it requires very little programming effort. It has been proved in [52] that it is optimal and efficient for problems on unstructured grids.

To interpret B^{curl} as an auxiliary space preconditioner, we choose $\mathscr{V} = \mathscr{V}(\text{curl}, \mathscr{T})$ and $\mathscr{W}_1 = \mathscr{W}_2 = \mathscr{V}(\text{grad}, \mathscr{T})$. The inner product for the smoother is induced by the diagonal matrix of A^{curl} and the inner product $\bar{A}_1, \bar{A}_2$ is induced by $(B^{\text{grad}})^{-1}$. The operator $\Pi_1 : \mathscr{W}_1 \to \mathscr{V}$ is the interpolation Π^{curl} and $\Pi_2 = \text{grad} : \mathscr{W}_1 \to \mathscr{V}$.

Theorem 25. *Suppose B^{grad} is an SPD matrix such that $((B^{\text{grad}})^{-1}u,u) \eqsim (u,u)_1$. Then the preconditioner B^{curl} defined by (98) admits the estimate*

$$\kappa(B^{\text{curl}}A^{\text{curl}}) \lesssim 1.$$

Proof. In view of Theorem 97, it suffices to verify properties (93), (94), and (95).

The property (94) is an easy consequence of Cauchy-Schwarz inequality and shape regularity of the mesh. We use the stability of the operator $\Pi_1 = \Pi^{\text{curl}}$ and $\Pi_2 = \text{grad}$ discussed in Section 5.1 and inequality $(u,u)_1 \lesssim ((B^{\text{grad}})^{-1}u,u)$ to get (94). To get (95), we can use the discrete regular decomposition in Section 5.1.4 and the inequality $((B^{\text{grad}})^{-1}u,u) \lesssim (u,u)_1$. This completes the proof.

We state a similar result for B^{div} below and leave the proof to readers.

Theorem 26. *Suppose B^{grad} is an SPD matrix such that $((B^{\text{grad}})^{-1}u,u) \eqsim (u,u)_1$. Then the preconditioner*

$$B^{\text{div}} = S^{\text{div}} + \Pi^{\text{div}}B^{\text{grad}}(\Pi^{\text{div}})^t + \text{curl } S^{\text{curl}}(\text{curl})^t + \text{curl}\,\Pi^{\text{curl}}B^{\text{grad}}(\Pi^{\text{curl}})^t(\text{curl})^t$$

admits the estimate

$$\kappa(B^{\text{div}}A^{\text{div}}) \lesssim 1.$$

For $H(\text{curl})$ systems, the preconditioners have been included and tested in LLNL’s *hypre* package [36, 37, 38] based on its parallel algebraic multigrid solver “BoomerAMG” [46]. It is a parallel implementation, almost a ‘black-box” as it requires only discrete gradient matrix plus vertex coordinates, it can handle complicated geometries and coefficient jumps, scales with the problem size and on large parallel machines, supports simplified magnetostatics mode, and can utilize Poisson matrices, when available. Extensive numerical experiments demonstrate that this preconditioner is also efficient and robust for more general equations (see Hiptmair and Xu [52], and Kolev and Vassilevski [54, 55]) such as

$$\text{curl}(\mu(x)\,\text{curl}\,u) + \sigma(x)u = f \tag{99}$$

where μ and σ may be discontinuous, degenerate, and exhibit large variations.

For this type of general equations, we may not expect that the simple Poisson solvers are sufficient to handle possible variations of μ and σ. Let us argue roughly what the right equations are to replace the Poisson equations. Let us assume our problems has sufficient regularity (e.g., Ω is convex). We then have

$$\|\operatorname{grad} u\|^2 \eqsim \|\operatorname{curl} u\|^2 + \|\operatorname{div} u\|^2.$$

If $u(=\operatorname{curl} w) \in N(\text{curl})^{\perp}$, then $\|\operatorname{grad} u\| = \|\operatorname{curl} u\|$. Roughly, we get the following equivalence:

$$(\mu \operatorname{curl} u, \operatorname{curl} u) + (\sigma u, u) \eqsim (\mu \operatorname{grad} u, \operatorname{grad} u) + (\sigma u, u),$$

which corresponds to the following operator:

$$L_1 u \equiv -\operatorname{div}(\mu(x) \operatorname{grad} u) + \sigma(x) u. \tag{100}$$

On the other hand, if $u, v \in N(\text{curl})$, $u = \operatorname{grad} p$ and $v = \operatorname{grad} q$,

$$(\mu \operatorname{curl} u, \operatorname{curl} v) + (\sigma u, v) = (\sigma \operatorname{grad} p, \operatorname{grad} q)$$

which corresponds to the following operator:

$$L_2 u \equiv -\operatorname{div}(\sigma(x) \operatorname{grad} p). \tag{101}$$

We obtain the following preconditioner for the general equation (99):

$$B^{\text{curl}} = S^{\text{curl}} + \Pi^{\text{curl}} B_1^{\text{grad}} (\Pi^{\text{curl}})^t + \text{grad } B_2^{\text{grad}} (\text{grad})^t$$

where is B_1^{grad} is a preconditioner for the operator in the equation (100) and B_2^{grad} is a preconditioner for the operator in the equation (101).

The $H(\text{div})$ systems arise naturally from numerous problems of practical importance, such as stabilized mixed formulations of the Stokes problem, least squares methods for $H(\text{grad})$ systems, and mixed methods for $H(\text{grad})$ systems, see [3, 88]. Motivated by [13], we have recently designed a compatible gauge AMG algorithm for $H(\text{div})$ systems in [14], and the numerical experiments demonstrate the efficiency and robustness of this algorithm.

Acknowledgements L. Chen was supported in part by NSF Grant DMS-0811272, and in part by NIH Grant P50GM76516 and R01GM75309. This work is also partially supported by the Beijing International Center for Mathematical Research.

R.H. Nochetto was partially supported by NSF Grant DMS-0807811.

J. Xu was partially by NSF DMS-0609727 and DMS 0749202, and by Alexander von Humboldt Research Award for Senior US Scientists.

References

1. B. Aksoylu and M. Holst. Optimality of multilevel preconditioners for local mesh refinement in three dimensions. *SIAM J. Numer. Anal.*, 44(3):1005–1025, 2006.
2. D. N. Arnold. Differential complexes and numerical stability. *Plenary address delivered at ICM 2002 International Congress of Mathematicians*, 2004.

3. D. N. Arnold, R. S. Falk, and R. Winther. Preconditioning in $H(div)$ and applications. *Math. Comp.*, 66:957–984, 1997.
4. D. N. Arnold, R. S. Falk, and R. Winther. Multigrid in $H(div)$ and $H(curl)$. *Numer. Math.*, 85:197–218, 2000.
5. D. N. Arnold, R. S. Falk, and R. Winther. Finite element exterior calculus, homological techniques, and applications. *Acta Numer.*, pages 1–155, 2006.
6. D. N. Arnold, A. Mukherjee, and L. Pouly. Locally adapted tetrahedral meshes using bisection. *SIAM J. Sci. Comput.*, 22(2):431–448, 2000.
7. D. N. Arnold, L. R. Scott, and M. Vogelius. Regular inversion of the divergence operator with Dirichlet boundary conditions on a polygon. *Ann. Scuola Norm. Sup. Pisa Cl. Sci. (4)*, 15(2):169–192 (1989), 1988.
8. R. E. Bank, A. H. Sherman, and A. Weiser. Refinement algorithms and data structures for regular local mesh refinement. In *Scientific Computing*, pages 3–17. IMACS/North-Holland Publishing Company, Amsterdam, 1983.
9. E. Bänsch. Local mesh refinement in 2 and 3 dimensions. *Impact of Computing in Science and Engineering*, 3:181–191, 1991.
10. R. Beck. Algebraic multigrid by component splitting for edge elements on simplicial triangulations. Technical Report SC 99-40, ZIB, Berlin, Germany, 1999.
11. T. C. Biedl, P. Bose, E. D. Demaine, and A. Lubiw. Efficient algorithms for Petersen's matching theorem. In *Symposium on Discrete Algorithms*, pages 130–139, 1999.
12. P. Binev, W. Dahmen, and R. DeVore. Adaptive finite element methods with convergence rates. *Numer. Math.*, 97(2):219–268, 2004.
13. P. B. Bochev, J. J. Hu, C. M. Siefert, and R. S. Tuminaro. An Algebraic Multigrid Approach Based on a Compatible Gauge Reformulation of Maxwell's Equations. *SIAM J. Sci. Comput.*, 31:557–583, 2008.
14. P. B. Bochev, C. M. Siefert, R. S. Tuminaro, J. Xu, and Y. Zhu. Compatible Gauge Approaches for $H(\mathrm{div})$ Equations. In *SNL-CSRI Proceeding*, 2007.
15. F. A. Bornemann and H. Yserentant. A basic norm equivalence for the theory of multilevel methods. *Numer. Math.*, 64:455–476, 1993.
16. J. H. Bramble. *Multigrid Methods*, volume 294 of *Pitman Research Notes in Mathematical Sciences*. Longman Scientific & Technical, Essex, England, 1993.
17. J. H. Bramble, J. E. Pasciak, and A. H. Schatz. The construction of preconditioners for elliptic problems by substructuring, IV. *Math. Comp.*, 53:1–24, 1989.
18. J. H. Bramble, J. E. Pasciak, J. Wang, and J. Xu. Convergence estimates for product iterative methods with applications to domain decomposition. *Math. Comp.*, 57:1–21, 1991.
19. J. H. Bramble, J. E. Pasciak, and J. Xu. Parallel multilevel preconditioners. *Math. Comp.*, 55(191):1–22, 1990.
20. J. H. Bramble and J. Xu. Some estimates for a weighted L^2 projection. *Math. Comp.*, 56:463–476, 1991.
21. A. Brandt. Multi-level adaptive solutions to boundary-value problems. *Math. Comp.*, 31:333–390, 1977.
22. J. M. Cascón, R. H. Nochetto and K. G. Siebert. Design and Convergence of AFEM in $H(\mathrm{div};\Omega)$. *Math. Models Methods Appl. Sci.*, 17: 1849–1881, 2007.
23. J. M. Cascón, C. Kreuzer, R. H. Nochetto, and K. G. Siebert. Quasi-optimal convergence rate for an adaptive finite element method. *SIAM J. Numer. Anal.*, 46(5):2524–2550, 2008.
24. T. F. Chan and T. P. Mathew. *Domain Decomposition Algorithms*, volume 3 of *Acta Numerica*, pages 61–143. Cambridge University Press, Cambridge, 1994.
25. T. F. Chan and W. L. Wan. Robust multigrid methods for nonsmooth coefficient elliptic linear systems. *J. Comput. Appl. Math.*, 123(1-2):323–352, 2000.
26. L. Chen. *i*FEM: an innovative finite element methods package in MATLAB. *Submitted*, 2009.
27. L. Chen, R. H. Nochetto, and J. Xu. Local multilevel methods on graded bisection grids: H^1 problem. Technical report, University of Maryland at College Park, 2007.
28. L. Chen, R. H. Nochetto, and J. Xu. Local multilevel methods on graded bisection grids: $H(curl$ and $H(div)$ systems. Technical report, University of Maryland at College Park, 2008.

29. L. Chen and C.-S. Zhang. A coarsening algorithm and multilevel methods on adaptive grids by newest vertex bisection. Technical report, Department of Mathematics, University of Maryland at College Park, 2007.
30. D. Cho, J. Xu, and L. Zikatanov. New estimates for the rate of convergence of the method of subspace corrections. *Numer. Math. Theor. Meth. Appl.*, 1:44-56, 2008.
31. W. Dahmen and A. Kunoth. Multilevel preconditioning. *Numer. Math.*, 63:315–344, 1992.
32. R. A. DeVore and G. G. Lorentz. *Constructive Approximation*. Spring-Verlag, New York, NY, 1993.
33. M. Dryja, M. V. Sarkis, and O. B. Widlund. Multilevel Schwarz methods for elliptic problems with discontinuous coefficients in three dimensions. *Numer. Math.*, 72(3):313–348, 1996.
34. M. Dryja, B. F. Smith, and O. B. Widlund. Schwarz analysis of iterative substructuring algorithms for elliptic problems in three dimensions. *SIAM J. Numer. Anal.*, 31, 1994.
35. R. G. Durán and M. A. Muschietti. An explicit right inverse of the divergence operator which is continuous in weighted norms. *Studia Math.*, 148(3):207–219, 2001.
36. R. D. Falgout, J. E. Jones, and U. M. Yang. Pursuing scalability for hypre's conceptual interfaces. *ACM Trans. Math. Software*, 31(3):326–350, 2005.
37. R. D. Falgout, J. E. Jones, and U. M. Yang. The design and implementation of hypre, a library of parallel high performance preconditioners. In *Numerical solution of partial differential equations on parallel computers*, volume 51 of *Lect. Notes Comput. Sci. Eng.*, pages 267–294. Springer, Berlin, 2006.
38. R. D. Falgout and U. M. Yang. Hypre: A Library of High Performance Preconditioners. In *International Conference on Computational Science (3)*, pages 632–641, 2002.
39. V. Girault and P. A. Raviart. *Finite Element Methods for Navier–Stokes Equations*. Springer-Verlag, Berlin, 1986.
40. G. Golub and C. van Loan. *Matrix Computations*. Johns Hopkins University Press, 1996.
41. I. G. Graham and M. J. Hagger. Unstructured additive Schwarz-conjugate gradient method for elliptic problems with highly discontinuous coefficients. *SIAM J. Sci. Comput.*, 20:2041–2066, 1999.
42. M. Griebel and P. Oswald. On additive Schwarz preconditioners for sparse grid discretization. *Numer. Math.*, 66:449–464, 1994.
43. M. Griebel and P. Oswald. On the abstract theory of additive and multiplicative Schwarz methods. *Numer. Math.*, 70:163–180, 1995.
44. W. Hackbusch. *Multigrid Methods and Applications*, volume 4 of *Computational Mathematics*. Springer–Verlag, Berlin, 1985.
45. W. Hackbusch. *Iterative Solution of Large Sparse Systems of Equations*, volume 95 of *Applied Mathematical Sciences*. Springer–Verlag, New York, translated and revised from the 1991 german original edition, 1994.
46. V. E. Henson and U. M. Yang. Boomeramg: a parallel algebraic multigrid solver and preconditioner. *Appl. Numer. Math.*, 41(1):155–177, 2002.
47. R. Hiptmair. Multigrid method for $H(div)$ in three dimensions. *Electron. Trans. Numer. Anal.*, 6:133–152, 1997.
48. R. Hiptmair. Canonical construction of finite elements. *Math. Comp.*, 68:1325–1346, 1999.
49. R. Hiptmair. Multigrid method for Maxwell's equations. *SIAM J. Numer. Anal.*, 36(1):204–225, 1999. unreadable.
50. R. Hiptmair. Finite elements in computational electromagnetism. *Acta Numer.*, pages 237–339, 2002.
51. R. Hiptmair and A. Toselli. Overlapping and multilevel schwarz methods for vector valued elliptic problems in three dimensions. *IMA Volumes in Mathematics and its Applications*, 2000.
52. R. Hiptmair and J. Xu. Nodal auxiliary space preconditioning in H(curl) and H(div) spaces. *SIAM J. Numer. Anal.*, 45(6):2483–2509, 2007.
53. R. Hiptmair and W. Zheng. Local Multigrid in $H(curl)$. *J. Comput. Math.*, 27:573–603, 2009.
54. T. Kolev and P. S. Vassilevski. Some experience with a H1-based auxiliary space AMG for H(curl) problems. Technical Report 221841, LLNL, 2006.

55. T. V. Kolev and P. S. Vassilevski. Parallel Auxiliary Space AMG For H(curl) Problems. *J. Comput. Math.*, 27:604–623, 2009.
56. I. Kossaczký. A recursive approach to local mesh refinement in two and three dimensions. *J. Comput. Appl. Math.*, 55:275–288, 1994.
57. Y. Lee, J. Wu, J. Xu, and L. Zikatanov. Robust subspace correction methods for nearly singular systems. *M3AS*, 17:1937–1963 , 2007.
58. A. Liu and B. Joe. Quality local refinement of tetrahedral meshes based on bisection. *SIAM J. Sci. Comput.*, 16(6):1269–1291, 1995.
59. J. Maubach. Local bisection refinement for *n*-simplicial grids generated by reflection. *SIAM J. Sci. Comput.*, 16(1):210–227, 1995.
60. W. F. Mitchell. *Unified Multilevel Adaptive Finite Element Methods for Elliptic Problems*. PhD thesis, University of Illinois at Urbana-Champaign, 1988.
61. W. F. Mitchell. A comparison of adaptive refinement techniques for elliptic problems. *ACM Trans. Math. Software (TOMS) archive*, 15(4):326 – 347, 1989.
62. W. F. Mitchell. Optimal multilevel iterative methods for adaptive grids. *SIAM J. Sci. Stat. Comp.*, 13:146–167, 1992.
63. R. Nochetto, K. Siebert, and A. Veeser. Theory of adaptive finite element methods: an introduction. In R. A. DeVore and A. Kunoth, editors, *Multiscale, Nonlinear and Adaptive Approximation*. Springer, 2009.
64. P. Oswald. On discrete norm estimates related to multilevel preconditioners in the finite element method. In *Constructive Theory of Functions, Proc. Int. Conf. Varna 1991*, pages 203–214, Sofia, 1992. Bulg. Acad. Sci.
65. P. Oswald. *Multilevel Finite Element Approximation, Theory and Applications*. Teubner Skripten zur Numerik. Teubner Verlag, Stuttgart, 1994.
66. P. Oswald. On the robustness of the BPX-preconditioner with respect to jumps in the coefficients. *Math. Comp.*, 68:633–650, 1999.
67. A. Plaza and G. F. Carey. Local refinement of simplicial grids based on the skeleton. *Appl. Numer. Math.*, 32(2):195–218, 2000.
68. A. Plaza and M.-C. Rivara. Mesh refinement based on the 8-tetrahedra longest-edge partition. *12th meshing roundtable*, pages 67–78, 2003.
69. M. C. Rivara. Mesh refinement processes based on the generalized bisection of simplexes. *SIAM J. Numer. Anal.*, 21:604–613, 1984.
70. M.-C. Rivara and P. Inostroza. Using longest-side bisection techniques for the automatic refinement fo Delaunay triangulations. *Int. J. Numer. Methods. Eng.*, 40:581–597, 1997.
71. M. C. Rivara and G. Iribarren. The 4-triangles longest-side partition of triangles and linear refinement algorithms. *Math. Comp.*, 65(216):1485–1501, 1996.
72. M. C. Rivara and M. Venere. Cost analysis of the longest-side (triangle bisection) refinement algorithms for triangulations. *Engineering with Computers*, 12:224–234, 1996.
73. J. W. Ruge and K. Stüben. Algebraic multigrid (AMG). In S. F. McCormick, editor, *Multigrid Methods*, volume 3 of *Frontiers in Applied Mathematics*, pages 73–130. SIAM, Philadelphia, PA, 1987.
74. A. Schmidt and K. G. Siebert. *Design of Adaptive Finite Element Software*, volume 42 of *Lecture Notes in Computational Science and Engineering*. Springer-Verlag, Berlin, 2005. The finite element toolbox ALBERTA, With 1 CD-ROM (Unix/Linux).
75. R. Scott and S. Zhang. Finite element interpolation of nonsmooth functions satisfying boundary conditions. *Math. Comp.*, 54:483–493, 1990.
76. E. G. Sewell. Automatic generation of triangulations for piecewise polynomial approximation. In *Ph. D. dissertation*. Purdue Univ., West Lafayette, Ind., 1972.
77. B. F. Smith. A domain decomposition algorithm for elliptic problems in three dimensions. *Numer. Math.*, 60:219–234, 1991.
78. R. Stevenson. Stable three-point wavelet bases on general meshes. *Numer. Math.*, V80(1):131–158, 1998.
79. R. Stevenson. Optimality of a standard adaptive finite element method. *Found. Comput. Math.*, 7(2):245–269, 2007.

80. R. Stevenson. The completion of locally refined simplicial partitions created by bisection. *Math. Comp.*, 77:227–241, 2008.
81. K. Stüben. An introduction to algebraic multigrid. In U. Trottenberg, C. Oosterlee, and A. Schüller, editors, *Multigrid*, pages 413–528. Academic Press, San Diego, CA, 2001.
82. A. Toselli and O. Widlund. *Domain Decomposition Methods: Algorithms and Theory*. Springer Series in Computational Mathematics, 2005.
83. C. T. Traxler. An algorithm for adaptive mesh refinement in n dimensions. *Computing*, 59(2):115–137, 1997.
84. C. Vuik, A. Segal, and J. A. Meijerink. An efficient preconditioned cg method for the solution of a class of layered problems with extreme contrasts in the coefficients. *J. Comput. Phys.*, 152(1):385–403, June 1999.
85. J. Wang. New convergence estimates for multilevel algorithms for finite-element approximations. *J. Comput. Appl. Math.*, 50:593–604, 1994.
86. J. Wang and R. Xie. Domain decomposition for elliptic problems with large jumps in coefficients. In *the Proceedings of Conference on Scientific and Engineering Computing*, pages 74–86. National Defense Industry Press, 1994.
87. O. B. Widlund. Some Schwarz methods for symmetric and nonsymmetric elliptic problems. In D. E. Keyes, T. F. Chan, G. A. Meurant, J. S. Scroggs, and R. G. Voigt, editors, *Fifth International Symposium on Domain Decomposition Methods for Partial Differential Equations*, pages 19–36, Philadelphia, 1992. SIAM.
88. B. I. Wohlmuth, A. Toselli, and O. B. Widlund. An iterative substructuring method for Raviart–Thomas vector fields in three dimensions. *SIAM J. Numer. Anal.*, 37(5):1657–1676, 2000.
89. H. Wu and Z. Chen. Uniform convergence of multigrid v-cycle on adaptively refined finite element meshes for second order elliptic problems. *Science in China: Series A Mathematics*, 49(1):1–28, 2006.
90. J. Xu. Counter examples concerning a weighted L^2 projection. *Math. Comp.*, 57:563–568, 1991.
91. J. Xu. Iterative methods by space decomposition and subspace correction. *SIAM Rev.*, 34:581–613, 1992.
92. J. Xu. The auxiliary space method and optimal multigrid preconditioning techniques for unstructured meshes. *Computing*, 56:215–235, 1996.
93. J. Xu and Y. Zhu. Uniform convergent multigrid methods for elliptic problems with strongly discontinuous coefficients. *M3AS*, 2007.
94. J. Xu and L. Zikatanov. The method of alternating projections and the method of subspace corrections in Hilbert space. *J. Amer. Math. Soc.*, 15:573–597, 2002.
95. J. Xu and J. Zou. Some nonoverlapping domain decomposition methods. *SIAM Rev.*, 40(4):857–914, 1998.
96. H. Yserentant. Old and new convergence proofs for multigrid methods. *Acta Numer.*, pages 285–326, 1993.
97. Y. Zhu. Domain decomposition preconditioners for elliptic equations with jump coefficients. *Numer. Linear Algebra Appl.*, 15:271–289, 2008.
98. L. Zikatanov. Two-sided bounds on the convergence rate of two-level methods. *Numer. Linear Algebra Appl.*, 15(5): 439-454, 2008.