# AN ODYSSEY INTO LOCAL REFINEMENT AND MULTILEVEL PRECONDITIONING I: OPTIMALITY OF THE BPX PRECONDITIONER *

BURAK AKSOYLU† AND MICHAEL HOLST‡

**Abstract.** In this article, we examine the Bramble-Pasciak-Xu (BPX) preconditioner in the setting of local 2D and 3D mesh refinement. While the available optimality results for the BPX preconditioner have been constructed primarily in the setting of uniformly refined meshes, a notable exception is the 2D result due to Dahmen and Kunoth, which established BPX optimality on meshes produced by a restricted class of local 2D red-green refinement. The purpose of this article is to extend the original 2D Dahmen-Kunoth result to several additional types of local 2D and 3D red-green (conforming) and red (non-conforming) refinement procedures. The extensions are accomplished through a 3D extension of the 2D framework in the original Dahmen-Kunoth work, by which the question of optimality is reduced to establishing that locally enriched finite element subspaces allow for the construction of a scaled basis which is formally Riesz stable. This construction in turn rests entirely on establishing a number of geometrical properties between neighboring simplices produced by the local refinement algorithms. These properties are then used to build Riesz-stable scaled bases for use in the BPX optimality framework. Since the theoretical framework supports arbitrary spatial dimension $d \geq 1$, we indicate clearly which geometrical properties, established here for several 2D and 3D local refinement procedures, must be re-established to show BPX optimality for spatial dimension $d \geq 4$. Finally, we also present a simple alternative optimality proof of the BPX preconditioner on quasiuniform meshes in two and three spatial dimensions, through the use of $K$-functionals and $H^s$-stability of $L_2$-projection for $s \geq 1$. The proof techniques we use are quite general; in particular, the results require no smoothness assumptions on the PDE coefficients beyond those required for well-posedness in $H^1$, and the refinement procedures may produce nonconforming meshes.

**Key words.** finite element approximation theory, multilevel preconditioning, BPX, two and three dimensions, local mesh refinement, red and red-green refinement.

**AMS subject classifications.** 65M55, 65N55, 65N22, 65F10

## 1. Introduction.

In this article, we analyze the impact of local adaptive mesh refinement on the stability of multilevel finite element spaces and on the optimality (linear space and time complexity) of multilevel preconditioners. Adaptive refinement techniques have become a crucial tool for many applications, and access to optimal or near-optimal multilevel preconditioners for locally refined mesh situations is of primary concern to computational scientists. The preconditioners which can be expected to have somewhat favorable space and time complexity in such local refinement scenarios are the hierarchical basis (HB) method, the Bramble-Pasciak-Xu (BPX) preconditioner, and the wavelet modified (or stabilized) hierarchical basis (WHB) method. While there are optimality results for both the BPX and WHB preconditioners in the literature, these are primarily for quasiuniform meshes and/or two space dimensions (with some exceptions noted below). In particular, there are few hard results in the literature on the optimality of these methods for various real-

*The first author was supported in part by the Burroughs Wellcome Fund through the LJIS predoctoral training program at UC San Diego, in part by NSF (ACI-9721349, DMS-9872890), and in part by DOE (W-7405-ENG-48/B341492). Other support was provided by Intel, Microsoft, Alias|Wavefront, Pixar, and the Packard Foundation. The second author was supported in part by NSF (CAREER Award DMS-9875856 and standard grants DMS-0208449, DMS-9973276, DMS-0112413), in part by DOE (SCI-DAC-21-6993), and in part by a Hellman Fellowship.

† Department of Computer Science, California Institute of Technology, Pasadena, CA 91125, USA (baksoylu@cs.caltech.edu).

‡ Department of Mathematics, University of California at San Diego, La Jolla, CA 92093, USA (mholst@math.ucsd.edu).

istic local mesh refinement hierarchies, especially in three space dimensions. In this article, the first in a series of three articles [2, 3] on local refinement and multilevel preconditioners, we assemble a number of such optimality results for the BPX preconditioner in local refinement scenarios, in both two and three space dimensions. (The material forming this trilogy is based on the first author's Ph.D. dissertation [1].) The second article [3] builds on the BPX results we present here to develop some analogous optimality results for the WHB method in local refinement settings. The main results in both articles are valid for any spatial dimension $d \geq 1$, for nonsmooth PDE coefficients $p \in L_\infty(\Omega)$, and allow for nonconforming meshes.

The problem class we focus on here and in [3] is linear second order partial differential equations (PDE) of the form:

$$-\nabla \cdot (p\ \nabla u) + q\ u = f, \quad u \in H_0^1(\Omega). \tag{1.1}$$

Here, $f \in L_2(\Omega)$, $p, q \in L_\infty(\Omega)$, $p : \Omega \to L(\mathbb{R}^d, \mathbb{R}^d)$, $q : \Omega \to \mathbb{R}$, where $p$ is a symmetric positive definite matrix function, and where $q$ is a nonnegative function. Let $\mathcal{T}_0$ be a shape regular and quasiuniform initial partition of $\Omega$ into a finite number of $d$ simplices, and generate $\mathcal{T}_1, \mathcal{T}_2, \ldots$ by refining the initial partition using either red-green or red local refinement strategies in $d = 2$ or $d = 3$ spatial dimensions. Denote as $\mathcal{S}_j$ the simplicial linear $C^0$ finite element space corresponding to $\mathcal{T}_j$ equipped with zero boundary values. The set of nodal basis functions for $\mathcal{S}_j$ is denoted by $\Phi^{(j)} = \{\phi_i^{(j)}\}_{i=1}^{N_j}$ where $N_j = \dim\ \mathcal{S}_j$ is equal to the number of interior nodes in $\mathcal{T}_j$, representing the number of degrees of freedom in the discrete space. Successively refined finite element spaces will form the following nested sequence:

$$\mathcal{S}_0 \subset \mathcal{S}_1 \subset \ldots \subset \mathcal{S}_j \subset \ldots \subset H_0^1(\Omega). \tag{1.2}$$

Although the mesh is nonconforming in the case of red refinement, $\mathcal{S}_j$ is used within the framework of conforming finite element methods for discretizing (1.1).

Let the bilinear form and the functional associated with the weak formulation of (1.1) be denoted as

$$a(u, v) = \int_\Omega p\ \nabla u \cdot \nabla v + q\ u\ v\ dx, \quad b(v) = \int_\Omega f\ v\ dx, \quad u, v \in H_0^1(\Omega).$$

We consider primarily the following Galerkin formulation: Find $u \in \mathcal{S}_j$, such that

$$a(u, v) = b(v), \quad \forall v \in \mathcal{S}_j. \tag{1.3}$$

The finite element approximation in $\mathcal{S}_j$ has the form $u^{(j)} = \sum_{i=1}^{N_j} u_i \phi_i^{(j)}$, where $u = (u_1, \ldots, u_{N_j})^T$ denotes the coefficients of $u^{(j)}$ with respect to $\Phi^{(j)}$. The resulting *discretization operator* $A^{(j)} = \{a(\phi_k^{(j)}, \phi_l^{(j)})\}_{k,l=1}^{N_j}$ determines the interaction of basis functions with respect to $a(\cdot, \cdot)$ and must be inverted numerically to determine the coefficients $u$ from the linear system:

$$A^{(j)} u = F^{(j)}, \tag{1.4}$$

where $F^{(j)} = \{b(\phi_l^{(j)})\}_{l=1}^{N_j}$. Our task is to solve (1.4) with optimal (linear) complexity in both storage and computation, where the finite element spaces $\mathcal{S}_j$ are built on locally refined meshes. The condition number $\kappa_{\Phi^{(j)}}(A^{(j)})$ of $A^{(j)}$ with respect to the chosen basis $\Phi^{(j)}$ provides an upper bound on the number of iterations required by conjugate gradient-type methods to produce an approximate solution (satisfying a given fixed tolerance) to a linear system involving $A^{(j)}$. Therefore, it is desirable to have an analysis framework for bounding the condition number produced by a given basis, with the goal of finding bases which produce uniformly bounded condition numbers (or at least condition numbers with slow growth).

The primary idea behind BPX and related multilevel preconditioners is the notion of a *stable* splitting of $u \in \mathcal{S}_J$,

$$u = \sum_{j=0}^{J} (\pi_j - \pi_{j-1}) u, \tag{1.5}$$

by linear operators $\pi_j : L_2 \to \mathcal{S}_j$ such that $\pi_j |_{\mathcal{S}_j} = I$ and $\pi_j \pi_k = \pi_{\min\{j,k\}}$, where $\pi_{-1} = 0$. The splitting (1.5) or equivalently the hierarchical ordering of the nodes gives rise to a hierarchical splitting

$$\mathcal{S}_j = \mathcal{S}_{j-1} \oplus \mathcal{S}_j^f, \tag{1.6}$$

where we will refer to $\mathcal{S}_j^f$ as the *slice space*, and where the superscript $f$ stands for *fine*. The slice space $\mathcal{S}_j^f$ is selected as a hierarchical complement of $\mathcal{S}_{j-1}$ in $\mathcal{S}_j$, namely $\mathcal{S}_j^f = (\pi_j - \pi_{j-1})\mathcal{S}_j$. By the properties of $\pi_j$ listed above, the two level direct decomposition becomes (1.6)

$$\mathcal{S}_j = \mathcal{S}_{j-1} \oplus (I - \pi_{j-1})\mathcal{S}_j. \tag{1.7}$$

The particular splitting (1.5) defines a preconditioner $B^{(J)}$ through

$$(B^{(J)} u, v) = \sum_{j=0}^{J} 2^{2jk} ((\pi_j - \pi_{j-1})u, (\pi_j - \pi_{j-1})v), \quad u, v \in \mathcal{S}_j, k = 1, \tag{1.8}$$

where $k$ is the smoothness parameter to accommodate higher order PDEs, and where the factors $2^{2j}$ are dimension independent and simply reflect the dyadic refinement. If $B^{(J)}$ is computationally feasible and is spectrally equivalent to $A(J)$:

$$\lambda_{B^{(J)}} (B^{(J)} u, u) \leq (A^{(J)} u, u) \leq \Lambda_{B^{(J)}} (B^{(J)} u, u), \tag{1.9}$$

then the efficiency of the preconditioner will be determined by $\frac{\Lambda_{B^{(J)}}}{\lambda_{B^{(J)}}}$ since the condition number then satisfies $\kappa({B^{(J)}}^{-1} A^{(J)}) \leq \frac{\Lambda_{B^{(J)}}}{\lambda_{B^{(J)}}}$. The spectral equivalence bounds $\Lambda_{B^{(J)}}$ and $\lambda_{B^{(J)}}$ are produced by analyzing the *preconditioner norm*:

$$\|u\|_{B^{(J)}}^2 \equiv \sum_{j=0}^{J} 2^{2j} \|(\pi_j - \pi_{j-1})u\|_{L_2}^2, \tag{1.10}$$

and in particular, by establishing that it is norm equivalent to the $H^1$-norm:

$$c_1 \ \|u\|_{B^{(J)}}^2 \leq \|u\|_{H^1}^2 \leq c_2 \ \|u\|_{B^{(J)}}^2. \tag{1.11}$$

Here, we should clarify that *stable* splitting [23, 24] means that the corresponding preconditioner will have favorable $\lambda_{B^{(J)}}$ and $\Lambda_{B^{(J)}}$, and in the best case, *optimal* bounds (meaning absolute constants).

The two level direct decomposition (1.7) and the choice of the finite element interpolation $I_j$ as $\pi_j$ was the introduction of the *hierarchical basis* (HB) methods [4, 5, 33, 35]. In local refinement settings, HB methods enjoy an optimal complexity of $O(N_j - N_{j-1})$ per iteration per level (resulting in $O(N_J)$ overall complexity) by employing only degrees of freedom corresponding to $\mathcal{S}_j^f$ by the virtue of (1.7). However, condition number of the HB preconditioner is not uniformly bounded, and while it grows only slowly in 2D, the growth is rapid in 3D. The insight of Bramble, Pasciak, and Xu in [11] was to replace $I_j$ with $Q_j$, the $L_2$-projection onto $\mathcal{S}_j$, producing the famed BPX preconditioner. While it was eventually shown that the original BPX preconditioner was optimal in uniform refinement settings [22, 23, 32], later variations [7, 9, 14, 24] also allowed for certain restricted types of local refinement. It should be noted that an alternative to the BPX preconditioner is the direct stabilization of the hierarchical basis due to Vassilevski and Wang [29, 30] and to Stevenson [27]; these techniques are discussed in detail in the companion paper [3]. The motivation behind these methods is that the BPX decomposition, $\mathcal{S}_j = \mathcal{S}_{j-1} \oplus (Q_j - Q_{j-1})\mathcal{S}_j$, gives rise to globally supported basis functions but with rapid decay, allowing for locally supported approximations [18].

Our primary interest here is the original BPX preconditioner, where the BPX norm is defined as the preconditioner norm (1.10) using $L^2$-projection:

$$\|u\|_{\text{BPX}}^2 \equiv \sum_{j=0}^{J} 2^{2j} \|(Q_j - Q_{j-1})u\|_{L_2}^2. \tag{1.12}$$

The best case in (1.9) would be to obtain absolute constants $c_1$, $c_2$ in the norm equivalence (1.11), so that the BPX preconditioner becomes optimal. In the original work [11] the following suboptimal result was established with regularity assumptions:

$$\frac{c_1}{J}\ \|u\|_{\text{BPX}}^2 \le \|u\|_{H^1}^2 \le c_2 J\ \|u\|_{\text{BPX}}^2. \tag{1.13}$$

There have been numerous contributions since the original work toward achieving optimal upper and lower bounds in (1.11) under various assumptions, but primarily in the uniform refinement setting [7, 32, 34, 37]. Oswald [22, 23] first proved the optimality result for both two- and three-dimensional problems in the quasiuniform setting using Besov space techniques, relying on the fact that the Sobolev space $H^1$ coincides with certain Besov spaces, and as a result the corresponding norms are equivalent. The two primary results on optimality of BPX in local refinement settings are due to Dahmen and Kunoth [14] and Bornemann and Yserentant [7]. Both works consider only two space dimensions, and in particular, the refinement strategies analyzed are restricted 2D red-green refinement and 2D red refinement, respectively. In this paper, we extend the framework developed in [14] to several types of practical, implementable, local red-green (conforming) and red (non-conforming) refinement procedures in both two and three space dimensions, and establish that the BPX preconditioner is optimal for the resulting finite element hierarchy.

***Outline of the paper.*** In §2, we outline some basic approximation theory tools for relating Besov and approximation spaces, such as *Jackson* and *Bernstein estimates*. In §3 we outline the theoretical framework used by Dahmen and Kunoth, including the norm equivalence results to be established in local refinement scenarios. In §4 and §5, we list the refinement conditions and additional rules for tetrahedralization and triangulation for $d = 3$ and $d = 2$, respectively. We give several theorems about the generation and size relations of the neighboring simplices, thereby establishing quasiuniformity for the simplices in the support of a basis function. In §6, we give some additional generation bounds for the red refinement scheme. In §7, we use these geometrical results to extend the 2D Dahmen-Kunoth results to the 2D and 3D local refinement procedures of interest here, establishing the desired norm equivalence (1.11). The geometrical properties established in §4, §5, and §6 lead to quasiuniformity of the support which gives rise to an $L_2$-stable Riesz basis; one can then establish the Bernstein estimate. While it is not possible to establish a Jackson estimate due to the nature of local adaptivity, in §8 the remaining inequality in the norm equivalence is handled directly using approximation theory tools, as in the original work [14]. In §9 we give an extremely simple BPX optimality proof for quasiuniform meshes using $K$-functionals and $H^s$-stability of $L_2$-projection for $s > 1$. Table 1.1 encapsulates the optimality results we establish in this article.

Table 1.1
*Collection of existing results and the ones proved in this article for the optimality of the BPX preconditioner. r and r-g stand for red and red-green respectively.*

| Refinement | 3D r | 2D r | 2D r in [7] | 3D r-g | 2D r-g |
|---|---|---|---|---|---|
| Reference | This article | This article | [7] | This article | [14] |
| PDE coefficient | $p \in L_\infty$ | $p \in L_\infty$ | $p \in C^1$ | $p \in L_\infty$ | $p \in L_\infty$ |

As a final remark, we note that the question of $H^s$-stability of $L_2$-projection (primarily for $s = 1$) onto finite element spaces built through various local refinement schemes is currently under intensive study in the finite element community due to its relationship to multilevel preconditioning. The existing theoretical results, due primarily to Carstensen [13] and Bramble-Pasciak-Steinbach [10] involve *a posteriori* verification of somewhat complicated mesh conditions after local refinement has taken place. If such mesh conditions are not satisfied, one has to redefine the mesh. However, an interesting consequence of the BPX optimality results for locally refined 2D and 3D meshes established here is $H^1$-stability of $L_2$-projection restricted to the same locally enriched finite element spaces. This result, which is established in [3] based on the results here, appears to be the first *a priori* $H^1$-stability result for $L_2$-projection on finite element spaces produced by practical and easily implementable 2D and 3D local refinement algorithms.

**2. Some approximation theory background.** Let $\Omega \subset \mathbb{R}^d$ be open, for arbitrary $k = 1, 2, \ldots$ and $h \in \mathbb{R}^d$, and define the subset

$$\Omega_{k,h} = \left\{ x \in \mathbb{R}^d : [x, x + kh] \subset \Omega \right\}.$$

Define the directional $k$-th order difference of $f \in L_p(\Omega)$ as

$$\left(\Delta_h^k f\right)(x) = \sum_{r=0}^{k} \left( \begin{array}{c} k \\ r \end{array} \right) (-1)^{k-r} f(x + rh), \quad x, h \in \mathbb{R}^d.$$

A finer scale of smoothness than differentiability, *modulus of smoothness*, is a central tool in the analysis.

Definition 2.1. *$L_p$-modulus of smoothness*

$$\omega_k(f, t, \Omega)_p = \sup_{|h| \le t} \|\Delta_h^k f\|_{L_p(\Omega_{k,h})}.$$

The Besov space and the approximation space are the two crucial spaces borrowed from approximation theory for norm estimates. Besov space becomes the primary function space setting by realizing the Sobolev space as a Besov space:

$$H^s(\Omega) \cong B^s_{2,2}(\Omega), \quad s > 0.$$

The analysis needed for functions in Sobolev space is done in the Besov space. The primary motivation for employing the Besov space stems from the fact that the characterization of functions which have a given upper bound for the error of approximation sometimes calls for a finer scale of smoothness. *Besov spaces* are defined to be the collection of functions $f \in L_p(\Omega)$ with a finite Besov norm defined as follows:

$$\|f\|_{B^s_{p,q}(\Omega)} = \|f\|_{L_p(\Omega)} + |f|_{B^s_{p,q}(\Omega)},$$

where the seminorm is given by

$$|f|_{B^s_{p,q}(\Omega)} = \|\{2^{sj}\omega_k(f, 2^{-j}, \Omega)_p\}_{j\in\mathbb{N}_0}\|_{l_q},$$

and $k$ is any fixed integer larger than $s$.

The functions with finite BPX norm (1.12) also form a function space. This class corresponds to a general class in approximation theory which is referred to as the *approximation space*. It is the collection of functions $f \in L_p(\Omega)$ with a finite approximation space norm defined as follows.

$$\|f\|_{A^s_{p,q}(\Omega)} = \|\{2^{sj}\|(Q_j - Q_{j-1})f\|_{L_p(\Omega)}\}_{j\in\mathbb{N}_0}\|_{l_q}.$$

The *Bernstein* and the *Jackson* estimates are the fundamental estimates utilized in the analysis. The Bernstein estimate describes smoothness (or differentiability if available) of the functions in $\mathcal{S}_J$.

DEFINITION 2.2. *A Bernstein estimate has the form:*

$$\omega_{k+1}(u, t)_p \le c\ (\min\{1, t2^J\})^\beta \|u\|_{L_p}, \quad u \in \mathcal{S}_J, \tag{2.1}$$

*where $c$ is independent of $u$ and $J$. Usually $k$ = degree of the element and $\beta > k$.* On the other hand, both the Jackson estimate and the approximation inequality characterize the approximation error in a coarser (or weaker) norm than the native norm on the space containing the given function. For a given subset $M \subset L_p(\Omega)$, consider the best approximation:

$$E_M(f)_p = \inf_{g\in M} \|f - g\|_{L_p}, \quad f \in L_p(\Omega).$$

We are interested in getting estimates from above for $E_M(f)_p$ when $M$ coincides with a finite element subspace. In the Besov space setting, $E$ is characterized by modulus of smoothness, and if a Sobolev space, $W^\alpha_p$, setting is available, then it can also be characterized by the Sobolev norm. This gives rise to different versions of the Jackson estimate.

DEFINITION 2.3. *The Jackson estimate for Besov spaces is defined as follows:*

$$E_{\mathcal{S}_J}(f)_p \le c\ \omega_\alpha(f, 2^{-J})_p, \quad f \in L_p(\Omega). \tag{2.2}$$

*The analog of (2.2) in Sobolev spaces is called the best approximation inequality:*

$$E_{\mathcal{S}_J}(f)_p \le c\ (2^{-J})^\alpha |f|_{W^\alpha_p}, \quad f \in W^\alpha_p(\Omega), \tag{2.3}$$

*where $c$ is a constant independent of $f$ and $J$, and $\alpha$ is an integer.*

The following norm equivalence (a part of Theorem 3.1) will be the primary step in order to establish the main norm equivalence (1.11).

$$\frac{c_1}{v_J^{(p)}}\|u\|_{A^s_{p,q}} \leq \|u\|_{B^s_{p,q}} \leq c_2\|u\|_{A^s_{p,q}}, \quad u \in \mathcal{S}_J. \tag{2.4}$$

The Bernstein estimates (2.1) are used for the upper bound and also dictates what needs to be done for the lower bound. The subtle term $v_J^{(p)}$ in (2.4) involves the error of the best approximation. This is precisely where the Jackson estimate comes into play in order to guarantee an absolute constant as the lower bound. Hence the Bernstein estimates and the Jackson (approximation) estimates complement each other for the purpose of achieving the norm equivalence (1.11).

Furthermore, if these estimates can be established, then one reaches an isomorphism between the approximation and Besov spaces. In particular, the embedding $A^s_{p,q} \hookrightarrow B^s_{p,q}$ becomes an implication of the Bernstein estimate (2.1). This is usually referred to as the upper bound in the norm equivalence (2.4). The Jackson estimate implies the embeddings in the opposite direction, $B^s_{p,q} \hookrightarrow A^s_{p,q}$, which establishes the lower bound of (1.11).

**3. Norm equivalences and optimality of the BPX preconditioner.** The work of Dahmen and Kunoth [14] established a theoretical framework for analyzing the optimality of BPX for 2D local refinement. Their work was in fact the only one to date which carefully examined a realistic, implementable type of 2D red-green refinement and rigorously established optimality results. Our analysis for the 3D case follows their outline for the 2D case. The focus of the analysis in the norm equivalence is the quantity $v_J^{(p)}$ defined as follows.

$$v_{j,J}^{(p)} = \sup_{u\in\mathcal{S}_J} \frac{\|u - Q_j u\|_{L_p}}{\omega_{k+1}(u, 2^{-j}, \Omega)_p}, \quad v_J^{(p)} = \max\left\{1, v_{j,J}^{(p)} : j = 0, \ldots, J\right\}. \tag{3.1}$$

This quantity can be seen as a characterization of how well one can reach the Jackson estimate (2.2) by using functions from $\mathcal{S}_J$. The target norm equivalence between the Besov space $B^1_{2,2}$ and the approximation space $A^1_{2,2}$ is established as we collect all the results given in [14] related to the BPX preconditioner under one big theorem as follows.

Theorem 3.1. *Suppose the Bernstein estimate (2.1) holds for some real number $\beta > k$. Then*

- *For each $0 < s < \min\{\beta, k + 1\}$, there exist constants $0 < c_1,\ c_2 < \infty$ independent of $u \in \mathcal{S}_J,\ J = 0, 1, \ldots,$ such that (2.4) holds:*

$$\frac{c_1}{v_J^{(p)}}\|u\|_{A^s_{p,q}} \leq \|u\|_{B^s_{p,q}} \leq c_2\|u\|_{A^s_{p,q}}, \quad u \in \mathcal{S}_J.$$

- *The following condition number estimate holds for the BPX preconditioner:*

$$\kappa(B^{(J)^{-1/2}} A^{(J)} B^{(J)^{-1/2}}) = O((v_J^{(2)})^2), \quad J \to \infty. \tag{3.2}$$

- *If in addition the Jackson estimate (2.3) holds for an integer $\alpha > k$ and $p = 2$, then*

$$v_J^{(p)} = O(1), \quad J \to \infty, \tag{3.3}$$

*holds, and as a result the condition number is optimal:*

$$\kappa(B^{(J)^{-1/2}} A^{(J)} B^{(J)^{-1/2}}) = O(1), \quad J \to \infty. \tag{3.4}$$

*Proof.* See Theorem 4.1, Proposition 4.1, Theorem 3.1, and Theorem 3.2 in [14], respectively. ▯

Due to the fact that the Jackson estimate cannot hold for local refinement, $v_J^{(p)}$ must be estimated directly (see §8 for details). In the sections to follow, the emphasis is on extending optimality results to local refinement. We extend the 2D red-green refinement result in [14] to an analogous 3D red-green refinement procedure. We give precise bounds on the size and generation relations between neighboring simplices in the tetrahedralization. Motivated by applications in computer graphics and animation, we also study 2D and 3D versions of the red refinement. The red refinement produces nonconforming meshes. The same analysis and proof technique done for red-green refinement naturally carries over to red refinement without violating any required assumptions. As a result, we can employ $v_J^{(p)}$ in the analysis both for conforming and nonconforming settings.

**4. 3D red-green refinement.** We turn our attention to the optimality of the BPX preconditioner for locally refined meshes in 3D. Dahmen and Kunoth [14] established an optimality statement for the BPX preconditioner using the classical *red-green* refinement procedure in 2D. We prove that the BPX preconditioner still preserves its optimality for an analog of this procedure in 3D. In fact, it is the natural extension of red-green refinement to 3D. The details of the procedure can be found in [6].

We first list a number of geometric assumptions we make concerning the underlying mesh. (The two-dimensional case is quite standard, so we only describe the more complicated three-dimensional case here.) Let $\Omega \subset \mathbb{R}^3$ be a polyhedral domain. We assume that the triangulation $\mathcal{T}_j$ of $\Omega$ at level $j$ is a collection of tetrahedra with mutually disjoint interiors which cover $\Omega = \bigcup_{\tau \in \mathcal{T}_j} \tau$. We want to generate successive refinements $\mathcal{T}_0, \mathcal{T}_1, \ldots$ which satisfy the following conditions:

ASSUMPTION 4.1. **Nestedness:** *Each tetrahedron (son) $\tau \in \mathcal{T}_j$ is covered by exactly one tetrahedron (father) $\tau' \in \mathcal{T}_{j-1}$, and any corner of $\tau$ is either a corner or an edge midpoint of $\tau'$.*

ASSUMPTION 4.2. **Conformity:** *The intersection of any two tetrahedra $\tau, \tau' \in \mathcal{T}_j$ is either empty, a common vertex, a common edge or a common face.*

ASSUMPTION 4.3. **Nondegeneracy:** *The interior angles of all tetrahedra in the refinement sequence $\mathcal{T}_0, \mathcal{T}_1, \ldots$ are bounded away from zero.*

A regular (red) refinement subdivides a tetrahedron $\tau$ into 8 equal volume subtetrahedra. We connect the edges of each face as in 2D regular refinement. We then cut off four subtetrahedra at the corners which are congruent to $\tau$. An octahedron with three parallelograms remains in the interior. Cutting the octahedron along the two faces of these parallelograms, we obtain four more subtetrahedra which are not necessarily congruent to $\tau$. We choose the diagonal of the parallelogram so that the successive refinements always preserve nondegeneracy [1, 6, 21, 36].

If a tetrahedron is marked for regular refinement, the resulting triangulation violates conformity A.4.2. Nonconformity is then remedied by irregular (green) refinement. In 3D, there are altogether $2^6 = 64$ possible edge refinements, of which 62 are irregular. One must pay extra attention to irregular refinement in the implementation due to the large number of possible nonconforming configurations. Bey [6] gives a methodical way of handling irregular cases. Using symmetry arguments, the 62 irregular cases can be divided into 9 different types. To ensure that the interior angles remain bounded away from zero, we enforce these additional conditions: (Similar assumptions were made in [14] in 2D analogue)

ASSUMPTION 4.4. *Irregular tetrahedra are not refined further.*

ASSUMPTION 4.5. *Only tetrahedra $\tau \in \mathcal{T}_j$ with $L(\tau) = j$ are refined for the construction of $\mathcal{T}_{j+1}$, where $L(\tau) = \min\{j : \tau \in \mathcal{T}_j\}$ denotes the level of $\tau$.*

One should note that the restrictive character of A.4.4 and A.4.5 can be eliminated by a modification on the sequence of the tetrahedralizations [6]. On the other hand, it is straightforward to enforce both assumptions in a typical local refinement algorithm by minor modifications of the supporting datastructures for tetrahedral elements (cf. [17]). In any event, the proof technique (see (8.7) and (8.8)) requires both assumptions hold. The last refinement condition enforced for the possible 62 irregularly refined tetrahedra is stated as the following.

ASSUMPTION 4.6. *If three or more edges are refined and do not belong to a common face, then the tetrahedron is refined regularly.* The *generation bound* for simplices in $\mathbb{R}^d$ which are adjacent on $d$ many vertices is the major result required in the support of a basis function so that (7.1) holds. The generation bound for simplices which are adjacent on $d-1, d-2, \dots$ many vertices follows by using the shape regularity and the generation bound established for $d$-vertex adjacency. We provide rigorous generation bound proofs for all the adjacency types mentioned in the lemmas to follow when $d = 3$.

LEMMA 4.1. *Let $\tau$ and $\tau'$ be two tetrahedra in $\mathcal{T}_j$ sharing a common face $f$. Then*

$$|L(\tau) - L(\tau')| \leq 1. \tag{4.1}$$

*Proof.* (The original proof may be found in [1]; the 2D version appeared in [14]. 3D extension works without additional framework.) If $L(\tau) = L(\tau')$, then $0 \leq 1$, there is nothing to show. Without loss of generality, assume that $L(\tau) < L(\tau')$. Proof requires a detailed and systematic analysis. To show the line of reasoning, we first list the facts used in the proof:

1. $L(\tau') \leq j$ because by assumption $\tau' \in \mathcal{T}_j$.
2. $L(\tau) < j$ $(L(\tau) < L(\tau') \leq j)$.
3. By assumption $\tau \in \mathcal{T}_j$, meaning that $\tau$ was never refined from the level it was born $L(\tau)$ to level $j$.
4. Let $\tau''$ be the father of $\tau'$. Then $L(\tau'') = L(\tau') - 1 < j$.
5. $L(\tau) < L(\tau')$ by assumption, implying $L(\tau) \leq L(\tau'')$.
6. By (3), $\tau$ belongs to all the triangulations from $L(\tau)$ to $j$, in particular $\tau \in \mathcal{T}_{L(\tau'')}$, where by (4) $L(\tau'') < j$.

$f$ is the common face of $\tau$ and $\tau'$ on level $j$. If $\tau'$ is obtained by regular refinement of its father $\tau''$, then $f$ is still the common face of $\tau$ and $\tau''$. But, by (6) both $\tau$, $\tau'' \in \mathcal{T}_{L(\tau'')}$. Then, this would violate conformity A.4.2 for level $L(\tau'')$. Hence, $\tau'$ must have been irregular.

On the other hand, $L(\tau) \leq L(\tau') - 1 = L(\tau'')$. Next, we proceed by eliminating the possibility that $L(\tau) < L(\tau'')$. If so, we repeat the above reasoning, and $\tau''$ becomes irregular. $\tau''$ is already the father of the irregular $\tau'$, contradicting A.4.4 for level $L(\tau'')$. Hence $L(\tau) = L(\tau'') = L(\tau') - 1$ concludes the proof. ☐

By A.4.4 and A.4.5, every tetrahedron at any $\mathcal{T}_j$ is geometrically similar to some tetrahedron in $\mathcal{T}_0$ or to a tetrahedron arising from an irregular refinement of some tetrahedron in $\mathcal{T}_0$. Then, there exist absolute constants $c_1,\ c_2$ such that

$$c_1 \ \mathrm{size}(\bar{\tau})\ 2^{-L(\tau)} \leq \mathrm{size}(\tau) \leq c_2 \ \mathrm{size}(\bar{\tau})\ 2^{-L(\tau)}, \tag{4.2}$$

where $\bar{\tau}$ is the father of $\tau$ in the initial mesh.

LEMMA 4.2. *Let $\tau$ and $\tau'$ be two tetrahedra in $\mathcal{T}_j$ sharing a common edge (two vertices). Then there exists a finite number $E$ depending on the shape regularity such that*

$$|L(\tau) - L(\tau')| \leq E. \tag{4.3}$$

*Proof.* Start with $\tau = \tau_1$, obtain the face-adjacent neighbor $\tau_2$ (either of the two faces), and then obtain the face-adjacent tetrahedron $\tau_3$ of $\tau_2$, repeat this process $\delta$ times until you reach $\tau' = \tau_\delta$. After you pick one of the two faces, the direction of the face-adjacent tetrahedra is determined. $\delta$ is always a finite number due to shape regularity. Lemma follows by face-adjacent neighbor relation (4.1). ∎

LEMMA 4.3. *Let $\tau$ and $\tau'$ be two tetrahedra in $\mathcal{T}_j$ sharing a common vertex. Then there exists a finite number $V$ depending on the shape regularity such that*

$$|L(\tau) - L(\tau')| \leq V. \tag{4.4}$$

*Proof.* Take one edge from $\tau$ and $\tau'$ where the edges meet at the common vertex. By shape regularity, there exist $\eta$ (a bounded number) many edge-adjacent tetrahedra between $\tau_1 = \tau$ and $\tau_\eta = \tau'$. By the above construction in Lemma 4.2, there exist $\delta_{1,2}$ many face-adjacent tetrahedra between $\tau_1$ and $\tau_2$. We repeat this process until we place $\delta_{\eta-1,\eta}$ many tetrahedra between $\tau_{\eta-1}$ and $\tau_\eta$. Hence, there exist $\sum_{i=1}^{\eta-1} \delta_{i,i+1}$ face-adjacent tetrahedra between $\tau$ and $\tau'$. Again, face-adjacent neighbor relation (4.1) concludes the lemma with $V = \sum_{i=1}^{\eta-1} \delta_{i,i+1}$. ∎

Consequently, simplices in the support of a basis function are comparable in size as indicated in (4.5). This is usually called *patchwise quasiuniformity*. Furthermore, it was shown in [1] that patchwise quasiuniformity (4.5) holds for 3D marked tetrahedron bisection by Joe and Liu [19] and for 2D newest vertex bisection by Sewell [26] and Mitchell [20]. Due to restrictive nature of the proof technique (see (8.7) and (8.8)), we focus on refinement procedures which obey A.4.4 and A.4.5.

LEMMA 4.4. *There is a constant depending on the shape regularity of $\mathcal{T}_j$ and the quasiuniformity of $\mathcal{T}_0$, such that*

$$\frac{\mathrm{size}(\tau)}{\mathrm{size}(\tau')} \leq c, \quad \forall \tau, \tau' \in \mathcal{T}_j, \quad \tau \cap \tau' \neq \emptyset. \tag{4.5}$$

*Proof.* $\tau$ and $\tau'$ are either face-adjacent ($d$ vertices), edge-adjacent ($d-1$ vertices), or vertex-adjacent, and are handled by (4.1), (4.3), (4.4), respectively.

$$\begin{aligned}\frac{\mathrm{size}(\tau)}{\mathrm{size}(\tau')} &\leq c\ 2^{|L(\tau)-L(\tau')|}\ \frac{\mathrm{size}(\bar{\tau})}{\mathrm{size}(\bar{\tau'})} \quad \text{(by (4.2))} \\ &\leq c\ 2^{\max\{1,E,V\}}\ \gamma^{(0)} \quad \text{(by (4.1), (4.3), (4.4) and quasiuniformity of } \mathcal{T}_0)\end{aligned}$$

∎

**5. Red refinement.** In red-green refinement, hanging nodes are closed by green refinement. In this sense, green refinement is auxiliary or secondary to red refinement. Red refinement has a number of attractive features when compared to red-green refinement. The type of red refinement under consideration here is exactly red-green refinement without green closure. This version of red refinement can be interpreted as a reduction of red-green refinement by deleting the impact of green closure. Hence, in essence a red mesh can be turned into a red-green mesh when needed. Observe that the red mesh in the middle is turned into a red-green mesh by green closure as in the right mesh in Figure 5.1. This correspondence is especially useful in inheriting certain properties in red refined meshes which are already established for red-green refinement.

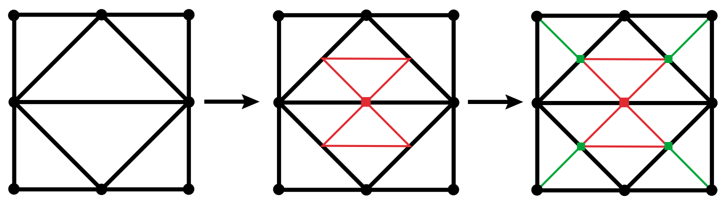

FIG. 5.1. *Left: Coarse DOF,* $N_0 = 8$. *Middle: a DOF created by red refinement,* $N_1^{red} = 9$. *Right: Green closure deployed,* $N_1^{red-green} = 13$.

Red refinement as a stand-alone procedure creates new degrees of freedom by pairwise quadrasection or octasection. The resulting hanging nodes are not closed, and therefore cannot be degrees of freedom (see the middle mesh in Figure 5.1 where a new degrees of freedom is represented by a small square). The initial triangulation $\mathcal{T}_0$ gives rise to nested, but possibly nonconforming triangulations; see the middle mesh in Figure 5.1. The father-son relationship (4.2) reduces to the simple expression:

$$\mathrm{size}(\tau) = \mathrm{size}(\bar{\tau})\; 2^{-L(\tau)}.$$

Subspaces generated by red refinement spaces are sufficiently rich to handle any desired type of singularity, but will not be as rich as the corresponding subspaces generated by red-green refinement. A function $u \in \mathcal{S}_j$ is determined by its values at degrees of freedom. Hanging nodes are always midpoints of edges connecting two degrees of freedom. The values at hanging nodes are computed by linear interpolation using the corresponding degrees of freedom at the ends of edges. Although the mesh is nonconforming, we have well-defined basis functions which satisfy the Lagrange property; see Figure 5.2.

A simplex in the red mesh can be expressed as a union of simplices in the corresponding red-green mesh. Hence, the red finite element space is a subspace of the corresponding red-green finite element space. Similarly, any simplex in $\mathcal{T}_j$ can be expressed as a union of simplices in the uniformly refined triangulations $\tilde{\mathcal{T}}_j$.

$$\tau \in \mathcal{T}_j, \quad \tau = \bigcup_{\tilde{\tau} \in \tilde{\mathcal{T}}_j} \tilde{\tau}. \tag{5.1}$$

Moreover, (5.1) implies that the basis functions which live on simplices of $\mathcal{T}_j$ can be uniquely expressed by basis functions living on simplices of $\tilde{\mathcal{T}}_j$. This gives rise to the most attractive property of red refinement: $\mathcal{S}_j$ is a true subspace of $\tilde{\mathcal{S}}_j$. The property (5.1) is no longer valid if red refinement is supplemented with the green refinement.

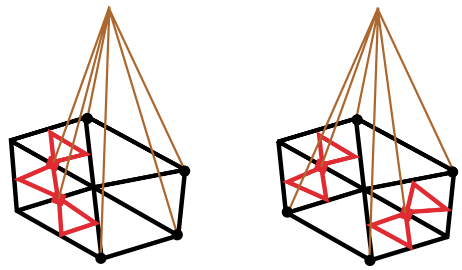

FIG. 5.2. *Basis functions on meshes created by two different red refinements. Left: Two DOF created on edge-adjacent simplices. Right: Two DOF created on non-edge-adjacent simplices.*

The fact that $\mathcal{S}_j \subset \tilde{\mathcal{S}}_j$ is quite convenient simply because the standard estimates such as inverse inequalities and Cauchy-Schwarz like estimates which naturally hold for $\tilde{\mathcal{S}}_j$ are inherited by $\mathcal{S}_j$ without additional effort. Red refinement is also preferred over red-green refinement in practical applications such as computer graphics. Two of the favorable practical properties of red refinement are as follows. Valence (number of edges meeting at a vertex) of all nodes except the coarsest ones is equal to 6, and the shape regularity constant (ratio of the diameter of circumscribing over the inscribed sphere) of the father simplex remains the same for any children in the ML hierarchy. Moreover, red refinement can be programmed using relatively simple datastructures such as quadtrees and octtrees.

Let us elaborate on the two different 2D red refinement procedures. The one in [7] enforces the difference of levels of two simplices to be *at most 1* if they have at least one common node. This brings a *patchwise uniform* refinement flavor and is closer to uniform refinement than the type of red refinement here. There is an advantage of this type of refinement: All the subsimplices of a subdivided simplex can be marked for further refinement. In our refinement, this holds only for a subset of the subsimplices. On the other hand, one can introduce DOF inside a given patch without uniformly refining the whole patch. This flexible behavior is exhibited in Figure 5.2. Our 2D red refinement guarantees the BPX optimality for $p \in L_\infty(\Omega)$. In addition, this framework supports an easy extension to any spatial dimension $d \geq 1$. The BPX optimality presented in [7] is restricted to $p \in C^1(\Omega)$ with $d = 2$.

One has to modify the red-green refinement assumptions for red refinement. Nestedness A.4.1 is automatically satisfied. However, conformity in the sense of A.4.2 is violated. Neighbor relations in A.4.2 are categorized as *regular adjacency*, and either of the alternative relations, namely, intersection at a quarter of a face, a half of an edge, and a hanging node is categorized as *irregular adjacency*. The latter category will be characterized as $f/4$, $e/2$, and $h$ respectively. In order to preserve nondegeneracy A.4.3, we substitute A.4.4 by the following analogous version designed for red refinement.

ASSUMPTION 5.1. *There can be at most one hanging node $h$ on a given edge in $\mathcal{T}_j$.*

A.4.5 and A.4.6 can remain as stated. A side remark about A.4.6 is that A.4.4 dictates a quadrasection rather than two bisections for a 2D simplex having two hanging nodes, and an octasection rather than irregular refinement for a 3D simplex having three or more hangings nodes that do not belong to a common face.

**6. Generation bounds.** Assume that $\tau$ is a child simplex formed by regular refinement, which now has an irregular adjacency with $\tau'$. Then green closure in the case of red-green refinement gives rise to a child $\tau''$ of $\tau'$ with $L(\tau'') = L(\tau') + 1$. By A.4.5, $\tau$ and $\tau''$ live on the same level. One returns to red refinement by undoing the green closure, and finds that $L(\tau) = L(\tau') + 1$. Consequently, we have the following generation bound for irregular adjacency:

$$|L(\tau) - L(\tau')| = 1, \quad \tau \cap \tau' = \{f/4, e/2, h\}. \tag{6.1}$$

For regular adjacency, the generation bound (4.1) is still valid for face adjacent $\tau$ and $\tau'$, neither of which has an irregular adjacency with the rest of the simplices. However, if either one, say $\tau$, has an irregular adjacency with any other simplex, then after green closure, a child $\tau''$ of $\tau'$ will have the generation bound (4.1) with $\tau$. Undoing the green closure can give rise to a case which can increase the generation bound by 1.

$$|L(\tau) - L(\tau')| \leq 2, \quad \tau \cap \tau' = f. \tag{6.2}$$

Since we have generation bounds (6.1) and (6.2) for face adjacent simplices for irregular and regular adjacencies respectively, a generation bound such as (4.3) for edge adjacent follows by shape regularity with a larger bound than the red-green bound due to irregular adjacencies. Similarly, an analog of (4.4) follows by an application of the bound for edge adjacent simplices, the bound for irregular adjacency for the case of hanging nodes, and by shape regularity.

Basis stability (7.10), and hence the Bernstein estimate (2.1) holds, and the ultimate goal, namely, the optimal BPX lower bound for red refinement, follows by exactly the same analysis as done for the case of red-green refinement.

**7. Establishing optimality of the BPX preconditioner.** In this section, we extend the Dahmen- Kunoth framework of to three spatial dimensions and the extension closely follows the original work. However, the general case for $d \geq 1$ spatial dimensions is never presented in the literature, and we present it for completeness.

Our motivation is to form a stable basis in the following sense [24].

$$\| \sum_{x_i \in \mathcal{N}_j} u_i \phi_i^{(j)} \|_{L_2(\Omega)} \simeq \| \{ \text{size}^{1/2}(\text{supp } \phi_i^{(j)}) \; u_i \}_{x_i \in \mathcal{N}_j} \|_{l_2}. \tag{7.1}$$

The basis stability (7.1) will then guarantee that the Bernstein estimate (2.1) holds, which is the first step in establishing the norm equivalence (2.4). For a stable basis, functions with small supports have to be augmented by an appropriate scaling so that $\|\phi_i^{(j)}\|_{L_2(\Omega)}$ remains roughly the same for all basis functions. This is reflected in size(supp $\phi_i^{(j)}$) by defining:

$$L_{j,i} = \min\{L(\tau) : \tau \in \mathcal{T}_j, \; x_i \in \tau\}. \tag{7.2}$$

Then

$$\text{size}(\text{supp } \phi_i^{(j)}) \simeq 2^{-dL_{j,i}}.$$

We prefer to use an equivalent stability and such basis is called $L_2$*-stable Riesz basis* (see section 2 and Remark 1 in [3]).

$$\| \sum_{x_i \in \mathcal{N}_j} \hat{u}_i \hat{\phi}_i^{(j)} \|_{L_2(\Omega)} \simeq \| \{ \hat{u}_i \}_{x_i \in \mathcal{N}_j} \|_{l_2}, \tag{7.3}$$

where $\hat{\phi}_i^{(j)}$ denotes the scaled basis, and the relationship between (7.1) and (7.3) is given as follows:

$$\hat{\phi}_i^{(j)} = 2^{d/2L_{j,i}}\ \phi_i^{(j)}, \quad \hat{u}_i = 2^{-d/2L_{j,i}}\ u_i, \quad x_i \in \mathcal{N}_j. \tag{7.4}$$

REMARK 7.1. *Our construction works for any d-dimensional setting with the scaling (7.4). However, it is not clear how to define face-adjacency relations for* $d > 3$. *If such relations can be defined through some topological or geometrical abstraction, then our framework naturally extends to d-dimensional local refinement strategies, and hence the optimality of the BPX preconditioner can be guaranteed in* $\mathbb{R}^d,\ d \geq 1$.

*The analysis is done purely with basis functions, completely independent of the underlying mesh geometry. Since basis functions in red refinement are well-defined (see Figure 5.2), hanging nodes do not pose any problems.*

The element mass matrix gives rise to the following useful formula.

$$\|g\|^2_{L_2(\tau)} = \frac{\text{volume}(\tau)}{(d+1)(d+2)}\ (\sum_{i=1}^{d+1} g(x_i)^2 + [\sum_{i=1}^{d+1} g(x_i)]^2), \tag{7.5}$$

where, $i = 1, \ldots, d+1$ and $x_i$ is a vertex of $\tau$, $d = 2, 3$. In view of (7.5), we have that

$$\|\hat{\phi}_i^{(j)}\|^2_{L_2(\Omega)} = 2^{dL_{j,i}}\ \frac{\text{volume}(\text{supp}\ \hat{\phi}_i^{(j)})}{(d+1)(d+2)}.$$

Since the min in (7.2) is attained, there exists at least one $\tau \in \text{supp}\ \hat{\phi}_i^{(j)}$ such that $L(\tau) = L_{j,i}$. By (4.2) we have

$$2^{L_{j,i}} \simeq \frac{\text{size}(\tau)}{\text{size}(\bar{\tau})}. \tag{7.6}$$

Also,

$$\text{volume}(\text{supp}\ \hat{\phi}_i^{(j)}) \simeq \sum_{i=1}^{M} \text{size}^3(\tau_i), \quad \tau_i \in \text{supp}\ \hat{\phi}_i^{(j)}. \tag{7.7}$$

By (4.5), we have

$$\text{size}(\tau_i) \simeq \text{size}(\tau). \tag{7.8}$$

Combining (7.7) and (7.8), we conclude

$$\text{volume}(\text{supp}\ \hat{\phi}_i^{(j)}) \simeq M\text{size}^3(\tau). \tag{7.9}$$

Finally then, (7.6) and (7.9) yield

$$2^{dL_{j,i}}\text{volume}(\text{supp}\ \hat{\phi}_i^{(j)}) \simeq M\ \frac{1}{\text{size}^3(\bar{\tau})}.$$

$M$ is a uniformly bounded constant by shape regularity. One can view the size of any tetrahedron in $\mathcal{T}_0$, in particular size of $\bar{\tau}$, as a constant. The reason is the following: A.4.4 and A.4.5 force every tetrahedron at any $\mathcal{T}_j$ to be geometrically similar to some tetrahedron in $\mathcal{T}_0$ or to a tetrahedron arising from an irregular refinement of some tetrahedron in $\mathcal{T}_0$, hence, to some tetrahedron of a fixed finite collection. Combining the two arguments above, we have established that

$$\|\hat{\phi}_i^{(j)}\|_{L_2(\Omega)} \simeq 1, \quad x_i \in \mathcal{N}_j. \tag{7.10}$$

Let $g = \sum_{x_i \in \mathcal{N}_j} \hat{u}_i \hat{\phi}_i^{(j)} \in \mathcal{S}_j$. For any $\tau \in \mathcal{T}_j$ we have that

$$\|g\|^2_{L_2(\tau)} \le c \sum_{x_i \in \mathcal{N}_{j,\tau}} |\hat{u}_i|^2 \|\hat{\phi}_i^{(j)}\|^2_{L_2(\Omega)}, \tag{7.11}$$

where $\mathcal{N}_{j,\tau} = \{x_i \in \mathcal{N}_j : x_i \in \tau\}$, which is uniformly bounded in $\tau \in \mathcal{T}_j$ and $j \in \mathbb{N}_0$. By the scaling (7.4), we get equality in the estimate below. The inequality is a standard inverse estimate where one bounds $g(x_i)$ using formula (7.5) and by handling the volume in the formula by (4.2):

$$|\hat{u}_i|^2 = 2^{-dL_{j,i}} |g(x_i)|^2 \le c\; 2^{-dL_{j,i}} 2^{dL_{j,i}} \|g\|^2_{L_2(\tau)}. \tag{7.12}$$

We sum up over $\tau \in \mathcal{T}_j$ in (7.11) and (7.12), and by using (7.10) we achieve stability (7.3). This allows us to establish the Bernstein estimate (2.1).

LEMMA 7.1. *For the scaled basis (7.4), the Bernstein estimate (2.1) holds for* $\beta = 3/2$

*Proof.* (7.10) with (7.11) and (7.12) assert that the scaled basis (7.4) is stable in the sense of (7.3). Hence, (2.1) holds by Theorem 4 in [24]. Note that the proof actually works independently of the spatial dimension. ∎

**8. Lower bound in the norm equivalence.** In a locally refined mesh, the Jackson estimate (2.3) holds only for functions whose singularities are somehow well-captured by the mesh geometry. For instance, if a mesh is designed to pick up the singularity at $x = 0$ of $y = 1/x$, then on the same mesh we will not be able to recover a singularity at $x = 1$ of $y = 1/(x-1)$. Hence the Jackson estimate (2.3) cannot hold for all functions $f \in W_p^k$. As pointed out above, some kind of a Jackson-like estimate (2.2) can hold for special functions. In order to get the lower bound, we focus on estimating $v_J^{(2)}$ directly, as in [14] for the 2D setting.

Let $\tau \in \mathcal{T}_j$ be a tetrahedron with vertices $x_1, x_2, x_3, x_4$. Clearly the restrictions of $\hat{\phi}_i^{(j)}$ to $\tau$ are linearly independent over $\tau$ where $x_i \in \{x_1, x_2, x_3, x_4\}$. Then, there exists a unique set of linear polynomials $\psi_1^\tau, \psi_2^\tau, \psi_3^\tau, \psi_4^\tau$ such that

$$\int_\tau \hat{\phi}_k^{(j)}(x,y,z)\psi_l^\tau(x,y,z)dxdydz = \delta_{kl}, \quad x_k, x_l \in \{x_1, x_2, x_3, x_4\}. \tag{8.1}$$

For $x_i \in \mathcal{N}_j$ and $\tau \in \mathcal{T}_j$, define a function for $x_i \in \tau$

$$\xi_i^{(j)}(x,y,z) = \begin{cases} \frac{1}{M_i}\psi_i^\tau(x,y,z), & (x,y,z) \in \tau \\ 0, & (x,y,z) \notin \text{supp } \hat{\phi}_i^{(j)} \end{cases}, \tag{8.2}$$

where $M_i$ is the number of tetrahedra in $\mathcal{T}_j$ in supp $\hat{\phi}_i^{(j)}$. By (8.1) and (8.2), we obtain

$$(\xi_k^{(j)}, \hat{\phi}_l^{(j)}) = \int_\Omega \xi_k^{(j)}(x,y,z)\hat{\phi}_l(x,y,z)dxdydz = \delta_{kl}, \quad x_k, x_l \in \mathcal{N}_j. \tag{8.3}$$

We can now define a quasi-interpolant, in fact a *projector* onto $\mathcal{S}_j$, such that

$$(\tilde{Q}_j f)(x,y,z) = \sum_{x_i \in \mathcal{N}_j} (f, \xi_i^{(j)}) \hat{\phi}_i^{(j)}(x,y,z). \tag{8.4}$$

One can easily observe by (7.10) and (8.3) that

$$\|\xi_i^{(j)}\|_{L_2(\Omega)} \simeq 1, \quad x_i \in \mathcal{N}_j, \; j \in \mathbb{N}_0. \tag{8.5}$$

Letting $\Omega_{j,\tau} = \bigcup\{\tau' \in \mathcal{T}_j : \; \tau \cap \tau' \neq \emptyset\}$, we can conclude from (7.10) and (8.5) that

$$\|\tilde{Q}_j f\|_{L_2(\tau)} = \| \sum_{x_k \in \mathcal{N}_{j,l}} (f, \xi_l^{(j)}) \hat{\phi}_k^{(j)} \|_{L_2(\tau)} \leq c \|f\|_{L_2(\Omega_{j,\tau})}. \tag{8.6}$$

We define now a subset of the triangulation where the refinement activity stops, meaning that all tetrahedra in $\mathcal{T}_j^*, \; j \leq m$ also belong to $\mathcal{T}_m$:

$$\mathcal{T}_j^* = \{\tau \in \mathcal{T}_j : \; L(\tau) < j, \; \Omega_{j,\tau} \cap \tau' = \emptyset, \; \forall \tau' \in \mathcal{T}_j \text{ with } L(\tau') = j\}. \tag{8.7}$$

Since there is no refinement activity on $\mathcal{T}_j^*$, the finite element spaces are frozen in the refinement hierarchy. That is, $g \in \mathcal{S}_m$ implies $g \in \mathcal{S}_j$ for $j \leq m$. Note that $\tilde{Q}_j g = g$ for $g \in \mathcal{S}_j$. Hence,

$$\|g - \tilde{Q}_j g\|_{L_2(\tau)} = 0, \quad \tau \in \mathcal{T}_j^*. \tag{8.8}$$

Using the fact that $\tilde{Q}_j$ fixes polynomials of degree at most 1 (i.e. $\Pi_1(\mathbb{R}^3)$) and (8.6), we arrive:

$$\begin{aligned} \|g - \tilde{Q}_j g\|_{L_2(\tau)} &\leq \|g - P\|_{L_2(\tau)} + \|\tilde{Q}_j (P - g)\|_{L_2(\tau)} \\ &\leq c \; \|g - P\|_{L_2(\Omega_{j,\tau})}, \quad \tau \in \mathcal{T}_j \setminus \mathcal{T}_j^*. \end{aligned} \tag{8.9}$$

We would like to bound the right hand side of (8.9) in terms of a modulus of smoothness in order to reach a Jackson-type estimate. Following [14], we utilize a modified modulus of smoothness:

$$\tilde{\omega}_k(f,t,\Omega)_p^p = t^{-s} \int_{[-t,t]^s} \|\Delta_h^k f\|_{L_p(\Omega_{k,h})}^p \; dh.$$

They can be shown to be equivalent [15]: there exist $0 < c_1, c_2 < \infty$ such that

$$c_1 \tilde{\omega}_k(f,t,\Omega)_p \leq \omega_k(f,t,\Omega)_p \leq c_2 \tilde{\omega}_k(f,t,\Omega)_p.$$

(The equivalence of the two moduli of smoothness $\omega_k$ and $\tilde{\omega}_k$ in the one-dimensional setting can be found in Lemma 5.1 in [15].)

For $\tau$ a simplex in $\mathbb{R}^d$ and $t = \text{size}(\tau)$, a Whitney estimate shows that [16, 22, 28]

$$\inf_{P \in \Pi_{k-1}(\mathbb{R}^d)} \|f - P\|_{L_p(\tau)} \leq c \tilde{\omega}_k(f,t,\tau)_p, \tag{8.10}$$

where $c$ depends only on the smallest angle of $\tau$ but not on $f$ and $t$. The reason why $\tilde{Q}_j$ works well for tetrahedralization in 3D is the fact that the Whitney estimate (8.10) remains valid for any spatial dimension. $\mathcal{T}_j \setminus \mathcal{T}_j^*$ is the part of the tetrahedralization $\mathcal{T}_j$ where refinement is active at every level. Then, in view of (4.5)

$$\text{size}(\Omega_{j,\tau}) \simeq 2^{-j}, \quad \tau \in \mathcal{T}_j \setminus \mathcal{T}_j^*.$$

Taking the inf over $P \in \Pi_{k-1}(\mathbb{R}^3)$ in (8.9) and using the Whitney estimate (8.10) we conclude

$$\|g - \tilde{Q}_j g\|_{L_2(\tau)} \leq c\tilde{\omega}_k(g, 2^{-j}, \Omega_{j,\tau})_2.$$

Recalling (8.8) and summing over $\tau \in \mathcal{T}_j \setminus \mathcal{T}_j^*$ gives rise to

$$\|g - \tilde{Q}_j g\|_{L_2(\Omega)} \leq c\tilde{\omega}_2(g, 2^{-j}, \Omega)_2.$$

It is time to switch from the modified modulus of smoothness to the standard one. Consequently,

$$\|g - Q_j g\|_{L_2(\Omega)} \leq \|g - \tilde{Q}_j g\|_{L_2(\Omega)} \leq c\ \omega_2(g, 2^{-j}, \Omega)_2,$$

and (3.1) yields

$$v_J^{(2)} = O(1), \quad J \to \infty.$$

**9. Additional results.** We present a simple optimality proof of the BPX preconditioner using $K$-functionals, based on a critical insight pointed out by Bornemann in [8].

DEFINITION 9.1.

$$K(f, t; L_p(\Omega), W_p^k(\Omega)) = \inf_{g \in W_p^k(\Omega)} \left\{ \|f - g\|_{L_p(\Omega)} + t|g|_{W_p^k(\Omega)} \right\}, \tag{9.1}$$

*is called the $K$-functional for $f \in L_p(\Omega)$.* Both modulus of smoothness and $K$-functionals give direct information on the smoothness of $f$. We will use the extended version of the Jackson estimate (2.3) where $\alpha \in \mathbb{R}$:

DEFINITION 9.2. *Let $\mathcal{S}_J \subset Y \subset L_p(\Omega)$. We say that the approximation inequality holds, if there exists $\alpha > 0$ satisfying*

$$E_{\mathcal{S}_J}(u)_p \leq c\ (2^{-J})^\alpha |u|_Y, \quad u \in Y. \tag{9.2}$$

It is well known that the $L_2$-projection is $H^s$-stable on quasiuniform meshes [12]:

$$|Q_j u|_{H^s} \leq c\ |u|_{H^s}, \quad \forall u \in H^s,\ 0 \leq s \leq 1. \tag{9.3}$$

In addition, the approximation inequality (9.2) holds in the following form:

$$E_{\mathcal{S}_j}(f)_2 = \|f - Q_j f\|_{L_2} \leq c\ (2^{-j})^s |f|_{H^s}, \quad f \in H^s, \quad 0 \leq s \leq 2.$$

Using (9.2) and the $H^s$-stability of $L_2$-projection:

$$\|Q_j u - Q_{j-1} u\|_{L_2} = \|Q_j u - Q_{j-1} Q_j u\|_{L_2} \leq c\ (2^{-j})^s |Q_j u|_{H^s} \leq c\ (2^{-j})^s |u|_{H^s}.$$

The canonical suboptimal lower bound for the BPX preconditioner on quasiuniform meshes can be obtained:

$$\|u\|_{\text{BPX}}^2 \leq c \sum_{j=0}^{J} |u|_{H^s}^2 = c\ J\ |u|_{H^s}^2, \quad s = 1.$$

Next, we present a novel approach by using $K$-functionals to obtain the optimal and suboptimal lower bound for quasiuniform meshes. We begin with a lemma that characterizes the components of the BPX norm in terms of the $K$-functional.

LEMMA 9.3. *Let the approximation inequality (9.2) hold, then*

$$\|(Q_j - Q_{j-1})u\|_{L_2} \le c\ K(u, 2^{-(j-1)s}; L_2, H^s). \tag{9.4}$$

*Proof.* Let $v \in H^s$, then

$$\begin{aligned}
\|(Q_j - Q_{j-1})u\|_{L_2} &\le \|u - Q_{j-1}u\|_{L_2} \\
&\le \|u - v\|_{L_2} + \|v - Q_{j-1}v\|_{L_2} + \|Q_{j-1}(u-v)\|_{L_2} \\
&\le 2\|u - v\|_{L_2} + \|v - Q_{j-1}v\|_{L_2} \\
&\le c\ (\|u - v\|_{L_2} + (2^{-(j-1)})^s |v|_{H^s})\ \text{(by (9.2))} \\
&\le c\ K(u, 2^{-(j-1)s}; L_2, H^s).
\end{aligned}$$

See Lemma 7.1 in [7] for an alternative proof. ∎

Lemma 9.3 requires the approximation inequality which cannot be satisfied on locally refined meshes for the reason stated in the beginning of §8. That is why our optimality results will, for the moment, be limited to uniform refinement. If one can establish the $H^s$-stability, $s > 1$, of the $L_2$-projection, then we will show that the optimal lower bound can be obtained. This approach will be taken in the next theorem and it dramatically simplifies the optimality proof as is first suggested in [1]. However, to the authors' knowledge, the $H^s$-stability, $s > 1$, of the $L_2$-projection has not appeared in the literature. Although the $H^s$-stability for $s > 1$ has not been carefully studied, there is consensus [25, 31] that it can be established.

THEOREM 9.4. *If the $L_2$-projection $Q_{j-1}$ is $H^s$-stable as in (9.3), then for*

- $s = 1 + \epsilon$, *we have* $\|u\|^2_{\text{BPX}} \le c|u|^2_{H^s}$,
- $s = 1$, *we have* $\frac{1}{J}\|u\|^2_{\text{BPX}} \le c|u|^2_{H^s}$.

*Proof.* $Q_{j-1} : L_2 \mapsto \mathcal{S}_{j-1}$, also we know that $\mathcal{S}_{j-1} \subset H^s$.

$$\begin{aligned}
K(u, 2^{-(j-1)s}; L_2, H^s) &\le \|u - Q_{j-1}u\|_{L_2} + (2^{-(j-1)})^s |Q_{j-1}u|_{H^s} \\
&\le c\ 2^{-(j-1)s} |u|_{H^s}.
\end{aligned}$$

The last step is due to the approximation inequality and the $H^s$-stability of the $L_2$-projection.

$$\begin{aligned}
\sum_{j=0}^{J} 2^{2j} \|(Q_j - Q_{j-1})u\|^2_{L_2} &\le c\ \sum_{j=0}^{J} 2^{2j} K(u, 2^{-(j-1)s}; L_2, H^s)^2\ \text{(by (9.4))} \\
&\le c\ 2^{2s}\ \sum_{j=0}^{J} 2^{2j(1-s)}\ |u|^2_{H^s} \\
&\le c\ |u|^2_{H^s} \quad (s = 1 + \epsilon > 1) \\
\text{or} &\le c\ \frac{1}{J} |u|^2_{H^s} \quad (s = 1).
\end{aligned}$$

∎

The remaining piece for BPX optimality is the upper bound, and it is established by using the Bernstein estimate (2.1) in the same way as in (2.4).

Remark 9.1. *It is interesting to note that the original BPX proof was improved to optimal bounds by realizing that the Jackson estimate holds for $H^s$, $s > 1$, an insight due perhaps to Bornemann [8]. The relationship between $H^s$-stability of $L_2$-projection and the Jackson estimate is currently unclear.*

**Acknowledgments.** The authors thank R. Bank, P. Vassilevski, and J. Xu for many enlightening discussions.

## REFERENCES

[1] B. Aksoylu, *Adaptive Multilevel Numerical Methods with Applications in Diffusive Biomolecular Reactions*, PhD thesis, Department of Mathematics, University of California, San Diego, La Jolla, CA, 2001.
[2] B. Aksoylu, S. Bond, and M. Holst, *An odyssey into local refinement and multilevel preconditioning III: Implementation and numerical experiments*, in Proceedings of the 7th Copper Mountain Conference on Iterative Methods, H. van der Vorst, ed., Copper Mountain, CO, 2002, SIAM J. Sci. Comput. Copper Mountain special issue, in review.
[3] B. Aksoylu and M. Holst, *An odyssey into local refinement and multilevel preconditioning II: Stabilizing hierarchical basis methods*, SIAM J. Numer. Anal., (2002). in review.
[4] R. E. Bank, *Hierarchical basis and the finite element method*, Acta Numerica, (1996), pp. 1–43.
[5] R. E. Bank, T. Dupont, and H. Yserentant, *The hierarchical basis multigrid method*, Numer. Math., 52 (1988), pp. 427–458.
[6] J. Bey, *Tetrahedral grid refinement*, Computing, 55 (1995), pp. 271–288.
[7] F. Bornemann and H. Yserentant, *A basic norm equivalence for the theory of multilevel methods*, Numer. Math., 64 (1993), pp. 455–476.
[8] F. A. Bornemann, *Interpolation Spaces and Optimal Multilevel Preconditioner*, in Proceedings of the 7th International Conference on Domain Decomposition Methods, D. Keyes and J. Xu, eds., Providence, 1994, CONM Series of the AMS, MR 95j:65155; Zbl. 818.65107, pp. 3–8.
[9] J. H. Bramble and J. E. Pasciak, *New estimates for multilevel algorithms including the V-cycle*, Math. Comp., 60 (1993), pp. 447–471.
[10] J. H. Bramble, J. E. Pasciak, and O. Steinbach, *On the stability of the $L^2$ projection in $H^1(\Omega)$*, Math. Comp., (2000). accepted.
[11] J. H. Bramble, J. E. Pasciak, and J. Xu, *Parallel multilevel preconditioners*, Math. Comp., 55 (1990), pp. 1–22.
[12] J. H. Bramble and J. Xu, *Some estimates for a weighted $L^2$ projection*, Math. Comp., 56 (1991), pp. 463–476.
[13] C. Carstensen, *Merging the Bramble-Pasciak-Steinbach and the Crouzeix-Thomee criterion for H1-stability of the L2-projection onto finite element spaces*, Math. Comp., (2000). Accepted. Berichtsreihe des Mathematischen Seminars Kiel 00-1 (2000).
[14] W. Dahmen and A. Kunoth, *Multilevel preconditioning*, Numer. Math., 63 (1992), pp. 315–344.
[15] R. A. DeVore and G. G. Lorentz, *Constructive Approximation*, Grundlehren der mathematischen Wissenschaften 303, Springer Verlag, Berlin Heidelberg, 1993.
[16] R. A. DeVore and V. A. Popov, *Interpolation of Besov spaces*, Trans. Amer. Math. Soc., 305 (1988), pp. 397–414.
[17] M. Holst, *Adaptive numerical treatment of elliptic systems on manifolds*, Advances in Computational Mathematics, 15 (2001), pp. 139–191.
[18] S. Jaffard, *Wavelet methods for fast resolution of elliptic problems*, SIAM J. Numer. Anal., 29 (1992), pp. 965–986.
[19] B. Joe and A. Liu, *Quality local refinement of tetrahedral meshes based on bisection*, SIAM J. Sci. Comput., 16 (1995), pp. 1269–1291.
[20] W. F. Mitchell, *Unified Multilevel Adaptive Finite Element Methods for Elliptic Problems*, PhD thesis, Computer Science, University of Illinois at Urbana-Champaign, Urbana, IL, 1988.
[21] M. E. G. Ong, *Uniform refinement of a tetrehedron*, SIAM J. Sci. Comput., 15 (1994), pp. 1134–1144.
[22] P. Oswald, *On function spaces related to finite element approximation theory*, Zeitschrift für Analysis und ihre Anwendungen, 9 (1990), pp. 43–64.
[23] ———, *On discrete norm estimates related to multilevel preconditioners in the finite element*

*method*, In K. G. Ivanov, P. Petrushev, and B. Sendov, editors, Proceedings International Conference on Constructive Theory of Functions, Varna 1991, (1992), pp. 203–214. Publ. House of Bulgarian Academy of Sciences.

[24] ———, *Multilevel Finite Element Approximation Theory and Applications*, Teubner Skripten zur Numerik, B. G. Teubner, Stuttgart, 1994.

[25] J. E. Pasciak, *Private communication*.

[26] E. G. Sewell, *Automatic generation of triangulations for piecewise polynomial approximation*, PhD thesis, Department of Mathematics, Purdue University, West Lafayette, IN, 1972.

[27] R. Stevenson, *A robust hierarchical basis preconditioner on general meshes*, Numer. Math., 78 (1997), pp. 269–303.

[28] E. A. Storozhenko and P. Oswald, *Jackson's theorem in the spaces* $L_p(\mathbb{R}^k)$, $0 < p < 1$, Siberian Math., 19 (1978), pp. 630–639.

[29] P. S. Vassilevski and J. Wang, *Stabilizing the hierarchical basis by approximate wavelets, I: Theory*, Numer. Linear Alg. Appl., 4 Number 2 (1997), pp. 103–126.

[30] ———, *Wavelet-like methods in the design of efficient multilevel preconditioners for elliptic PDEs*, in Multiscale Wavelet Methods For Partial Differential Equations, W. Dahmen, A. Kurdila, and P. Oswald, eds., Academic Press, 1997, ch. 1, pp. 59–105.

[31] J. Xu, *Private communication*.

[32] ———, *Iterative methods by space decomposition and subspace correction*, SIAM Rev., 34 (1992), pp. 581–613.

[33] H. Yserentant, *On the multilevel splitting of finite element spaces*, Numer. Math., 49 (1986), pp. 379–412.

[34] ———, *Two preconditioners based on the multi-level splitting of finite element spaces*, Numer. Math., 58 (1990), pp. 163–184.

[35] ———, *Old and new convergence proofs for multigrid methods*, Acta Numerica, (1993), pp. 285–326.

[36] S. Zhang, *Multilevel iterative techniques*, PhD thesis, Pennsylvania State University, 1988.

[37] X. Zhang, *Multilevel Schwarz methods*, Numer. Math., 63 (1992), pp. 521–539.