A Fast Algorithm for Quadrature by Expansion in Three Dimensions

Matt Wala*[1] and Andreas Klöckner†[1]

[1]Department of Computer Science, University of Illinois at Urbana-Champaign

March 29, 2019

Abstract

This paper presents an accelerated quadrature scheme for the evaluation of layer potentials in three dimensions. Our scheme combines a generic, high order quadrature method for singular kernels called Quadrature by Expansion (QBX) with a modified version of the Fast Multipole Method (FMM). Our scheme extends a recently developed formulation of the FMM for QBX in two dimensions, which, in that setting, achieves mathematically rigorous error and running time bounds. In addition to generalization to three dimensions, we highlight some algorithmic and mathematical opportunities for improved performance and stability. Lastly, we give numerical evidence supporting the accuracy, performance, and scalability of the algorithm through a series of experiments involving the Laplace and Helmholtz equations.

1 Introduction

Integral equation methods are an attractive approach for the solution of boundary value problems of elliptic partial differential equations (PDEs). The mathematical features that make integral equation methods attractive in two dimensions also hold in the three dimensional case, where their impact is felt even more drastically: exterior problems pose no more complication than their interior counterparts, a dimensional reduction in the number of degrees of freedom is achieved, and the conditioning of the numerical discretization mirrors that of the physical problem. However, due to the difficulties of scale and engineering involved, the practical realization in three dimensions is more difficult than in two.

The premise of integral equation methods is based on the reformulation of the underlying differential equation in integral form. The solution to a homogeneous elliptic boundary value problem (BVP) may be represented as a *layer potential*, such as the single-layer potential $\mathcal{S}$, a surface convolution integral over the boundary Γ:

$$\mathcal{S}\mu(x) \coloneqq \int_\Gamma \mathcal{G}(x,y)\mu(y)\,dS(y) \tag{1}$$

where the density function $\mu : \Gamma \to \mathbb{R}$ is unknown, and $\mathcal{G}$ is the free-space Green's function for the (homogeneous) PDE. For instance, for the Laplace equation $\triangle u = 0$ in three dimensions,

$$\mathcal{G}(x,y) \coloneqq (4\pi)^{-1}|x-y|_2^{-1}. \tag{2}$$

A number of challenges when solving a BVP with integral equation methods are apparent, particularly in three dimensions. First, solving the BVP requires the linear operator representing the restriction of $\mathcal{S}$ (or, depending on the integral equation formulation, some other layer potential) to the boundary Γ. When the restriction of $\mathcal{S}$ to the boundary is discretized, it becomes a finite dimensional linear operator $\mathcal{L}$. Unlike the

*wala1@illinois.edu

†andreask@illinois.edu

discretization of differential operators, the matrix representation of $\mathcal{L}$ is dense. Even considering the effects of dimensional reduction (from volume to surface), the number of degrees of freedom for three dimensional problems is sufficiently large that explicit formation of a matrix for $\mathcal{L}$ can be prohibitively expensive.

A second challenge remains the same as the two dimensional case: obtaining suitable, low complexity quadrature for the singular integrals involving the Green's function. While quadrature techniques have long been studied for two dimensional kernels, they are often built to be special-purpose, and those having three dimensional analogs are comparatively fewer. A related but more subtle issue is quadrature for the nearly singular integrals of the potential that arise when the evaluation point is close to but not on the boundary.

A rough overview of the subject of singular quadrature is given in [37] to which we refer the reader. Some notable schemes featuring singular quadrature rules for one, two, and three dimensional problems include [29, 40, 21, 28, 42, 3, 12, 26, 22, 30, 51, 32, 9, 59, 6, 17, 54, 31, 52, 10, 1, 41, 11, 27, 7, 8, 60, 43]. Most schemes intended for use when the number of evaluation or source points is large, such as for the application of layer potentials like (1), feature an acceleration component. A useful tool for this has been the Fast Multipole Method [24], e.g. in [59]. A variety of other acceleration methods have been utilized, such as fast direct solvers (e.g. [7]), recursive compressed inverse preconditioning [29], particle-mesh Ewald summation (e.g. [60]), or methods based on the Fast Fourier Transform (e.g. [42]).

Quadrature by Expansion (QBX, [37]) is a quadrature method that has been recently developed that promises to unify the treatment of the two- and three-dimensional, on-surface and off-surface cases for layer potentials like (1), including those with hypersingular kernels. QBX is applicable in this generality since it only relies on the analyticity of the underlying potential as a function inside or outside the domain, and, under mild assumptions, the existence of a smooth extension of the potential onto the boundary of the domain. Additionally, with some care, QBX is amenable to acceleration with the Fast Multipole Method, as we demonstrate in this contribution.

Because the FMM forms local expansions of source potentials, it appears well suited to a scheme known as 'global' QBX, in which a local expansion mediates the potential due to all sources in the geometry. The first published practical realization of a QBX-FMM coupling, in [47, 48] is a global scheme that achieves high accuracy and acceleration in two dimensions. However, in this algorithm, the error introduced by the FMM does not obey an error bound of the form $O((1/2)^{p_{\text{fmm}}+1})$, where p_{fmm} is the FMM order, which may be expected of a 'point' FMM in two dimensions [46]. In order to achieve a given accuracy tolerance $\varepsilon > 0$, one requires an often considerably greater FMM order than would be applicable for a 'point' FMM, and, additionally, the amount by which to increase the order must be empirically determined. A more recent version [57] of the QBX FMM redevelops the FMM algorithm with a guaranteed error bounds resembling that of the 'point' FMM. Both of these algorithms are based on an 'analytical' FMM. In contrast, the contribution in [49] develops Quadrature by Kernel-Independent Expansion, a 'numerical' version of QBX meant for use with the kernel-independent FMM.

Other recent research on QBX [34] has focused on automating the selection of parameters for quadrature, radius, and expansion order. Although thus far most of the work on QBX has been restricted to two dimensions, theoretical work anticipates the extensibility of QBX to the three dimensional setting [16, 35]. An exception is [33] which uses QBX in three dimensions over spheroidal bodies.

An alternative to global QBX recently demonstrated to be viable in three dimensions is 'local' QBX. In contrast to global QBX, in local QBX only the field in a neighborhood of each target point is mediated through local expansions [2]. This enables more geometric flexibility when placing expansion centers. This scheme is straightforward to integrate with the FMM, as only an approximation to the 'point' far field is required. Despite these advantages over global QBX, local QBX appears to require a generically higher quadrature order, and thus higher quadrature oversampling to control the additional error introduced by the transition from the QBX near-field to the point-source-based computation of the far field. The contribution [53] develops a three-dimensional local QBX algorithm with optimizations to decrease the cost of applying the QBX expansions.

This paper describes an accelerated global QBX scheme in three dimensions which builds and extends on GIGAQBX, our previous scheme for two dimensions featuring rigorous error bounds [57]. While much of the theory and many algorithmic aspects are directly analogous to the two dimensional case, we introduce a number of enhancements that help keep the scheme feasible and practical:

- We replace the geometry processing and refinement scheme of [48] with one that is applicable to surfaces in three dimensions. Specifically, we introduce new measures of quadrature resolution, and we replace

the 2-to-1 length requirement in that scheme (which does not straightforwardly generalize to three dimensions) with a two-stage refinement scheme that separates the calculation of the QBX expansion radii from the mesh resolution of the quadrature discretization. This is the subject of Sections 3 and 4.

- As part of this geometry processing, we report on an empirically effective criterion that aids in controlling the truncation error based on mesh element geometry, presented in Section 3.3.
- Algorithm 3 presents a simplified version of the target-to-center association algorithm of [48].
- We provide an improved version of the 'target confinement rule' of [57]. This rule governs the relationship between boxes as they occur in the GIGAQBX FMM and the QBX expansions used to approximate the layer potential near the source geometry and thus plays a central role in determining the cost of the algorithm. In [57], we used 'square' target confinement regions, i.e. ones whose geometry is governed by the ℓ^∞ norm. By more closely matching the true convergence behavior of the QBX expansions through the definition of the target confinement region with the help of the ℓ^2 norm, we obtain a considerable cost reduction. Amidst error estimates generalized to the three-dimensional case, this is the subject of Section 5.
- We derive complexity estimates and state conditions under which one may expect linear time complexity of the algorithm in three dimensions. This is the subject of Section 5.5.

Before we turn to the subject of these contributions, the next section presents background material on the subject of QBX in three dimensions and the considerations required for accuracy and acceleration.

2 Mathematical Preliminaries

As a model problem, consider the exterior Neumann problem for the Laplace equation in three dimensions, for a smooth bounded domain Ω. Given continuous Neumann boundary data g, the problem is to find u such that

$$\begin{aligned} \triangle u &= 0 \quad && \text{in } \mathbb{R}^3 \setminus \Omega, \\ \partial_n u &= g \quad && \text{on } \partial\Omega, \\ \lim_{|x|\to\infty} u &= 0. \end{aligned}$$

Here, the notation ∂_n indicates the derivative with respect to the outward unit normal.

The solution to this problem may be represented as $u \coloneqq \mathcal{S}\mu$, a single-layer potential over the boundary $\Gamma = \partial\Omega$, using an unknown density function μ. The properties of the operator $\mathcal{S}$ imply that the Laplace PDE and the far field boundary conditions are immediately satisfied by this representation in the exterior domain. The Neumann boundary condition, on Γ, together with the jump-relations for layer potentials [39] entails that μ satisfies the integral equation of the second kind

$$-\frac{\mu}{2} + \mathcal{S}'\mu = g,$$

where the operator $\mathcal{S}'$ is defined as

$$\mathcal{S}'\mu(x) \coloneqq PV \int_\Gamma \left(\partial_{n(x)}\mathcal{G}(x,y)\right) \mu(y)\, dS(y).$$

After discretization of the integral equation, its solution by iterative methods requires repeatedly applying the operator $\mathcal{S}'$. Similarly, the evaluation of the BVP solution u requires applying the operator $\mathcal{S}$. In this section we will focus on the evaluation of $\mathcal{S}$, although what is said in this section applies with little additional work to $\mathcal{S}'$ or other layer potentials.

When the evaluation (or 'target') point x is sufficiently far from Γ, the approximate evaluation of the integral $\mathcal{S}\mu(x)$ can be accomplished accurately using a high-order composite quadrature rule for smooth functions, as the integrand itself is a smooth function. For x nearer to the boundary, it is well known that

the singularity of the integrand presents resolution problems for smooth quadrature rules, which we address through the use of QBX.

In what follows, we use the notation

$$B_p(r, c) \coloneqq \{y \in \mathbb{R}^3 : |c - y|_p < r\}$$

to denote the open ball with respect to the ℓ^p-norm around center c with radius r. $\overline{B_p}(r, c)$ denotes the closure of that ball. In particular, we make use of $\overline{B_\infty}(r, c)$ and $\overline{B_2}(r, c)$. $\overline{B_2}(r, c)$ denotes the closed Euclidean ball of radius r centered at c. $\overline{B_\infty}(r, c)$ denotes the closed cube of radius r centered at c.

2.1 QBX Discretization

The idea of QBX is to use the smoothness of the potential for purposes of close and on-surface evaluation to recover a high-order accurate approximation everywhere in the domain. This is accomplished through formation of a local expansion of the potential near the source geometry and analytic continuation of the local expansion towards the boundary.

Throughout this paper we make use of spherical harmonic expansions. The expansion of the Laplace potential in spherical harmonics is based on writing the Green's function (2) using the following identity valid for $a, b \in \mathbb{R}^3$ with $|a|_2 < |b|_2$:

$$\mathcal{G}(a, b) = \sum_{n=0}^{\infty} \frac{1}{2n+1} \frac{|a|_2^n}{|b|_2^{n+1}} \sum_{m=-n}^{n} Y_n^m(\theta_a, \phi_a) Y_n^{-m}(\theta_b, \phi_b). \tag{3}$$

Here, (θ_a, ϕ_a) and (θ_b, ϕ_b) refer to the polar and azimuthal spherical coordinates of, respectively, a and b, i.e. $\theta = \cos^{-1}(z/r), \phi = \operatorname{atan2}(y, x)$. The spherical harmonic function Y_n^m of order m and degree n, $|m| \leq n$, is defined as

$$Y_n^m(\theta, \phi) \coloneqq \sqrt{\frac{2n+1}{4\pi} \frac{(n-|m|)!}{(n+|m|)!}} \cdot P_n^{|m|}(\cos\theta) e^{im\phi} \tag{4}$$

where P_n^m is the associated Legendre function of order m and degree n.

The identity (3) may be used in the formation of a *local expansion* centered at a center $c \in \mathbb{R}^3$ as follows. Given a source point $s \in \mathbb{R}^3$, define a doubly-indexed sequence $\langle L_n^m \rangle$ of *local coefficients* by

$$L_n^m \coloneqq \frac{1}{2n+1} \frac{1}{|s-c|_2^{n+1}} Y_n^{-m}(\theta_{s-c}, \phi_{s-c}), \quad |m| \leq n \tag{5}$$

where the subscripted $\theta_{(\cdot)}$ and $\phi_{(\cdot)}$ from this section onward refer to the polar and azimuthal spherical coordinates of the vector argument. Then the local expansion evaluated at a target $t \in \mathbb{R}^3$ may be written as

$$\mathcal{G}(s, t) = \sum_{n=0}^{\infty} \sum_{m=-n}^{n} L_n^m |t-c|_2^n Y_n^m(\theta_{t-c}, \phi_{t-c}).$$

A *p-th order local expansion* is one in which the index of the outer summation goes from 0 to p. (Some authors, e.g. [23, 53], follow the convention of defining the local coefficient (5) using the Y_n^m, the complex conjugate of Y_n^{-m}. Both (5) and the latter definition yield equivalent expansions, since the outer partial sums of (3) are real [53, eqn. (20)].)

Next, we describe the details of QBX. The QBX-based approximation of layer potentials may be thought of as occurring in three distinct steps.

2.2 First Approximation Step: Truncation

In the first stage, a local expansion of the potential is formed and truncated. For a selection of points $\{x_i\}_{i=1}^{N_C/2}$ on the surface Γ, we define a collection of N_C expansion centers $c_i^\pm$

$$c_i^\pm \coloneqq x_i \pm r(x_i)\hat{n}(x_i) \tag{6}$$

where $\hat{n}(x)$ is a unit-length normal vector to the surface Γ at x, and $r(x)$ is a yet-to-be-determined expansion radius.

The local coefficients $\langle (L_i^\pm)_n^m \rangle$ associated with the expansion at $c_i^\pm$ may be defined through the integrals

$$(L_i^\pm)_n^m = \frac{1}{2n+1} \int_\Gamma \frac{\mu(s)}{|s - c_i^\pm|_2^{n+1}} Y_n^{-m}(\theta_{s-c_i^\pm}, \phi_{s-c_i^\pm})\, dS(s). \tag{7}$$

Then the p-th order QBX local expansion at a target $t \in \overline{B_2}(r(x_i), c_i^\pm)$ may be evaluated as

$$\mathcal{S}\mu(t) \approx \sum_{n=0}^{p} \sum_{m=-n}^{n} (L_i^\pm)_n^m |t - c_i^\pm|_2^n Y_n^m(\theta_{t-c_i^\pm}, \phi_{t-c_i^\pm}). \tag{8}$$

The error incurred through the truncation of (8) to order p may be as in Lemma 1. (While the reference [16, Thm. 3.1] discusses the Helmholtz case, the Laplace case follows analogously.)

Lemma 1 (QBX truncation error in three dimensions, cf. [16, Thm 3.1])**.** *Suppose that Γ is smooth, non-self-intersecting and let $r > 0$. Let the local coefficients $\langle L_n^m \rangle$ be defined as in (7) and the expansion centers $\{c_i^\pm\}_{i=1}^{N_C/2}$ as in (6). Let $c \in \{c_i^+, c_i^-\}$ be a center for which $\overline{B}(r, c) \cap \Gamma = \{x_i\}$ for some $1 \le i \le N_C/2$. Then for each $p > 0$ and $\delta > 0$, there is a constant $M_{p,\delta}$ such that*

$$\left| \mathcal{S}\mu(x_i) - \sum_{n=0}^{p} \sum_{m=-n}^{n} L_n^m |x_i - c|_2^n Y_n^m(\theta_{x_i-c}, \phi_{x_i-c}) \right| \le M_{p,\delta} r^{p+1} \|\mu\|_{W^{3+p+\delta,2}(\Gamma)}. \tag{9}$$

2.3 Second Approximation Step: Quadrature

In the second stage, we apply numerical quadrature to discretize the integrals for the computation of the expansion coefficients in (7). We assume that the smooth, non-self-intersecting surface Γ is tessellated into individual, disjoint surface elements Γ_k so that

$$\Gamma = \bigcup_{k=1}^{K} \Gamma_k.$$

Each Γ_k is described by a smooth mapping function $\Psi_k : E \to \mathbb{R}^3$, where E is a two-dimensional reference element. We assume that the mapping Jacobian Ψ_k' has full rank everywhere. The integral (7) can then be split into contributions from each element as

$$(L_i^\pm)_n^m = \frac{1}{2n+1} \sum_k \int_{\Gamma_k} \frac{\mu(s)}{|s - c_i^\pm|_2^{n+1}} Y_n^{-m}(\theta_{s-c_i^\pm}, \phi_{s-c_i^\pm})\, dS(s),$$

and, for each element Γ_k, written as an integral over the reference element E using

$$\begin{aligned} &\int_{\Gamma_k} \frac{\mu(s)}{|s - c_i^\pm|_2^{n+1}} Y_n^{-m}(\theta_{s-c_i^\pm}, \phi_{s-c_i^\pm})\, dS(s) \\ &\qquad = \iint_E \frac{\mu(\Psi_k(s_1, s_2))}{|\Psi_k(s_1, s_2) - c_i^\pm|_2^{n+1}} \tilde{Y}_n^{-m}(\Psi_k(s_1, s_2) - c_i^\pm) |\partial_{s_1} \Psi_k(s_1, s_2) \times \partial_{s_2} \Psi_k(s_1, s_2)|_2\, ds_1\, ds_2 \end{aligned} \tag{10}$$

where we have introduced the notation $\tilde{Y}_n^{-m}(x) \coloneqq Y_n^{-m}(\theta_x, \phi_x)$ for brevity.

The integral (10) may be discretized using quadrature over the reference element E. As an example, we will assume the reference element is the bi-unit tensor product element $[-1, 1]^2$ and consider the discretization of the integral with a tensor product quadrature rule for smooth functions. (In practice, our implementation uses a triangular reference element with nodes and weights based on [58].) A tensor product rule is based on iterated evaluation of a one-dimensional q-point quadrature rule

$$Q_q \left\{ \int_{-1}^{1} f(y)\, dy \right\} = \sum_{j=1}^{q} w_j f(y_j). \tag{11}$$

After repeated application of (11), the integral (10) becomes

$$\sum_{j_1=1}^{q}\sum_{j_2=1}^{q} w_{j_1} w_{j_2} \frac{\mu(\Psi_k(y_{j_1}, y_{j_2}))}{|\Psi_k(y_{j_1}, y_{j_2}) - c_i^{\pm}|_2^{n+1}} \tilde{Y}_n^{-m}(\Psi_k(y_{j_1}, y_{j_2}) - c_i^{\pm}) |\partial_{s_1}\Psi_k(y_{j_1}, y_{j_2}) \times \partial_{s_2}\Psi_k(y_{j_1}, y_{j_2})|_2. \quad (12)$$

Neglecting geometry, it is straightforward if tedious to obtain estimates of the quadrature error incurred in (12). Such estimates are roughly analogous to prior results for curves embedded in two dimensions [16]. Compared with the two-dimensional case, the main difference in the element-wise estimate is the loss of a power of r, owing to the difference in free space Green's functions:

Lemma 2 (QBX quadrature error for tensor product elements)**.** *Let $\Gamma_k = [0,h]^2 \times \{0\}$ be a flat, square element and $\Psi_k : [-1,1]^2 \to [0,h]^2 \times \{0\}$ be given by $\Psi_k(x_1, x_2) = \frac{1}{2}h(x_1 + 1, x_2 + 1, 0) \in \mathbb{R}^3$. Let the expansion center be at a distance $r > 0$ from Γ_k, and consider a q-point Gauss-Legendre rule with points $\{y_j\}_{j=1}^q$ and weights $\{w_j\}_{j=1}^q$. Then there is a constant $C_{p,q} > 0$ such that for all $h > 0$ and $r > 0$*

$$\left| \sum_{j_1=1}^{q}\sum_{j_2=1}^{q} w_{j_1} w_{j_2} \frac{\mu(\Psi_k(y_{j_1}, y_{j_2}))}{|\Psi_k(y_{j_1}, y_{j_2}) - c_i^{\pm}|_2^{n+1}} \tilde{Y}_n^{-m}(\Psi_k(y_{j_1}, y_{j_2}) - c_i^{\pm}) |\partial_{s_1}\Psi_k(y_{j_1}, y_{j_2}) \times \partial_{s_2}\Psi_k(y_{j_1}, y_{j_2})|_2 \right.$$
$$\left. - \int_{\Gamma_k} \frac{\mu(s)}{|s - c_i^{\pm}|_2^{n+1}} Y_n^{-m}(\theta_{s-c_i^{\pm}}, \phi_{s-c_i^{\pm}})\, dS(s) \right| \leq C_{p,q}\,(1+h) \left(\frac{h}{4}\right)^{2q+1} \frac{1}{r^{n+1}} \left[1 + \left(\frac{1}{r}\right)^{2q}\right] \|\mu\|_{C^{2q}}. \quad (13)$$

This estimate pertains to the error in the quadrature contribution of one element to the coefficient $(L_i^{\pm})_n^m$. Similar estimates can be obtained for curved elements, although one must take into account the effects of the occurring mapping derivatives, both from the use of the substitution rule, and within the argument of μ. When the error contribution of the form (13) is summed over all the elements, this yields a quadrature error estimate for the coefficient $(L_i^{\pm})_n^m$. A factor of r^{-n} in (13) is dampened by the term $|t - c_i^{\pm}|_2^n$ when evaluating the summation (8) for the local expansion of the single-layer potential, leaving a quadrature error that scales essentially like $O((h/(4r))^{2q+1}\|\mu\|_{C^{2q}})$ for small enough h and r. An analysis that yields significantly more precise estimates for tensor product rules over elements can be found in [35].

2.4 Third Approximation Step: Acceleration

The third approximation applied in the rapid, QBX-based evaluation of layer potentials like (1) arises due to acceleration. The formation of local expansions (8) at all centers covering a neighborhood of Γ requires $O(NM)$ operations, where N is the number of source points and M is the number of target points. Interpreting (8) as the local expansion of a potential due to a finite set of source charges in space suggests that such expansion could be amenable to acceleration with the Fast Multipole Method (FMM).

Recall that the original version of the FMM (e.g. [24]) is designed to evaluate *point potentials*, which are potentials of the form

$$\Phi(x_i) \coloneqq \sum_{j=1}^{N} w_j \mathcal{G}(x_i, y_j) \quad (i = 1, \ldots, M). \quad (14)$$

Here, $\{x_i\}_{i=1}^M \subseteq \mathbb{R}^3$ is the set of target points, and $\{y_j\}_{j=1}^N \subseteq \mathbb{R}^3$ is the set of source points with the weights $\{w_j\}_{j=1}^N \subseteq \mathbb{R}$. We shall call FMMs which evaluate these types of summations 'point FMMs' in the remainder of this paper. In contrast, FMMs directed toward evaluation of QBX expansions for global QBX ([48, 49, 57]) can be described as a modification of the FMM where QBX centers are treated as a special kind of target, at which the FMM *forms a local expansion* rather than evaluating a point potential.

The capability of forming a local expansion of a point potential is an algorithmic component of the point FMM used in the far field approximation. Because the algorithmic machinery is already present in the point FMM, it would be appear to be a fairly natural step to modify the point FMM to form local expansions at the QBX centers. The first published version of the QBX FMM [48] operationally follows the point FMM algorithm, but replaces the point evaluations at the QBX centers with the formation of a local expansion. In particular, this allows it to reuse the intermediate local expansions formed by the FMM for purposes of QBX evaluation.

Unfortunately, for a given FMM expansion order, the accuracy attained by the algorithm in [48] is generically lower than what would be expected of a point FMM. Specifically, one does not observe a purely additive error of magnitude proportional to $c^{p_{\text{fmm}}+1}$, where c is a convergence factor (e.g. $c = 1/2$ for the Laplace FMM in two dimensions [46] or $c = 3/4$ for the Laplace FMM in three dimensions [45]) and p_{fmm} is the approximation order used. The reason for this loss of accuracy is discussed in detail in [57]. In short, for accurate evaluation of expansions of the form (8), *all* local coefficients need to be adequately approximated. This entails the ability not only to approximate a point potential but also *its derivatives* to a certain order of accuracy. The point FMM was not designed with the goal of providing accuracy estimates for this evaluation pattern. The procedure suggested in [48] is to set the FMM order to ${p_{\text{fmm}}}' = p_{\text{fmm}} + p_{\text{add}}$, where p_{fmm} is the FMM order required for the point FMM to achieve a specified tolerance, and $p_{\text{add}} > 0$ is an empirically determined quantity that depends on p_{fmm} and the accuracy tolerance. This strategy works in practice, although it comes with some disadvantages. First, higher order multipole and local expansions are expensive and more difficult to implement stably [19]. Second, error estimates covering this use are, to the best of our knowledge, not available.

In [57], an extension to the QBX FMM of [48] was developed to provide accuracy guarantees similar to the point FMM. This was done starting with the analytical result that QBX disks act like 'targets with extent.' What this suggests in practice is that the 'near field' of a QBX disk should be redefined to be proportional to the size of the disk, so that the field in a region nearby each disk is evaluated directly. The tree that is built over the computational domain needs to be aware of this change.

Perhaps the conceptually simplest scheme that matches the accuracy of the point FMM is to enforce that a QBX disk (in the two dimensional context, or ball in three dimensions) must be contained entirely inside the box that owns the disk, so that any QBX disk that cannot fit in a child box remains in the parent box. However, enforcing that a QBX disk must fit entirely inside a box is computationally expensive, since disks may be suspended at 'high' levels of the tree (near the root) leading to direct interactions with large parts of the geometry. To reduce the cost associated with suspending QBX disks at high levels, the algorithm in [57] allows QBX disks to protrude beyond their boxes by a fixed multiple of the box size. This allows them to settle lower down the tree (away from the root) at a level where the box size is commensurate with their diameter. This modification, which was termed a *target confinement rule*, retains the linear scaling of the FMM under mild assumptions on the geometry, while also permitting for control over the error introduced by acceleration.

The scheme described in this paper is a generalization and enhancement of the scheme described in [57], which we term the 'GIGAQBX FMM', for 'Geometric Global Accelerated QBX'. In Section 3, we describe a framework for ensuring accuracy in the application of QBX with no acceleration over arbitrary smooth geometries. Error estimates for FMM translations are derived in Section 4. The algorithm is described in Section 5. We close with numerical experiments in Section 6.

3 Accuracy Control for QBX on Surfaces

Since the cost of computational methods dealing with three-dimensional geometries is typically far greater than that of methods applied to two-dimensional geometries, and since that cost is directly related to the resolution supplied, it is not surprising that careful control of resolution and accuracy plays an important role in maintaining efficiency. For QBX, two related conditions must be satisfied to ensure accurate evaluation of the layer potential at any point in $\mathbb{R}^3$. First, the truncation and quadrature errors at the QBX centers must be adequately controlled, even given a geometry that does not necessarily satisfy the preconditions of Lemma 1 and Lemma 2. Second, every target needing QBX evaluation must be associated to a QBX center.

In this section, we describe a computational framework for establishing these conditions. While our presentation focuses on the setting of source *surfaces* embedded in three-dimensional space, the described approach has an immediate analog for curves embedded in two-dimensional space, permitting the computationally unified treatment of both cases.

The prior geometry processing scheme introduced in [48] had the potential to cause what one might call a 'chain reaction' of refinements, where a refinement based on insufficient quadrature resolution might trigger an element bisection, in turn moving expansion centers associated with the bisected elements, which might trigger further resolution-based refinements, and so on, in particular on surfaces. The main contribution of

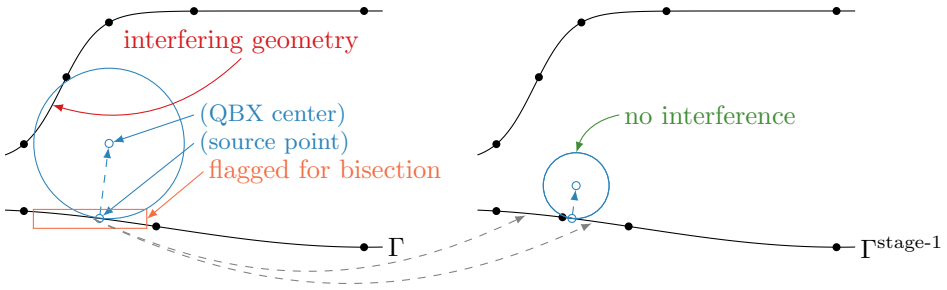


Figure 1: An illustration of stage-1 refinement for a two dimensional geometry. Element (panel) boundaries are show with black dots. The element associated with the QBX disk shown is bisected due to the presence of interfering geometry.

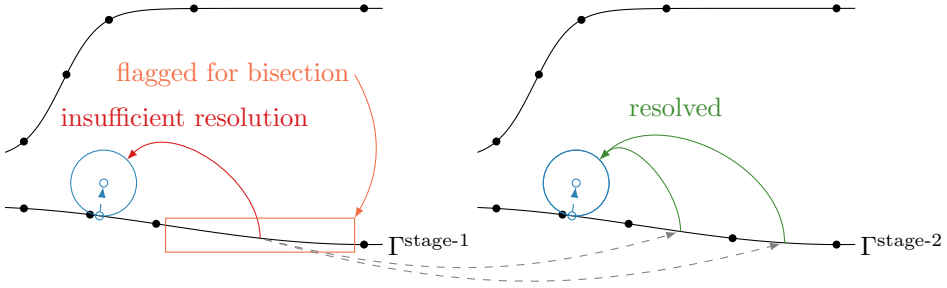


Figure 2: An illustration of stage-2 refinement for a two dimensional geometry, continuing from the previous figure. The element adjacent to the QBX disk shown is bisected because of insufficient quadrature resolution.

this section is a *multi-stage* approach that not only separates concerns between different causes for refinement, but also entirely avoids unnecessary 'chain reactions' between them.

3.1 Overview

We commence our discussion with an outline of a procedure for efficiently detecting and remedying potential sources of truncation and quadrature inaccuracy in arbitrary smooth geometries. From an initial, user-supplied, unstructured mesh, the process creates a set of three related, unstructured discretizations satisfying different invariants. We term these the '*stage-1 discretization*', the '*stage-2 discretization*', and the '*stage-2 quadrature discretization*'.

For concreteness, we describe these discretization in terms of triangles, with the understanding that generalizations to other types of reference element (e.g. squares) are expected to be straightforward. The stage-1 and stage-2 discretizations are interpolatory/unisolvent, i.e. a unique polynomial in the mapped polynomial space $P^{N_t} \circ \Psi_k^{-1}$ may be reconstructed from the nodal degrees of freedom, where N_t is the polynomial degree of the 'target function space' in which layer potentials are evaluated. We choose reference unit nodes following [55] and use nodal values at their mapped counterparts as degrees of freedom for the representation of the density and the geometry. Interpolation operators transport information, particularly on-surface density values, through the discretizations in the following order:

$$\text{stage-1} \quad \rightarrow \quad \text{stage-2} \quad \rightarrow \quad \text{stage-2 quadrature.}$$

Shortly, we will summarize the primary features of these discretizations and the algorithms used to obtain them. Afterwards, the remainder of this section supplies detailed analysis and algorithms, particularly for stage-1 refinement and stage-2 refinement (and additionally target association).

A key mechanism for maintaining scalability for algorithms in this section which require examination of non-local portions of the geometry is the *area query*, introduced in [48]. Briefly, given a spatial partitioning of the geometry into an octree and, for each box in the octree, a stored list of adjacent, equal-or-larger *peer*

boxes, an area query can efficiently find the set of leaf boxes intersecting a cubic region $\{x \in \mathbb{R}^3 : |x-c|_\infty \leq r\}$, for given c and r. For details, we refer to Section A in the appendix.

Stage-1 Discretization. Algorithm 1 of Section 3.3 produces the stage-1 discretization from the user-supplied mesh. The stage-1 discretization is a locally refined mesh fitting the geometry description which ensures that (a) sufficient resolution to represent the density and the geometry is available, and that (b) the assumptions of Lemma 1 are satisfied, i.e. specifically that the expansion balls of (6) are undisturbed by quadrature sources (cf. Section 3.3). Expansion radii $r(x)$ are chosen proportional to a resolution measure of the stage-1 discretization (cf. Section 3.2). Thus, locally bisecting triangular elements $\triangle$ associated with expansion balls disturbed by other geometry will shrink the associated expansion ball, helping to ensure that, potentially after a number of refinement cycles, the expansion ball clears the interfering geometry. Figure 1 gives an illustration of this bisection process on a portion of a two-dimensional geometry.

The stage-1 discretization also incorporates a novel '*scaled-curvature criterion*' to control for truncation error based on an empirically effective heuristic involving the local curvature of the mesh elements. See Section 3.3.1.

Stage-2 Discretization. The stage-2 discretization is generated using Algorithm 2 of Section 3.4, starting with the stage-1 discretization. The role of the stage-2 discretization is to ensure that enough quadrature resolution is available to satisfy the resolution requirement implied by the estimate (13) when applied between close elements, i.e. that the quadrature contribution to the approximation of the layer potential is asymptotically as accurate from nearby elements as it is from the element that spawns the QBX center. As with the stage-1 discretization, this discretization is obtained through iterative bisection of offending source elements. As an illustration of the potential issues that this discretization controls for, consider the two-dimensional geometric situation depicted in Figure 2. In this figure, the situation illustrated on the left leads to inaccuracies as the contribution from the large source element is not adequately resolved relative to the size of the target QBX disk, and bisection suffices to ensure adequate resolution.

Stage-2 Quadrature Discretization. The stage-2 quadrature discretization results from oversampling (i.e., increasing the order) of the source quadrature nodes of the stage-2 discretization. (Thus, the stage-2 quadrature discretization shares the same mesh as the stage-2 discretization.) The stage-2 quadrature discretization is optimized for the highest possible quadrature order achievable at a given node count, to control the quadrature error in Lemma 2, at the expense of unisolvence. In three dimensions, our implementation uses quadrature nodes and weights for the triangle based on [58].

Target Association. Lastly, we require a tool to compute a mapping from targets needing QBX evaluation to QBX centers. Algorithm 3 of Section 3.5 provides this capability. Compared with the similar target association algorithm in [48], this algorithm presents a simplified procedure for locating sources or QBX centers close to a given target, at the expense of performing two area queries instead of one.

3.2 Quantifying Quadrature Resolution on Surfaces

At the core of our accuracy control mechanism lies a measurement of quadrature resolution in the underlying high-order quadrature used to drive QBX. In our case, these are quadrature rules based on [58]. To accomplish this measurement, we define a modified element mapping $\tilde{\Psi}_k : \tilde{E} \to \mathbb{R}^3$, where $\tilde{E}$ is the 'bi-unit' equilateral triangle with vertices

$$v_1 = \begin{bmatrix} -1 \\ -1/\sqrt{3} \end{bmatrix}, \quad v_2 = \begin{bmatrix} 1 \\ -1/\sqrt{3} \end{bmatrix}, \quad v_3 = \begin{bmatrix} 0 \\ 2/\sqrt{3} \end{bmatrix}$$

serving as the modified reference element. We define a function

$$\eta_k(x) \coloneqq 2\sigma_1(\tilde{\Psi}_k'(\tilde{\Psi}_k^{-1}(x))) \quad \text{for } x \in \Gamma_k, \tag{15}$$

where $\sigma_1(A)$ denotes the largest singular value of a matrix A. The factor of two normalizes out the edge length of $\tilde{E}$. $\eta_k(x)$ computes an approximate local 'stretch factor' of the mapping $\tilde{\Psi}_k$ at the point x. Since

η_k may be discontinuous between adjacent elements, it is only unambiguously defined when the point x does not lie on the boundary of Γ_k, necessitating the subscript k to avoid ambiguity. η_k can serve as an analog of the 'speed' of the one-dimensional parametrization of a curve segment. It is crucial that $\tilde{E}$ be equilateral to ensure that η_k measures resolution independently of vertex ordering. We further define

$$\eta_k \coloneqq \max_{x \in \Gamma_k} \eta_k(x)$$

as a per-element maximum of the corresponding per-source-point function.

This resolution measure provides the basis for our choice of the expansion radii

$$r_k \coloneqq \frac{1}{2} \eta_k^{\text{stage-1}}. \tag{16}$$

The quantity $\eta_k^{\text{stage-1}}$ is simply η_k of (15) computed in reference to the stage-1 discretization, defined below. Allowing a rough analogy between the 'panel length' h_k of [48] and η_k makes the choices of expansion radii of [48] coincide with ours.

3.3 Stage-1 Refinement: Managing Truncation Error

Lemma 1 requires that the expansion ball be clear of source geometry except for the target point. For smooth, non-self-intersecting geometries, our method ensures that this condition is satisfied without explicit user involvement, through an approach analogous to that in [48]. Algorithm 1 describes the procedure.

Algorithm 1 operates by using area queries to find all source geometry that protrudes into the QBX expansion balls (cf. Section 2.2) and marking the elements that spawned the obstructed expansion balls for bisection. Bisection will lead the expansion radius (16) to shrink by way of a reduction of η_k, both of which will drop by a factor of 2 as a result of bisection. This is repeated until no more interfering geometry is found. To prevent the source point that spawned the center from being found and causing refinement, we reduce the size of the queried area by a factor of $\varepsilon_{\text{exp-disturb}}$. In practice, we choose $\varepsilon_{\text{exp-disturb}} = 0.025$. The discretization appears to be fairly insensitive to the choice of this parameter, which is plausible given our chosen quadrature margins (cf. Section 3.4). Values as large as 0.2 empirically cause little or no loss in accuracy.

Algorithm 1: Bisect source elements whose expansion balls encounter interfering source geometry

---

Require: The geometry discretized as a set of targets, sources, and expansion centers.
Ensure: By repeated bisection that the expansion radii r_k are sufficiently small that $\overline{B_2}(r_k, c_i^{\pm}) \cap \Gamma = \{x_i\}$ for $x_i \in \Gamma_k$ (as $\varepsilon_{\text{exp-disturb}} \to 0^+$).

repeat
  Create an octree on the computational domain containing all sources, expansion balls, and targets.
  for all expansion balls $\overline{B_2}((1 - \varepsilon_{\text{exp-disturb}})r_k, c_i^{\pm})$ **do**
    Perform an area query of radius r_k centered at $c_i^{\pm}$.
    if the query returned a source point s such that $|c_i^{\pm} - s|_2 < (1 - \varepsilon_{\text{exp-disturb}})r_k$ **then**
      Mark the element containing x_i for bisection.
    end if
  end for
  if elements were marked for bisection **then**
    Bisect the marked elements.
  end if
until no elements were marked for bisection

3.3.1 Truncation Error and 'Scaled-Curvature'

The realization that interference from nearby geometry, through derivatives of the surface parametrization, contributes to the degradation in the truncation error motivates a 'localized' criterion that controls the

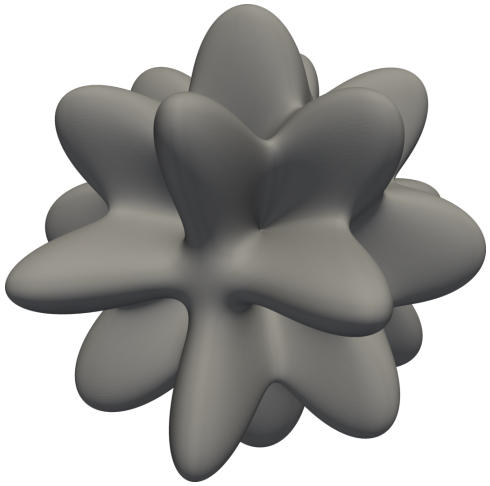


Figure 3: The 'urchin' test geometry γ_8 that we use for many computational experiments in this paper. See (31) for the warping function used to obtain γ_8. The geometry is represented by an unstructured triangular mesh. Starting with an icosahedron, the high-order triangular elements are repeatedly warped and adaptively bisected to resolve the element mapping functions Ψ_k. This resolved geometry is then processed according to Section 3.

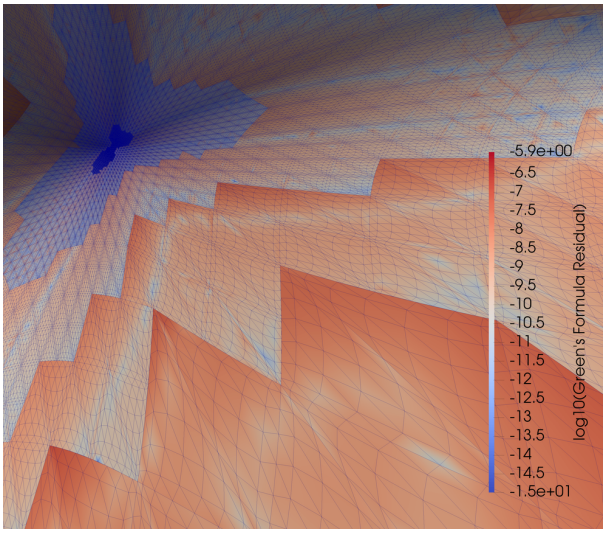


Figure 4: Five levels of scaled-curvature-guided refinement of the stage-1 mesh, shown on a small 'trough' part of γ_8. See Section 3.3 for details of the refinement method. The coloring shows the base-10 logarithm of the residual in Green's formula $\mathcal{S}(\partial_n u) - \mathcal{D}(u) = u/2$ (i.e., roughly, the number of accurate digits). 'Red' indicates largest residual, corresponding to around six accurate digits.

'amount of curvature' on each element of Γ.

The motivation for this may be most clearly explained by considering QBX in the two-dimensional case. Consider a smooth closed non self-intersecting curve $\Gamma \subseteq \mathbb{C}$ with arc length parametrization $w : [0, L] \to \Gamma$ such that $\overline{B_2}(0, r) \cap \Gamma = \{z\}$ (Figure 5). The single-layer potential evaluated at z for a density function $\sigma : \Gamma \to \mathbb{R}$ takes the form

$$S\sigma(z) = -\frac{1}{2\pi} \int_0^L \sigma(w(t)) \log |w(t) - z| \, dt.$$

We use the complex-valued logarithm, which satisfies $\operatorname{Re} \log y = \log |y|$ for all $|y| > 0$, to rewrite the above as

$$S\sigma(z) = -\frac{1}{2\pi} \operatorname{Re} \int_0^L \sigma(w(t)) \log \left(1 - \frac{z}{w(t)}\right) dt + S\sigma(0).$$

By expanding the kernel $\log(1 - z/w(t))$ in a Taylor series, we obtain the expression for the truncation error in an n-th order QBX expansion centered at the origin, which is:

$$e_n(z) = -\frac{1}{2\pi} \operatorname{Re} \int_0^L \sigma(w(t)) \sum_{k=n+1}^{\infty} \frac{1}{k} \left(\frac{z}{w(t)}\right)^k dt.$$

We integrate this expression by parts $1 \le p \le n$ times, assuming $\sigma \in C^\infty(\Gamma)$. At each step we replace the term w^{-k} with $-1/((k-1)w')\partial_t w^{-k+1}$, obtaining [16]

$$e_n(z) = -\frac{1}{2\pi} \operatorname{Re} \left((-z)^p \int_0^L D_t^p [\sigma(w(t))] \sum_{k=n+1}^{\infty} \frac{(k-p-1)!}{k!} \left(\frac{z}{w(t)}\right)^{k-p} dt \right)$$

where the differential operator D_t is defined through $D_t g = \partial_t [g(t)/w'(t)]$. For example, the values of

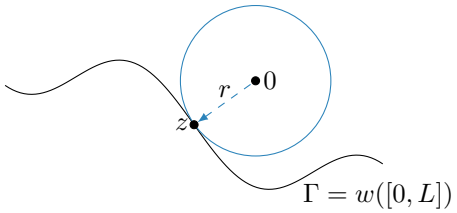


Figure 5: 2D QBX geometric evaluation scenario for the single-layer potential $\mathcal{S}\sigma$ in Section 3.3.1, for a segment of the closed curve Γ.

$D_t^p[\sigma(w(t))]$ for $p = 1$ and $p = 2$ are

$$D_t[\sigma(w(t))] = \sigma'(w(t)) - \frac{\sigma(w(t))w''(t)}{w'(t)^2},$$
$$D_t^2[\sigma(w(t))] = \sigma''(w(t)) - \frac{2w''(t)\sigma'(w(t))}{w'(t)^2} - \frac{w^{(3)}(t)\sigma(w(t))}{w'(t)^3} + \frac{3\sigma(w(t))w''(t)^2}{w'(t)^4}.$$

Via these expressions, it is clear that the higher derivatives of the curve parametrization w have some influence on the truncation error. In the expression for the truncation error, we assume without loss of generality that $|w'(t)| = 1$. The next higher derivative $w''(t)$, whose magnitude represents the curvature at parameter t, is the first derivative whose magnitude is not controlled. However, the contribution of the term $w''(t)$ to the truncation error may be dampened by ensuring that r is chosen to locally enforce that $|rw''(t)| \leq \kappa_{\max}$, for some constant $\kappa_{\max} > 0$, in a neighborhood of the target. Many other factors that enter into the truncation error, so this is at best a heuristic motivation. Nevertheless, we have been able to develop this insight into a practically useful criterion.

We realize this criterion in three dimensions as follows. Let $k_1(x)$ and $k_2(x)$ be the principal curvatures of Γ at $x \in \Gamma$. We require that

$$\kappa_k(x) \coloneqq \max(|k_1(x)|, |k_2(x)|) \cdot \eta_k(x) \leq \kappa_{\max} \quad (x \in \Gamma_k), \tag{17}$$

noting that $\kappa_k(x)$ is unit-less and invariant to scaling. A rough geometric interpretation is that the condition stipulates that a single source element may at most cover a certain angle of a tangent circle within the planes of principal curvature. Since $\eta_k(x)$ is reduced by bisecting elements and since $k_1(x)$ and $k_2(x)$ are independent of parametrization, $\kappa_k(x)$ can be effectively managed through refinement by bisection.

In our computational experiments, we use $\kappa_{\max} = 0.8$ with good success. As an example, consider the mesh of Figure 4, which shows a small section of the test geometry γ_8 (shown in Figure 3, see Section 6 for a more thorough description). The mesh shown in the figure exhibits five levels of bisection-based refinement that were triggered by the criterion (17). The surface coloring in the figure shows a logarithmic measure of the accuracy with which QBX evaluates layer potentials on Γ, where the color red indicates the highest levels of error encountered, corresponding to roughly six accurate digits. We observe that each successive level of refinement exhibits growth of the error up to roughly the 'red' level of accuracy, at which point further refinement is triggered. At least in the scenario shown, (17) exhibits remarkable sharpness and reliability.

We refer to (17) as the '*scaled-curvature criterion*'. Its application requires the reasonably accurate evaluation of two derivatives of the geometry, which may not be practical in all settings—notably when Γ is discretized using purely affine element mappings. Nonetheless, the availability of such a criterion yields substantial efficiency gains in the application of the QBX in high-accuracy settings, leading us to report on its discovery in this context. Despite strong heuristic motivation and encouraging computational results, the evidence supporting the criterion is empirical at this point. Furthermore, since this criterion leads to improvements of the resolution of the integrand, the application of this criterion also leads to reduction of the quadrature error, but at present it is not clear which of these error reduction effects (quadrature or truncation) dominates. We leave a detailed discussion and potential proofs of its properties for future work.

The steps in this section work together to manage truncation error under the assumption that all coefficient integrals are computed exactly. The resulting discretization is called the *stage-1 discretization*.

3.4 Stage-2 Refinement: Ensuring Accurate Coefficient Quadrature

To ensure the accurate computation of the coefficient integrals associated with the expansion centers spawned by the stage-1 discretization, we introduce a separate *stage-2 discretization* that may, depending on some criteria, be bisected into smaller elements than are present in the stage-1 discretization, to provide additional resolution for the high-order quadrature underlying QBX. An analogous refinement step was present in the geometry processing in [48], however an important difference between that scheme and ours is that it used only a single discretization, creating an artificial interdependence between quadrature-based refinement and the choice of the expansion radii, impacting the robustness of the refinement procedure. Our approach leaves the expansion radii fixed once the stage-1 discretization is determined, removing this unnecessary entanglement of the two stages.

In this section, it will be necessary to distinguish between different refinement iterations of the stage-2 mesh. We refer to different refinement iterations numbered using a superscript, e.g. the notation for the k-th element of the n-th iteration is $\Gamma_k^{\text{stage-2},n}$. Each iteration consists of K_n elements, so that $\Gamma = \bigcup_{k=1}^{K_n} \Gamma_k^{\text{stage-2},n}$. The initial iteration $\Gamma^{\text{stage-2},1}$ is the same as the stage-1 mesh which is the output of Algorithm 1.

Let a center c be spawned by some target element $\Gamma_k^{\text{stage-2},1}$. We seek to control the quadrature error at c due to a source panel $\Gamma_j^{\text{stage-2},1}$, with $j \neq k$ in general. Our primary concern in this section is the quadrature error contribution when $\Gamma_j^{\text{stage-2},1}$ is 'close' to c, i.e., the minimal Euclidean distance $d_2(\Gamma_j^{\text{stage-2},1}, c)$ between the center c and $\Gamma_j^{\text{stage-2},1}$ is sufficiently small to threaten accuracy. We can reexamine and further annotate the quadrature estimate of Lemma 2 to obtain a quadrature estimate for this regime. We assume that the quadrature error essentially takes the form

$$|\text{quadrature error}| \leq C \left(\frac{h_j^{\text{stage-2},1}}{4 d_2(\Gamma_j^{\text{stage-2},1}, c)} \right)^Q \|\mu\|, \quad 1 \leq j \leq K_1 \tag{18}$$

where Q is the order of accuracy of the quadrature ($Q = 2q + 1$ in Lemma 2). That is, the main factors governing the error are the ratio of the source panel size $h_j^{\text{stage-2},1}$ to the center distance, the density norm $\|\mu\|$ (with the choice of norm depending on the quadrature rule), and the order of quadrature accuracy Q. For simplicity, we may consider $h_j^{\text{stage-2},1}$ asymptotically equivalent to $\eta_j^{\text{stage-2},1}$, our quadrature resolution measure.

For concreteness, let a tolerance on the quadrature error *relative to* $\|\mu\|$ be given by ε. We seek to ensure that, for all centers c, refinement produces an iteration n of the discretization such that for all centers $c \in \{c_i^{\pm}\}_{i=1}^{N_C/2}$, we have a quadrature order Q such that

$$C \left(\frac{h_j^{\text{stage-2},n}}{4 d_2(\Gamma_j^{\text{stage-2},n}, c)} \right)^Q \leq C' \left(\frac{\eta_j^{\text{stage-2},n}}{4 d_2(\Gamma_j^{\text{stage-2},n}, c)} \right)^Q \leq \varepsilon, \quad 1 \leq j \leq K_n. \tag{19}$$

The change of leading constant from C to C' absorbs the asymptotic factors involved when switching from $h_j^{\text{stage-2},1}$ to $\eta_j^{\text{stage-2},1}$.

In the case of the 'self-interaction', i.e. when the center c was spawned by a point on the element $\Gamma_k^{\text{stage-2},1}$, for some $1 \leq k \leq K_1$, we may combine the fact that $d_2(\Gamma_k^{\text{stage-2},1}, c) = r_k$ with our choice (16) of r_k, to obtain that (19) implies we must assure

$$C' \left(\frac{\eta_k^{\text{stage-2},1}}{2\eta_k^{\text{stage-1}}} \right)^Q \leq \varepsilon. \tag{20}$$

For efficiency, we seek to avoid having to refine any element to evaluate its self-interaction absent other constraints. This implies, temporarily assuming in this situation that $\eta_k^{\text{stage-2},1} = \eta_k^{\text{stage-1}}$, that there is precisely one free parameter to attain (20), the order of accuracy of quadrature Q.

For only the self-interaction (20), it would suffice to simply choose Q to provide the required accuracy. However, the global nature of our algorithm compels us to use the same quadrature resolution for *all* targets, and so this simple strategy is not necessarily sufficient on its own. Instead, for the benefit of the treatment of the 'non-self interaction' from other elements whose resolution may differ, we choose a higher value of Q

so that the coefficient integrals for a hypothetical source element larger by a factor of (4/3) would still attain the required level of accuracy in its coefficient integrals. In other words, we actually choose Q to assure the stronger bound

$$C'\left(\frac{(4/3)\eta_k^{\text{stage-2,1}}}{2\eta_k^{\text{stage-1}}}\right)^Q \leq \varepsilon. \tag{21}$$

To ensure accuracy of the non-self-interactions, say for all $1 \leq j \leq K_1$, we detect expansion centers within a distance of $0.5 \cdot (3/4)\eta_j^{\text{stage-2,1}}$ from $\Gamma_j^{\text{stage-2,1}}$ using area queries originating from source points on each $\Gamma_j^{\text{stage-2,1}}$. If any such centers are found, $\Gamma_j^{\text{stage-2,1}}$ is bisected. The bisection reduces $\eta_j^{\text{stage-2,1}}$ by a factor of two, producing $\eta_{j'}^{\text{stage-2,2}} = \frac{1}{2}\eta_j^{\text{stage-2,1}}$ for all refined elements $1 \leq j' \leq K_2$ whose parent element is j. Bisection does *not* affect the placement of any expansion centers, so that only the numerator of the bound (18) is affected.

After a sufficient number (say n) of iterations, the refinement process ensures that we always have $d(\Gamma_j^{\text{stage-2},n}, c) \geq 0.5 \cdot (3/4)\eta_j^{\text{stage-2},n}$, so that

$$C\left(\frac{h_j^{\text{stage-2},n}}{4d_2(\Gamma_j^{\text{stage-2},n}, c)}\right)^Q \leq C'\left(\frac{\eta_j^{\text{stage-2},n}}{4 \cdot 0.5 \cdot (3/4)\eta_j^{\text{stage-2},n}}\right)^Q = C'\left(\frac{(4/3)\eta_j^{\text{stage-2},n}}{2\eta_j^{\text{stage-2},n}}\right)^Q \leq \varepsilon, \quad 1 \leq j \leq K_n$$

where Q was chosen above so as to ensure accurate quadrature in this case, cf. (21) and the text before it. One particular consequence of these parameter choices is that, at a resolution change in the stage-1 discretization (perhaps as the result of bisection), the stage-2 refinement scheme just described produces a 'buffer zone' of stage-2-refined elements around the stage-1 refinement fringe.

The halving of $\eta^{\text{stage-2}}$ through bisection implies that the set of 'endangered' centers found in the current iteration will be equal to or a superset of that found in the following iteration. For smooth, non-self-intersecting geometries, the associated procedure, detailed in Algorithm 2, is guaranteed to terminate. As a last step, the stage-2 discretization resulting from Algorithm 2 is upsampled to use a sufficient number of quadrature nodes p_{quad} to achieve order of accuracy Q, obtaining the *stage-2 quadrature* discretization whose nodes are used as source particles for our fast algorithm, detailed in Section 5.

The objective of this contribution is to clarify the asymptotic relationships between the geometric variables, not to provide concrete estimates of the constants involved. As such, we give any specific factors (such as in (16)) merely for concreteness. We claim that the choices described here are adequate to illustrate the behavior of the scheme, and to obtain a practically viable method for layer potential evaluation. However, we make no claim of optimality for the chosen parameters. The contribution [35] contains more precise bounds suggesting that a fully quantitative understanding may be attainable. We leave this for future work.

Algorithm 2: Bisect stage-2 source elements until sufficient quadrature resolution is available

---

Require: The stage-1 discretization has been determined in accordance with Section 3.3.
Ensure: The quadrature accuracy condition (19) holds for all centers and all source elements $\Gamma_k^{\text{stage-2}}$.

Initialize the stage-2 discretization $\Gamma^{\text{stage-2}}$ to be identical to the stage-1 discretization.
repeat
  Create an octree on the computational domain containing all source points in the stage-2 discretization and expansion centers.
  for all stage-2 elements $\Gamma_k^{\text{stage-2}}$ **do**
    for all source points $x_i^{\text{stage-2}} \in \Gamma_k^{\text{stage-2}}$ **do**
      Perform an area query of radius $0.5 \cdot (3/4)\eta_k^{\text{stage-2}}$ centered at $x_i^{\text{stage-2}}$.
      if the query returned an expansion center c such that $|c - x_i^{\text{stage-2}}|_2 \leq 0.5 \cdot (3/4)\eta_k^{\text{stage-2}}$ **then**
        Mark the element $\Gamma_k^{\text{stage-2}}$ for bisection.
      end if
    end for
  end for
  if elements were marked for bisection **then**

Bisect the marked elements.
end if
until no elements were marked for bisection

3.5 Associating Targets with QBX Centers

The computed set of quadrature discretizations ensures that the QBX expansions at the chosen set of centers can be computed accurately (ignoring error from acceleration). A final issue to be solved by geometry processing is to determine, for each target point, whether evaluation of the potential with QBX is needed or whether unmodified smooth high-order quadrature suffices. Algorithm 3 describes a procedure for associating targets to QBX centers. The algorithm proceeds in two stages, which consist of *identifying* endangered targets and *associating* targets to an expansion center using area queries. The algorithm produces a mapping from targets to associated centers, and it also flags endangered targets that could not be associated to any centers.

In the first stage, targets that require QBX evaluation are determined based on their proximity to endangering source particles. Section 3.4 implies that a 'danger zone' of radius $r_s^{\text{danger}} = \eta_k/2$ exists around each source particle $s \in \Gamma_k$. Using area queries around each source point of size r_s^{danger}, every target that some source endangers can be identified efficiently.

In the second stage, an area query around each expansion center having the same radius as the expansion ball is used to associate endangered targets to centers. In practice, because the expansion balls leave gaps in the coverage of the source danger zone, we have found it useful to allow a target to be matched to a center c if it is within a ball of radius $r_c(1 + \varepsilon_{\text{tgt}})$, where ε_{tgt} is some tolerance value and r_c is the expansion ball radius. Although such usage is not necessarily governed by theoretical guarantees, experience suggests that a small value of ε_{tgt} decreases the chance of having unassociated endangered targets without appearing to have an adverse impact on accuracy. If increasing ε_{tgt} still leads to unassociated targets, another available strategy is to refine near the targets which could not be associated, or to introduce additional QBX centers.

A natural extension of this algorithm is to include a *side preference* for each target. This amounts to associating a target only if the side of the center matches the desired side (interior/exterior) for the target. The need for this arises when performing on-surface evaluation for layer potentials where the limiting value depends on the direction of approach. See [37] for further discussion.

The algorithm we use in our implementation to obtain the results Section 6 is mildly more complicated than Algorithm 3, owing to being designed to run in parallel. The main difference is that in our implementation the area query takes the 'point of view' of the targets, rather than the sources or the centers. This potentially can result in better load distribution when large numbers of targets are clustered in one part of the geometry. Nevertheless, the output of our implemented algorithm is functionally identical to output of the algorithm in this section.

Algorithm 3: Associate near-source/'endangered' target points with QBX centers

Require: The geometry discretized as a set of targets, sources, and expansion centers.
Require: A target association tolerance $\varepsilon_{\text{tgt}} \geq 0$.
Ensure: Computes a partial function from targets to expansion centers and flags the set of targets that could not be associated.

Create an octree on the computational domain containing all sources, targets, and expansion centers.

Find endangered targets
for all source points $s \in \Gamma$ **do**
  Perform an area query of radius r_s^{danger} centered at s.
  for all targets t in boxes returned by the query **do**
    if $|t - s|_2 \leq r_s^{\text{danger}}$ **then**
      Mark t as endangered.
    end if
  end for
end for

Find centers for endangered targets
for all expansion centers c **do**
  Perform an area query of radius $r_c(1 + \varepsilon_{\text{tgt}})$ centered at c.
  for all endangered targets t in boxes returned by the query **do**
    if $|t - c|_2 \leq r_c(1 + \varepsilon_{\text{tgt}})$ **and** c is the closest center to t encountered so far **then**
      Associate t to c.
    end if
  end for
end for

Flag targets that could not be associated
for all endangered targets t **do**
  if t is not associated to a center **then**
    Flag t.
  end if
end for

4 Error Estimates for FMM Translations

In [57], error estimates were presented for the GIGAQBX FMM that applied to the 2D Laplace kernel with complex Taylor expansions. In this section, we present their analogs in three dimensions. We restrict our attention to the spherical harmonic approximation of the three-dimensional Laplace potential (2). This section lays the groundwork for showing the conditions under which a 3D GIGAQBX-style FMM for this potential can be expected to have the same convergence factor as a 3D point FMM.

4.1 Overview

The main difference between 'point' estimates for an FMM and the estimates in this section is that the object being approximated is a local expansion rather than a 'point' value (see Section 4.3). Nevertheless, readers familiar with the error estimates for the point FMM will recognize a number of similarities in this section with the error estimates from the point case.

Types of Translations. First, there is a direct correspondence in these error estimates with the evaluation scenarios for the point FMM. For instance, see [23, Lem. 3.2] for the local case (cf. Hypothesis 2), [23,

Thm. 3.5.4] for the multipole case (cf. Hypothesis 1), and [23, Thm. 3.5.5] for the multipole-to-local case (cf. Hypothesis 3).

'Sized' Targets. Second, the results in this section empirically confirm that, for purposes of FMM accuracy, the local expansion that is to be approximated behaves much like a 'sized target'. What this means is that, for a given evaluation scenario, the accuracy that is expected is similar to the accuracy expected if the local expansion were a set of targets for point evaluation.

Accuracy Dependence on Intermediate Expansion Order. Last, the results suggest that the accuracy chiefly depends on the order of the intermediate multipole/local expansions used and not the 'final' order of the local expansion (i.e. the QBX order). The results also suggest it might be possible to find an estimate independent of the final expansion order. With regards to the dependence on the order of the intermediate expansion, the error behavior mirrors the original FMM.

4.2 Analytical Preliminaries

Local expansions have already been introduced in Section 2.1. Recall that the p-th order local expansion L_p due to a source $s \in \mathbb{R}^3$ centered at $c \in \mathbb{R}^3$, with coefficients $\langle L_n^m \rangle$ given by (5) and evaluated at a target $t \in \mathbb{R}^3$ takes the form

$$L_p(t) = \sum_{n=0}^{p} \sum_{m=-n}^{n} L_n^m |t-c|_2^n Y_n^m(\theta_{t-c}, \phi_{t-c}).$$

This expansion converges as long as $|t-c|_2 < |s-c|_2$.

A *multipole expansion* due to a source $s \in \mathbb{R}^3$ with center $c \in \mathbb{R}^3$ is defined via the coefficients $\langle M_n^m \rangle$ given by

$$M_n^m := \frac{1}{2n+1} |s-c|_2^n Y_n^m(\theta_{s-c}, \phi_{s-c}). \tag{22}$$

The expression $M_p(t)$ for a p-th order multipole expansion evaluated at $t \in \mathbb{R}^3$ takes the form

$$M_p(t) = \sum_{n=0}^{p} \sum_{m=-n}^{n} \frac{M_n^m}{|t-c|_2^{n+1}} Y_n^{-m}(\theta_{t-c}, \phi_{t-c}).$$

The multipole expansion converges for $|t-c|_2 > |s-c|_2$.

Translation operators allow for the shifting of centers of expansions, or the conversion of multipole expansions to local expansions. If the coefficients of the original expansion are notated as $\langle B_n^m \rangle$, the translation operator amounts to a linear transformation $\langle B_n^m \rangle \mapsto \langle (B')_n^m \rangle$ to new coefficients $\langle (B')_n^m \rangle$. We will say that the expansion with coefficients $\langle B_n^m \rangle$ has order p if the coefficients $\langle B_n^m \rangle$ satisfy $B_n^m = 0$ for $n > p$. Equivalently, in that case the $\langle B_n^m \rangle$ may be thought of as a vector of $(p+1)^2$ coefficients $B_0^0, B_1^{-1}, B_1^0, B_1^1, \ldots$.

We denote the translation of a local expansion with order p from a center $c \in \mathbb{R}^3$, to a local expansion with order q at a center $c' \in \mathbb{R}^3$, by $\langle (B')_n^m \rangle = \mathrm{L2L}_{p\to q}^{c\to c'} \langle B_n^m \rangle$. See [23, Lem. 3.2] for the explicit formula for this operator.

We similarly define the operators $\mathrm{M2M}_{p\to q}^{c\to c'}$ for multipole-to-multipole translation and $\mathrm{M2L}_{p\to q}^{c\to c'}$ for multipole-to-local translation. An explicit formula for both of these operators may be found respectively in [23, Thm. 3.5.4] and [23, Thm. 3.5.5].

The following error estimate pertains to the truncation error of multipole and local expansions.

Proposition 3 (Accuracy of multipole and local expansions, based on [23, Lem. 3.2.4])**.** *Let $s, c, t \in \mathbb{R}^3$.*

(a) Consider the local expansion of $\mathcal{G}(s,\cdot)$ centered at c and evaluated at t, with the coefficients $\langle L_n^m \rangle$ defined as in (5). Let $r = |t-c|_2$ and $\rho = |s-c|_2$. If $r < \rho$, then

$$\left| \mathcal{G}(s,t) - \sum_{n=0}^{p} \sum_{m=-n}^{n} L_n^m |t-c|_2^n Y_n^m(\theta_{t-c}, \phi_{t-c}) \right| \leq \frac{1}{4\pi} \frac{1}{\rho - r} \left(\frac{r}{\rho} \right)^{p+1}.$$

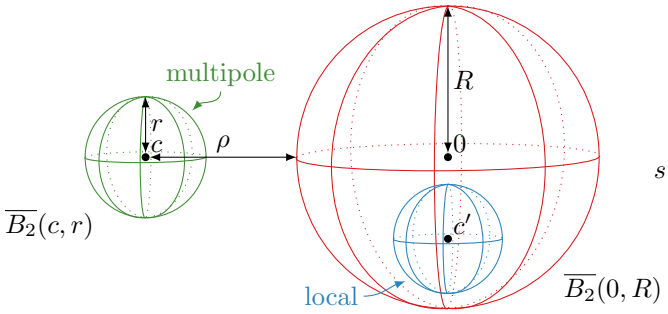


Figure 6: Geometric depiction of the use of a multipole expansion to approximate the local expansion of a potential. The multipole expansion is formed at the center c and translated to c'. This provides the geometric setting for the situations described in Sections 4.3.1 and 4.3.3.

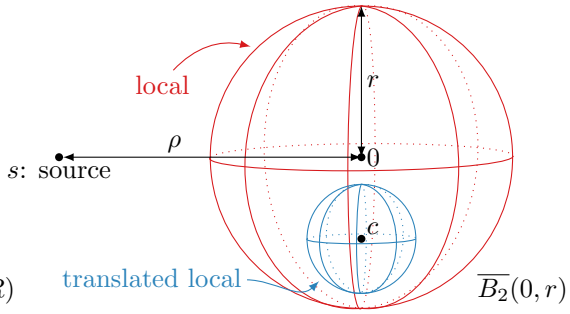


Figure 7: Geometric depiction of the use of an intermediate local expansion to approximate the local expansion of a potential. The local expansion is formed at the origin and translated to c. This provides the geometric setting for the situation described in Section 4.3.2.

(b) Next, consider the multipole expansion of $\mathcal{G}(s,\cdot)$ centered at c and evaluated at t, with the coefficients $\langle M_n^m \rangle$ defined as in (22). Let $r = |s-c|_2$ and $\rho = |t-c|_2$. If $r < \rho$, then

$$\left| \mathcal{G}(s,t) - \sum_{n=0}^{p} \sum_{m=-n}^{n} \frac{M_n^m}{|t-c|_2^{n+1}} Y_n^{-m}(\theta_{t-c}, \phi_{t-c}) \right| \leq \frac{1}{4\pi} \frac{1}{\rho - r} \left(\frac{r}{\rho} \right)^{p+1}.$$

4.3 Accuracy of GIGAQBX FMM Translations

Recall the difference between the notion of accuracy in a point FMM and the notion of accuracy used in the GIGAQBX FMM. A point FMM computes the value of a potential at a point x. The accuracy is measured by

$$\text{Point FMM Accuracy} = \left| \Phi(x) - \tilde{\Phi}_p(x) \right|, \tag{23}$$

where $\tilde{\Phi}_p$ is the point FMM's p-th order approximation to the point potential Φ (14).

In contrast, the correct error metric to use in the GIGAQBX FMM measures the FMM's ability to approximate the q-th order local expansion of the potential at a generic target x from a generic center c. Let $\langle L_n^m \rangle$ denote the coefficients of the local expansion of Φ centered at c, and let $\langle (\tilde{L}_p)_n^m \rangle$ denote the coefficients of the local expansion of the p-th order approximation $\tilde{\Phi}_p$ centered at c. Then the accuracy may be measured by

$$\text{GIGAQBX FMM Accuracy} = \left| \sum_{n=0}^{q} \sum_{m=-n}^{n} (\tilde{L}_p)_n^m |x-c|_2^n Y_n^m(\theta_{x-c}, \phi_{x-c}) - \sum_{n=0}^{q} \sum_{m=-n}^{n} L_n^m |x-c|_2^n Y_n^m(\theta_{x-c}, \phi_{x-c}) \right|. \tag{24}$$

The formulas (24) and (23) are related in that the point FMM error (23) is a special case of (24) with an expansion radius of zero. The error estimates in this section will be of the form (24).

Given the intended usage pattern in the FMM, and using an intermediate expansion order $p \in \mathbb{N}_0$ and target (QBX) expansion order $q \in \mathbb{N}_0$, we present a study of the error for the following translation chains:

- Source $\to$ Multipole(p) $\to$ Local(q) (Section 4.3.1)
- Source $\to$ Local(p) $\to$ Local(q) (Section 4.3.2)
- Source $\to$ Multipole(p) $\to$ Local(p) $\to$ Local(q) (Section 4.3.3).

In this analysis, it suffices to consider at most one intermediate local or multipole expansion of order p. This is because the potential that results via a sequence of p-th order local-to-local translations only depends on

the source and the initial expansion center. Similarly, the potential that results via a sequence of p-th order multipole-to-multipole translations only depends on the source and the final expansion center.

We refer the reader to Appendix B for the details of the numerical experiments used to obtain the results in this section.

Remark 4 (Error estimates for multiple sources)**.** *The following sections work with a single unit-strength source charge, but can be straightforwardly extended to the case of an ensemble of m charges $s_1, s_2, \ldots, s_m$ with strengths $q_1, q_s, \ldots, q_m$. The corresponding error bound is scaled by $\sum_{i=1}^{m} |q_i|$.*

4.3.1 Multipole Accuracy

Recall that a multipole expansion is an 'outgoing' approximation to the field due to a set of sources at any point sufficiently far away from the expansion center. In this section, we consider the ability of a local expansion obtained through translation from a multipole expansion to approximate the local expansion of a potential. We make use of the following geometric situation, illustrated in Figure 6. Let $R > 0$ and $\rho > r > 0$. Consider a closed ball of radius r centered at c, with $|c|_2 = R + \rho$, containing a unit-strength source s. Also let a ball of radius R centered at the origin contain points t, c' satisfying $|c'|_2 \leq R$ and $|c' - t|_2 \leq R - |c'|_2$.

Suppose a p-th order multipole expansion M_p with coefficients $\langle (T_p^c)_n^m \rangle$ is formed at c due to the source s. Next suppose that this is translated to a q-th order local expansion with coefficients $\langle (T_q^{c'})_n^m \rangle = \mathrm{M2L}_{p\to q}^{c\to c'} \langle (T_p^c)_n^m \rangle$. The local expansion of the potential $\mathcal{G}(s, \cdot)$ due to s centered at c' may be written using the local coefficients $\langle (L_q^{c'})_n^m \rangle$. If one were to use the coefficients $\langle (T_q^{c'})_n^m \rangle$ as approximations to the coefficients $\langle (L_q^{c'})_n^m \rangle$ for evaluation of the local expansion at the target t, the approximation error would be the quantity $E_M(q)$ defined as

$$E_M(q) = \left| \sum_{n=0}^{q} \sum_{m=-n}^{n} (T_q^{c'})_n^m |t - c'|_2^n Y_n^m(\theta_{t-c'}, \phi_{t-c'}) - \sum_{n=0}^{q} \sum_{m=-n}^{n} (L_q^{c'})_n^m |t - c'|_2^n Y_n^m(\theta_{t-c'}, \phi_{t-c'}) \right|. \tag{25}$$

The following observation suggests a bound on E_M. Define the error function

$$\mathcal{E}(x) = \mathcal{G}(s, x) - M_p(x).$$

Observe that the quantity (25) is the magnitude of the q-th order local expansion of $\mathcal{E}$ centered at c' and evaluated at t. It can be shown that this expansion must converge, in the sense that $\lim_{q\to\infty} E_M(q) = \mathcal{E}(t)$. Therefore, by Proposition 3, we may expect that for every $\varepsilon > 0$, it is the case that for all sufficiently large q that

$$E_M(q) \leq (1 + \varepsilon) \left((4\pi)^{-1} / (\rho - r) \right) (r/\rho)^{p+1}.$$

It is also easy to see that $E_M(0) \leq ((4\pi)^{-1}/(\rho - r))(r/\rho)^{p+1}$, as the 0-th order expansion is just the quantity $\mathcal{E}(0)$, which satisfies this bound by Proposition 3. Given this asymptotic behavior and the fact that it holds for $q = 0$, it is at least plausible that the bound $E_M(q) \leq C((4\pi)^{-1}/(\rho - r))(r/\rho)^{p+1}$, for some $C > 0$, should hold for all q. We formulate this statement as the following hypothesis.

Hypothesis 1 (Source $\to$ Multipole(p) $\to$ Local(q))**.** *For the situation described above, there exists a constant $C > 0$ independent of R, p, q, ρ, r, s, c', and t such that the error in the multipole-mediated approximation to the local expansion of the potential satisfies the bound*

$$\left| \sum_{n=0}^{q} \sum_{m=-n}^{n} (T_q^{c'})_n^m |t - c'|_2^n Y_n^m(\theta_{t-c'}, \phi_{t-c'}) - \sum_{n=0}^{q} \sum_{m=-n}^{n} (L_q^{c'})_n^m |t - c'|_2^n Y_n^m(\theta_{t-c'}, \phi_{t-c'}) \right| \leq \frac{1}{4\pi} \frac{C}{\rho - r} \left(\frac{r}{\rho} \right)^{p+1}.$$

This hypothesis would imply that multipole translation in the GIGAQBX FMM obeys essentially the same bound as the corresponding point FMM error. This bound is true if and only if the approximations to $\mathcal{E}$ never significantly 'overshoot' the true value of $\mathcal{E}(t)$. The mathematical details of the proof of this hypothesis turn out to be fairly involved and its complete mathematical resolution is pending. However the numerical evidence is strongly in favor of this bound. We found, as described in Appendix B, that the inequality in Hypothesis 1 holds numerically with $C = 1.002$. Although the tests we performed cannot be exhaustive, the small value of C obtained by numerical testing is consistent with the truth of the hypothesis. We use Hypothesis 1 as an error estimate in the remainder of this paper and we expect to present a proof of this hypothesis at a later date.

4.3.2 Local Accuracy

Recall that the role of the local expansion is complementary to the multipole expansion, since the local expansion represents the potential in a neighborhood of an expansion center. This section considers the case where a local expansion is formed due to a potential at one center, subsequently translated to a second center, and used to approximate the local expansion of the potential at the second center. We make use of the following geometric situation, illustrated in Figure 7. Let $\rho > r > 0$ and suppose that a source particle is placed at s, with $|s|_2 = \rho$. Let $t, c \in \overline{B_2}(0, r)$ with $|t - c|_2 \leq r - |c|_2$.

The potential due to the source s can be described in a q-th order local expansion centered at c with coefficients $\langle (L_q^c)_n^m \rangle$. Consider a p-th order local expansion of the potential centered at the origin with coefficients $\langle (T_p^0)_n^m \rangle$. Suppose this expansion is translated to a q-th order expansion at c with coefficients $\langle (T_q^c)_n^m \rangle$ given by $\langle (T_q^c)_n^m \rangle = \mathrm{L2L}_{p\to q}^{0\to c} \langle (T_p^0)_n^m \rangle$. If the coefficients $\langle (T_q^c)_n^m \rangle$ are used in place of $\langle (L_q^c)_n^m \rangle$ as an approximation to the q-th order expansion for evaluation at a target t, the approximation error is the quantity $E_L(q)$ given by

$$E_L(q) = \left| \sum_{n=0}^{q} \sum_{m=-n}^{n} (T_q^c)_n^m |t - c|_2^n Y_n^m(\theta_{t-c}, \phi_{t-c}) - \sum_{n=0}^{q} \sum_{m=-n}^{n} (L_q^c)_n^m |t - c|_2^n Y_n^m(\theta_{t-c}, \phi_{t-c}) \right|. \tag{26}$$

The asymptotic behavior of $E_L(q)$ is similar to that of $E_M(q)$ in the previous section. In other words, for fixed p, we have $E_L(0) \leq ((4\pi)^{-1}/(\rho - r))(r/\rho)^{p+1}$ and, for all $\varepsilon > 0$ it is the case for sufficiently large q that $E_L(q) \leq (1 + \varepsilon)((4\pi)^{-1}/(\rho - r))(r/\rho)^{p+1}$. This motivates the following hypothesis about the behavior of $E_L(q)$ for all q.

Hypothesis 2 (Source $\to$ Local(p) $\to$ Local(q))**.** *For the situation described above, there exists a constant $C > 0$ independent of p, q, ρ, r, s, c, and t such that the error in the local-mediated approximation to the local expansion of the potential satisfies the bound*

$$\left| \sum_{n=0}^{q} \sum_{m=-n}^{n} (T_q^c)_n^m |t - c|_2^n Y_n^m(\theta_{t-c}, \phi_{t-c}) - \sum_{n=0}^{q} \sum_{m=-n}^{n} (L_q^c)_n^m |t - c|_2^n Y_n^m(\theta_{t-c}, \phi_{t-c}) \right| \leq \frac{1}{4\pi} \frac{C}{\rho - r} \left(\frac{r}{\rho} \right)^{p+1}.$$

Similar to the previous section, we evaluated the truth of this hypothesis numerically. We observed that the estimate for $E_L(q)$ empirically satisfies Hypothesis 2 with $C = 1.001$. As a result, we make use of Hypothesis 2 as an error estimate in this paper.

4.3.3 Multipole-to-Local Accuracy

In this section, we work with the same geometrical situation as given in Section 4.3.1 and illustrated in Figure 6. We consider the accuracy achieved when a multipole expansion is converted into a local expansion which is then translated to a third center, and used to approximate the local expansion of the potential at that center. Similar to Section 4.3.1, consider a p-th order multipole expansion due to the source s formed at the center c with coefficients $\langle (T_p^c)_n^m \rangle$. Suppose this expansion is translated to a p-th order local expansion T_p^0 centered at the origin, with coefficients $\langle (T_p^0)_n^m \rangle = \mathrm{M2L}_{p\to p}^{c\to 0} \langle (T_p^c)_n^m \rangle$. Then suppose that this expansion is translated to a q-th order local expansion centered at another center c', with the coefficients $\langle (T_q^{c'})_n^m \rangle$ given by $\langle (T_q^{c'})_n^m \rangle = \mathrm{L2L}_{p\to q}^{0\to c'} \langle (T_p^0)_n^m \rangle$.

The coefficients $\langle (T_q^{c'})_n^m \rangle$ are those of a local expansion that approximates the q-th order local expansion of the potential $\mathcal{G}(s, \cdot)$, centered at c'. Let $\langle (L_q^{c'})_n^m \rangle$ be the local coefficients of the approximated expansion. The error is given by

$$E_{M2L}(q) = \left| \sum_{n=0}^{q} \sum_{m=-n}^{n} (T_q^{c'})_n^m |t - c'|_2^n Y_n^m(\theta_{t-c'}, \phi_{t-c'}) - \sum_{n=0}^{q} \sum_{m=-n}^{n} (L_q^{c'})_n^m |t - c'|_2^n Y_n^m(\theta_{t-c'}, \phi_{t-c'}) \right|. \tag{27}$$

In the limit as $q \to \infty$ we have

$$\lim_{q\to\infty} E_{M2L}(q) = |\mathcal{G}(s, t) - T_p^0(t)|.$$

To simplify the analysis of this quantity, we introduce the p-th order local expansion of $\mathcal{G}(s,\cdot)$ at the origin, denoted by L_p^0. By the triangle inequality,

$$|\mathcal{G}(s,t) - T_p^0(t)| \leq |\mathcal{G}(s,t) - L_p^0(t)| + |L_p^0(t) - T_p^0(t)|.$$

The quantity $|\mathcal{G}(s,t) - L_p^0(t)|$ may be bounded by Proposition 3. Since the quantity $|L_p^0(t) - T_p^0(t)|$ is the difference between the local expansion of a multipole and the local expansion of the point potential at the origin, we may bound it with Hypothesis 1. This yields the bound

$$|\mathcal{G}(s,t) - T_p^0(t)| \leq \frac{1}{4\pi}\left[\frac{1}{\rho - r}\left(\frac{R}{R + (\rho - r)}\right)^{p+1} + \frac{C}{\rho - r}\left(\frac{r}{\rho}\right)^{p+1}\right].$$

Similar to Sections 4.3.1 and 4.3.2, this quantity is an asymptotic bound on E_{M2L} in q. This leads to the following hypothesis.

Hypothesis 3 (Source $\to$ Multipole(p) $\to$ Local(p) $\to$ Local(q))**.** *For the situation described above, there exists a constant $C > 0$ independent of R, p, q, ρ, r, s, c, c', and t such that the error in the multipole and local mediated approximation to the local expansion of the potential satisfies the bound*

$$\left|\sum_{n=0}^{q}\sum_{m=-n}^{n}(T_q^{c'})_n^m |t - c'|_2^n Y_n^m(\theta_{t-c'}, \phi_{t-c'}) - \sum_{n=0}^{q}\sum_{m=-n}^{n}(L_q^{c'})_n^m |t - c'|_2^n Y_n^m(\theta_{t-c'}, \phi_{t-c'})\right|$$
$$\leq \frac{C}{4\pi}\left[\frac{1}{\rho - r}\left(\frac{R}{R + (\rho - r)}\right)^{p+1} + \frac{1}{\rho - r}\left(\frac{r}{\rho}\right)^{p+1}\right].$$

Numerically, we observed that Hypothesis 3 appears to hold with $C = 1.001$. Similar to Sections 4.3.1 and 4.3.2, the low value of C that was observed is consistent with the truth of the hypothesis. We therefore make use of Hypothesis 3 as an error estimate in the remainder of this paper.

5 The GIGAQBX Algorithm in Three Dimensions

This section is concerned with the precise statement of the GIGAQBX algorithm and its complexity and accuracy analysis. The algorithm is presented in Sections 5.2 and 5.3. The accuracy and complexity analyses follow in Sections 5.4 and 5.5.

5.1 Overview

For the benefit of readers familiar with the point FMM, the original QBX FMM [48], or the two-dimensional version of GIGAQBX presented in [57], we point out the main differences in this section.

Target Confinement Rule. To prevent inaccurate contributions from entering the QBX local expansion, while still maintaining the efficiency enabled by the use of a tree, the design of the GIGAQBX algorithm adopts the point of view of QBX centers as 'targets with extent' that each have their own near-field. The realization of this idea is that GIGAQBX only permits QBX expansion balls to exist in a box if they do not extend beyond an (ℓ^2) radius surrounding the box, where the length of the radius is proportional to the box size, so that the near-field of the box ends up being an over-approximation of the near-field of the target with extent.

We call this modification the target confinement rule. During box subdivision in tree construction, if a ball cannot be placed in the child box due to this restriction, it remains in the parent box.

Particles Owned by Non-Leaf Boxes. It follows from the previous paragraph that the GIGAQBX algorithm, unlike the point FMM, allows for particles (specifically, QBX centers) to be 'owned' by non-leaf 'ancestor' boxes. The most important implication of this design is that interaction lists involving direct evaluations at particles (List 1 and List 3), as well as the FMM step of evaluation of far-field local expansions, must be redefined to incorporate the possibility of evaluation at non-leaf boxes.

Two-Away Near Neighborhood. To obtain a good convergence factor in two or three dimensions, it is convenient to consider the 'near-field' to consist of both a box's nearest neighbors and also its second nearest neighbors. This is not a new modification, having been present in the original three-dimensional FMM [23].

'Close' and 'Far' Lists. In order to actually obtain the accuracy guarantees provided by the target confinement rule, we disallow certain box near-field interactions that are too close to the target confinement region from using expansion mediation that would otherwise be mediated by expansions in the point FMM. Specifically, the fields associated with List 3 and List 4 are subdivided into 'close' and 'far' lists, where the close lists are evaluated directly via point-to-QBX-local interactions, and the far lists maintain sufficient separation to allow for normal expansion mediation.

Changes from the 2D Version of GIGAQBX. Perhaps the most significant difference with the two-dimensional version [57] is the use of an ℓ^2 target confinement region to confine the QBX expansion centers, whereas the previous version used an ℓ^∞ (square) region. The use of the ℓ^2 region improves the efficiency of the scheme, especially in three dimensions. Other than this, the definitions of the interaction lists and the statement of the algorithm itself are largely unchanged from the two-dimensional case.

5.2 Definitions and Interaction Lists

5.2.1 Computational Domain

The computational domain for the algorithm is an octree whose axis-aligned root box contains all sources, targets, and expansion centers ('particles') as well as the entirety of each expansion disk. Starting with the root box, the octree is refined by repeated subdivision of boxes that contain more than $n_{\max}$ particles until no more non-empty boxes can be produced, pruning any empty childless boxes. A childless box is also called a *leaf box*. Each box, including non-leaves, conceptually 'owns' a subset of particles. Upon subdivision, non-expansion center particles are placed from the parent into the spatially appropriate child box. Expansion centers are only placed in the child box if the expansion ball can fit within the target confinement region of the child. Otherwise, they remain owned by the parent.

5.2.2 Notation

For a box b in the octree, we will use $|b|$ to denote the (ℓ^∞) radius of b. The *target confinement region* ('TCR', also $\mathsf{TCR}(b)$) of a box b with center c is $\overline{B_2}(\sqrt{3}|b|(1+t_f), c)$, where t_f is the *target confinement factor* ('TCF'), where $\sqrt{3}|b|$ is half the box diagonal.

The *k-near neighborhood* of a box b with center c is the region $\overline{B_\infty}(|b|(1+2k), c)$.

The *k-colleagues* of a box b are boxes of the same level as b that are contained inside the k-near neighborhood of b. T_b denotes the set of 2-colleagues of a box b.

We say that two boxes are *k-well-separated* if they are on the same level and are not k-colleagues.

The parent of b is denoted $\mathsf{Parent}(b)$. The set of ancestors is $\mathsf{Ancestors}(b)$. The set of descendants is $\mathsf{Descendants}(b)$. $\mathsf{Ancestors}$ and $\mathsf{Descendants}$ are also defined in the natural way for sets of boxes.

A box owning a point or QBX center target is called a *target box*. A box owning a source quadrature node is called a *source box*. Ancestors of target boxes are called *target-ancestor boxes*.

Definition 1 (Adequate separation relation, $\prec$)**.** *We define a relation $\prec$ over the set of boxes and target confinement regions within the tree, with $a \prec b$ to be understood as 'a is adequately separated from b, relative to the size of a'.*

We write $a \prec \mathsf{TCR}(b)$ for boxes a and b if the ℓ^2 distance from the center of a to the boundary of $\mathsf{TCR}(b)$ is at least $3|a|$.

We write $\mathsf{TCR}(a) \prec b$ for boxes a and b if the ℓ^∞ distance from the center of a to the boundary of b is at least $3|a|(1+t_f)$. (This implies that the ℓ^2 distance is at least $3|a|(1+t_f)$.)

We write $a \not\prec b$ to denote the negation of $a \prec b$.

Because the size of the TCR is proportional to the box size, $\mathsf{Parent}(a) \prec b$ implies $a \prec b$. We call this property the 'monotonicity' of '$\prec$'.

5.2.3 Conventional Interaction Lists

The four conventional interaction lists in the FMM are defined in this section, with two modifications to the standard definition. First, non-leaf boxes are allowed as target boxes. Thus, lists normally associated with only leaf boxes (Lists 1 and 3) may be associated with arbitrary boxes in the tree. Second, our definition makes use of a near neighborhood of a box that is two boxes wide.

List 1 consists of interactions with adjacent boxes. These interactions are carried out directly, without acceleration through any expansions.

Definition 2 (List 1, U_b)**.** *For a target box b, U_b consists of all leaf boxes from among* $\mathsf{Descendants}(b) \cup \{b\}$ *and the set of boxes adjacent to b.*

List 2 consists of interactions with same-level boxes. These interactions have sufficient separation for an accurate multipole-to-local mediation. List 2 is *downward-propagating*, which means that the interactions received by the box are translated downward to the box's descendants via local-to-local translation.

Definition 3 (List 2, V_b)**.** *For a target or target-ancestor box b, V_b consists of the children of the 2-colleagues of b's parent that are 2-well-separated from b.*

List 3 consists of interactions where the source box is in the near field (T_b) of the target box, but not adjacent to the target box. Unlike List 2, the separation is insufficient for accurate multipole-to-local mediation. List 3 is not downward-propagating, and it is usually mediated with a multipole-to-target interaction.

Definition 4 (List 3, W_b)**.** *For a target box b, a box $d \in \mathsf{Descendants}(T_b)$ is in W_b if d is not adjacent to b and, for all $w \in \mathsf{Ancestors}(d) \cap \mathsf{Descendants}(T_b)$, w is adjacent to b.*

The following are immediate consequences of this definition:

- Any box in W_b is strictly smaller than b.
- Any box $d \in W_b$ is separated from b by at least the width of d.
- List 3 of b contains the immediate children of non-adjacent 2-colleagues of b.

List 4 consists of interactions where the target box is in the near field of the source box, but not adjacent to it. Like List 3, the separation is insufficient for accurate multipole-to-local interaction. Unlike List 3, List 4 *is* downward-propagating, and one may form a local expansion of the field from the source box that can be propagated to the descendants.

Definition 5 (List 4, X_b)**.** *For a target or target-ancestor box b, a source box d is in List 4 of b if d is a 2-colleague of some ancestor of b and d is adjacent to* $\mathsf{Parent}(b)$ *but not b itself. Additionally, a source box d is in X_b if d is a 2-colleague of b and d is not adjacent to b.*

The following are immediate consequences of this definition:

- Any box in X_b is at least as large as b.
- Any box in X_b is separated from b by at least the width of b.
- For any $d \in X_b$, either $b \in W_d$ or d is a 2-colleague of b.

Our FMM does not make use of List 3 and List 4 directly. Instead, the field from these lists is sub-partitioned into 'close' and 'far' lists. This is explained in the next section.

5.2.4 Close and Far Lists

Because of inadequate separation from the target confinement region, our algorithm cannot make use of the interaction lists W_b and X_b as they would be used normally: via, respectively, a multipole-to-target or source-to-local interaction. To manage this issue while still maintaining the benefits of FMM acceleration, we partition the contribution of the sources contained in these lists into 'close' and 'far' lists. The field due to a 'close' list cannot be mediated through intermediate expansions for accuracy reasons, and so it is

evaluated directly (i.e., using point-to-QBX-local expansion) at the QBX expansion centers. In contrast, the field due to a 'far' list is sufficiently far from the target confinement region that mediation via an intermediate expansion is permissible from the standpoint of accuracy.

For case of List 3, the 'close' list consists of source boxes in the near field of the target box which are not adequately separated from the TCR of the target box. Interactions from these source boxes must be accumulated directly. The 'far' list is the smallest possible 'complement' of this list in the sense that the close and far lists must cover the entire near field mediated by List 3, and the 'far' lists contains boxes that are as large as permissible given the target confinement restrictions. The 'far' list may be mediated via multipole-to-target interaction.

Definition 6 (List 3 close, W_b^{close})**.** *For a target box b, a leaf box d is said to be in W_b^{close} if $d \in \mathsf{Descendants}(W_b) \cup W_b$ such that $d \not\prec \mathsf{TCR}(b)$.*

Definition 7 (List 3 far, W_b^{far})**.** *For a target box b, a box d is in W_b^{far} if $d \in \mathsf{Descendants}(W_b) \cup W_b$ such that $d \prec \mathsf{TCR}(b)$ and, for all $w \in \mathsf{Ancestors}(d) \cap (\mathsf{Descendants}(W_b) \cup W_b)$, $w \not\prec \mathsf{TCR}(b)$.*

While $W_b^{\text{close}} \cup W_b^{\text{far}} \subseteq W_b$ does not hold in general, it is generally the case that $W_b^{\text{close}} \cup W_b^{\text{far}} \subseteq \mathsf{Descendants}(W_b) \cup W_b$.

For the case of List 4, the 'close' list consists of source boxes from which the TCR of the target box is not adequately separated. List 4 close is evaluated directly *only at the targets in the box*, and is not downward-propagating. List 4 far is downward-propagating. It consists of boxes form which the TCR of the target box is adequately separated. By monotonicity of '$\prec$', this means that the TCR of the descendants is also adequately separated from the sources, ensuring accurate downward propagation. To ensure that the field of List 4 close is propagated to the descendants, a box in List 4 close of the parent is placed in List 4 far of a descendant as soon as the TCR of the descendant is adequately separated from it.

Definition 8 (List 4 close, X_b^{close})**.** *Let b be a target or target-ancestor box. A box d is in X_b^{close} if for some $w \in \mathsf{Ancestors}(b) \cup \{b\}$ we have $d \in X_w$ and furthermore $\mathsf{TCR}(b) \not\prec d$.*

Definition 9 (List 4 far, X_b^{far})**.** *Let b be a target or target-ancestor box. A box $d \in X_b$ is in List 4 far if $\mathsf{TCR}(b) \prec d$. Furthermore, if b has a parent, a box $d \in X_{\mathsf{Parent}(b)}^{\text{close}}$ is in List 4 far if $\mathsf{TCR}(b) \prec d$.*

As with List 3, it is generally not the case that $X_b^{\text{close}} \cup X_b^{\text{far}} \subseteq X_b$. However, $X_b^{\text{close}} \cup X_b^{\text{far}}$ only contains boxes from a List 4 of b or an ancestor of b.

Remark 5 (Performance optimization for List 3 far)**.** *One can always remove a box from W_b^{far} and place its leaf descendants in W_b^{close}, without adverse impact on accuracy since a multipole-to-QBX-local interaction is replaced with a direct one. When the number of sources in the descendants is small, doing this decreases the computational cost associated with evaluation of the potential due to the sources. We make use of this possibility in the complexity analysis.*

5.3 Formal Statement of Algorithm

5.3.1 Notation

We make use of the following notation, which is a slightly modified version of the notation used in [57]. The following notation refers to 'point' potentials evaluated at a target not requiring QBX owned by a box b: (a) $\mathsf{P}_b^{\text{near}}(t)$ denotes the potential at a target point t due to all sources in $U_b \cup W_b^{\text{close}} \cup X_b^{\text{close}}$; and (b) $\mathsf{P}_b^{W}(t)$ denotes the potential at a target t due to all sources in W_b^{far}. Let c be a QBX center owned by box b. The following notation refers to potentials evaluated with QBX mediation: (a) $\mathsf{L}_c^{\text{qbx,near}}(t)$ denotes the (QBX) local expansion of the potential at the center c, evaluated at target t, due to all sources in $U_b \cup W_b^{\text{close}} \cup X_b^{\text{close}}$; (b) $\mathsf{L}_c^{\text{qbx},W}(t)$ denotes the (QBX) local expansion at the center c, evaluated at t, due to all sources in W_b^{far}; and (c) $\mathsf{L}_c^{\text{qbx,far}}(t)$ denotes the (QBX) local expansion at the center c, evaluated at t, due to all sources not in $U_b \cup W_b \cup X_b^{\text{close}}$. Lastly, given a box b, M_b and $\mathsf{L}_b^{\text{far}}$ refer respectively to the multipole and local expansions associated with the box.

5.3.2 Algorithmic Parameters

The parameters to the algorithm are p_{fmm}, the FMM order; p_{qbx}, the QBX order; p_{quad}, the upsampled quadrature node count; and the target confinement factor t_f. The choice of these parameters is based on the splitting of the overall error in the scheme into truncation error (Lemma 1), quadrature error (Lemma 2), and acceleration error (Theorem 6).

Of these sources of error, perhaps the one that is most straightforward to control with algorithmic parameters is the acceleration error. Given a tolerance $\varepsilon > 0$, choosing $t_f \leq 0.85$ allows setting $p_{\text{fmm}} \approx |\log_{4/3} \varepsilon|$ to guarantee that the relative error for this component is on the order of ε. (See Theorem 6 below.)

In contrast, the topic of parameter selection for QBX and quadrature order remains an area of active research. On the subject of quadrature order, we have found the advice in [57], which suggests the choice of a generically high order to ensure the smallness of the quadrature error term, to be a useful guide. Assuming the quadrature error is suitably controlled, we have observed that p_{fmm} is a practical upper bound on p_{qbx}, as the error due to acceleration empirically appears to decrease most slowly of the various sources of error (see Section 6).

The complete statement of the algorithm is given in Algorithm 4.

Algorithm 4: GIGAQBX FMM in Three Dimensions

Require: The maximum number of FMM targets/sources $n_{\max}$ per box for octree refinement and a target confinement factor t_f are chosen.
Require: The input geometry and targets are preprocessed according to Section 3.
Require: Based on the precision ε to be achieved, a QBX order p_{qbx}, an FMM order p_{fmm}, and an oversampled quadrature node count p_{quad} are chosen (see Sections 2 and 3).
Ensure: An accurate approximation to the potential at all target points is computed.

Stage 1: Build tree
Create a octree on the computational domain containing all sources, targets, and QBX centers.
repeat
  Subdivide each box containing more than $n_{\max}$ particles into eight children, pruning any empty child boxes. If an expansion center cannot be placed in a child box with target confinement factor t_f due to its radius, it remains in the parent box.
until each box can no longer be subdivided or an iteration produced only empty child boxes

Stage 2: Form multipoles
for all boxes b **do**
  Form a p_{fmm}-th order multipole expansion M_b centered at b due to sources owned by b.
end for
for all boxes b in postorder **do**
  For each child of b, shift the center of the multipole expansion at the child to b. Add the resulting expansions to M_b.
end for

Stage 3: Evaluate direct interactions
for all boxes b **do**
  For each conventional target t owned by b, add to $\mathsf{P}_b^{\text{near}}(t)$ the contribution due to the interactions from sources owned by boxes in U_b to t.
end for
for all boxes b **do**
  For each QBX center c owned by b, add to the expansion $\mathsf{L}_c^{\text{qbx,near}}$ the contribution due to the interactions from U_b to c.
end for

Stage 4: Translate multipoles to local expansions
for all boxes b **do**
For each box $d \in V_b$, translate the multipole expansion M_d to a local expansion centered at b. Add the resulting expansions to obtain $\mathsf{L}_b^{\text{far}}$.
end for

Stage 5(a): Evaluate direct interactions due to W_b^{close}
Repeat Stage 3 with W_b^{close} instead of U_b.

Stage 5(b): Evaluate multipoles due to W_b^{far}
for all boxes b **do**
For each conventional target t owned by b, evaluate the multipole expansion M_d of each box $d \in W_b^{\text{far}}$ to obtain $\mathsf{P}_b^W(t)$.
end for
for all boxes b **do**
For each QBX center c owned by b, add to the expansion $\mathsf{L}_c^{\text{qbx},W}$ the contributions due to the multipole expansion M_d of each box $d \in W_b^{\text{far}}$.
end for

Stage 6(a): Evaluate direct interactions due to X_b^{close}
Repeat Stage 3 with X_b^{close} instead of U_b.

Stage 6(b): Form locals due to X_b^{far}
for all boxes b **do**
Convert the field of every particle owned by boxes in X_b^{far} to a local expansion about b. Add to $\mathsf{L}_b^{\text{far}}$.
end for

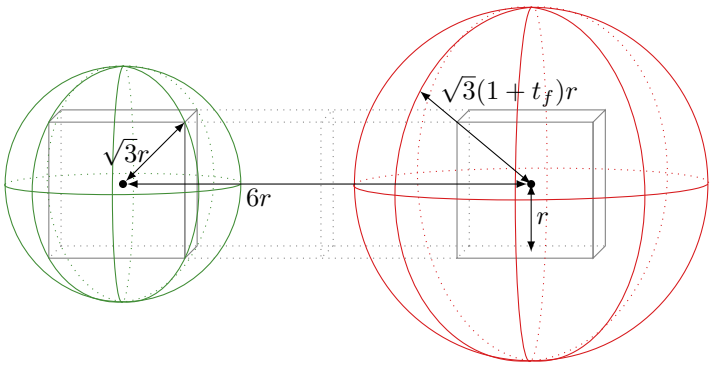
Stage 7: Propagate local expansions downward
for all boxes b in preorder **do**
For each child d of b, shift the center of the local expansions $\mathsf{L}_b^{\text{far}}$ to the child. Add the resulting expansions to $\mathsf{L}_d^{\text{far}}$ respectively.
end for

Stage 8: Form local expansions at QBX centers
for all boxes b **do**
For each QBX center c owned by b, translate $\mathsf{L}_b^{\text{far}}$ to c, obtaining $\mathsf{L}_c^{\text{qbx,far}}$.
end for

Stage 9: Evaluate final potential at targets
for all boxes b **do**
For each conventional target t owned by b, evaluate $\mathsf{L}_b^{\text{far}}(t)$.
Add $\mathsf{P}_b^{\text{near}}(t), \mathsf{P}_b^W(t), \mathsf{L}_b^{\text{far}}(t)$ to obtain the potential at t.
end for
for all boxes b **do**
For each target t associated to a QBX center c owned by b, add $\mathsf{L}_c^{\text{qbx,near}}(t), \mathsf{L}_c^{\text{qbx},W}(t), \mathsf{L}_c^{\text{qbx,far}}(t)$ to obtain the QBX local expansion due to c evaluated at t.
end for

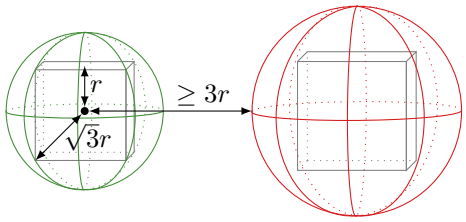
5.4 Accuracy

We carry out accuracy estimates of algorithm assuming the truth of the hypotheses in Section 4. These hypotheses imply that the accuracy of the FMM in approximating the unaccelerated version of the potential is mainly determined by the choice of t_f, with a smaller value of t_f leading to more accurate results. Choosing a value of $t_f \lessapprox 0.85$ recovers a convergence factor approximately the same as the 1-away point FMM, which is 3/4 [45].



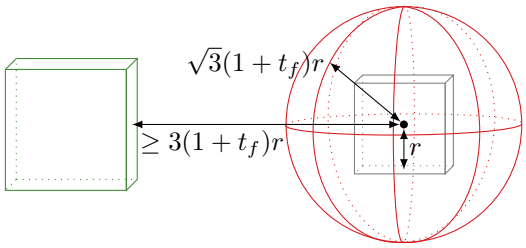
$$\frac{\text{furthest target}}{\text{closest source}} \leq \frac{\sqrt{3}(1+t_f)}{6-\sqrt{3}}$$

Figure 8: Convergence factor calculation for List 2. On the left is the source box and on the right are the target box and target confinement region.



$$\frac{\text{furthest source}}{\text{closest target}} \leq \frac{\sqrt{3}}{3}$$

Figure 9: Convergence factor calculation for List 3 far. The source box is on the left and the target box and target confinement region are on the right.



$$\frac{\text{furthest target}}{\text{closest source}} \leq \frac{\sqrt{3}}{3}$$

Figure 10: Convergence factor calculation for List 4 far. The source box is on the left and the target box and target confinement region are on the right.

Theorem 6 (Accuracy estimate for GIGAQBX algorithm)**.** *Fix a target confinement factor* $0 \leq t_f < 2\sqrt{3} - 2 \approx 1.47$*. Let* R *denote the radius of the smallest box in the tree and let* $\omega = \min\left(3-\sqrt{3}, 6-2\sqrt{3}-\sqrt{3}t_f\right)$*. Assuming the truth of Hypotheses 1, 2, and 3, there exists a constant* $M > 0$ *such that for every target point* $x \in \mathbb{R}^3$*, we have*

$$\left|\mathcal{S}_{\mathrm{QBX}(p_{\mathrm{qbx}},N)}\mu(x) - \mathcal{G}_{p_{\mathrm{fmm}}}[\mathcal{S}_{\mathrm{QBX}(p_{\mathrm{qbx}},N)}]\mu(x)\right| \leq \frac{MA}{\omega R} \max\left(\frac{\sqrt{3}(1+t_f)}{6-\sqrt{3}}, \frac{\sqrt{3}}{3}\right)^{p_{\mathrm{fmm}}+1}.$$

Here $\mathcal{S}_{\mathrm{QBX}(p_{\mathrm{qbx}},N)}\mu(x)$ *denotes the* p_{qbx}*-th order approximation to the single-layer potential given by QBX,* $\mathcal{G}_{p_{\mathrm{fmm}}}[\cdot]$ *denotes the approximation formed by the GIGAQBX FMM of order* p_{fmm}*, and* $A = \|\mu\|_\infty \sum_{i=1}^N |w_i|$*, where the* $\{w_i\}_{i=1}^N$ *are the quadrature weights. The constant* M *is independent of* t_f*,* ω*,* μ*, the particle distribution, and the QBX and FMM orders.*

In particular, for $t_f \leq (6\sqrt{3}-7)/4 \approx 0.85$*, we have*

$$\left|\mathcal{S}_{\mathrm{QBX}(p_{\mathrm{qbx}},N)}\mu(x) - \mathcal{G}_{p_{\mathrm{fmm}}}[\mathcal{S}_{\mathrm{QBX}(p_{\mathrm{qbx}},N)}]\mu(x)\right| \leq \frac{MA}{\omega R}\left(\frac{3}{4}\right)^{p_{\mathrm{fmm}}+1}.$$

Proof. Without loss of generality, we will assume that x is associated to a QBX center c. The proof of this

Table 1: Complexity of each stage of the GIGAQBX algorithm.

Stage	Modeled Operation Count	Note
Stage 1	NL	There are N total particles with at most L levels of refinement.
Stage 2	$N_S {p_{\text{fmm}}}^2 + N_B {p_{\text{fmm}}}^3$	$N_S {p_{\text{fmm}}}^2$ for forming multipoles and the rest for shifting multipoles upward, with each shift costing ${p_{\text{fmm}}}^3$.
Stage 3	$(27(N_C + N_S)n_{\max} + N_C M_C){p_{\text{qbx}}}^2$	Lemma 11
Stage 4	$875 N_B {p_{\text{fmm}}}^3$	Lemma 12
Stage 5	$N_C M_C {p_{\text{qbx}}}^2 + 124 L N_S n_{\max} {p_{\text{qbx}}}^2$	Lemma 13
Stage 6	$375 N_B n_{\max} {p_{\text{fmm}}}^2 + 250 N_C n_{\max} {p_{\text{qbx}}}^2$	Lemma 15
Stage 7	$8 N_B {p_{\text{fmm}}}^3$	The cost of shifting a local expansion downward is ${p_{\text{fmm}}}^3$. There are at most 8 children per box.
Stage 8	$N_C {p_{\text{fmm}}}^3$	Cost of translating the box local expansions to N_C centers.
Stage 9	$N_T {p_{\text{qbx}}}^2$	Cost of evaluating QBX expansions at N_T targets.

statement follows from applying the results of Section 4 to the definitions of the interaction lists in Section 5. The potential at c is the sum of the contributions $\mathsf{L}_c^{\text{qbx,near}}(t)$, $\mathsf{L}_c^{\text{qbx},W}(t)$, and $\mathsf{L}_c^{\text{qbx,far}}(t)$.

The potential due to $\mathsf{L}_c^{\text{qbx,near}}(t)$ must arrive via a direct interaction. This contribution incurs no (acceleration) error.

The potential due to $\mathsf{L}_c^{\text{qbx},W}(t)$, which arrives via a multipole-to-QBX-local interaction, incurs an error of at most

$$\frac{CA}{(3-\sqrt{3})R}\left(\frac{\sqrt{3}}{3}\right)^{p_{\text{fmm}}+1} \tag{28}$$

where $C > 0$ is some constant. See Hypothesis 1 and Figure 9.

The potential due to $\mathsf{L}_c^{\text{qbx,far}}(t)$ arrives via an interaction of $X_{b'}^{\text{far}}$ or $V_{b'}$, where b' is the box that owns c or an ancestor box. The contribution that comes from $X_{b'}^{\text{far}}$ arrives via a local-to-QBX-center interaction, and incurs an error of at most

$$\frac{CA}{(3-\sqrt{3})R}\left(\frac{\sqrt{3}}{3}\right)^{p_{\text{fmm}}+1} \tag{29}$$

where $C > 0$ is some constant. See Hypothesis 2 and Figure 10.

Lastly, the contribution due to all $V_{b'}$ interactions arrives via a multipole-to-local-to-QBX-center interaction and incurs an error of at most

$$\begin{aligned}\frac{CA}{(6-2\sqrt{3}-\sqrt{3}t_f)R}&\left(\left(\frac{\sqrt{3}(1+t_f)}{6-\sqrt{3}}\right)^{p_{\text{fmm}}+1}+\left(\frac{\sqrt{3}}{6-\sqrt{3}(1+t_f)}\right)^{p_{\text{fmm}}+1}\right)\\ &\leq \frac{2CA}{(6-2\sqrt{3}-\sqrt{3}t_f)R}\left(\frac{\sqrt{3}(1+t_f)}{6-\sqrt{3}}\right)^{p_{\text{fmm}}+1}\end{aligned} \tag{30}$$

where $C > 0$ is some constant. To obtain this bound, we have used the stated assumption that $t_f < 2\sqrt{3}-2$. See Hypothesis 3 and Figure 8 for the convergence factor calculation.

The result follows from combining (30), (28), and (29). □

5.5 Complexity

In this section we discuss the (time) complexity of the GIGAQBX algorithm in three dimensions. The costs presented in this section report the asymptotic number of 'modeled floating point operations' (or 'modeled

Table 2: Parameters to the complexity analysis.

Parameter	Note
p_{fmm}	FMM order.
p_{qbx}	QBX order.
$n_{\max}$	Desired bound on the number of particles per box (in some cases there may be more, see Section C.2).
t_f	Target confinement factor.
N_B	Number of boxes in the tree.
L	Number of levels in the tree.
N_S	Number of sources in the tree.
N_C	Number of QBX centers in the tree.
N_T	Number of targets in the tree.
N	$N \coloneqq N_S + N_C + N_T$; the number of 'particles'.
M_C	The average number of source particles inside $\overline{B_\infty}(4\sqrt{3}r_c t_f, c)$, taken over all QBX centers c (see also Section C.2).

flops') performed by the algorithm. The parameters we introduce in the complexity analysis are summarized in Table 2. Table 1 and Theorem 7 in this section provide a summary of the complexity analysis. For extended details, see Appendix C.

5.5.1 Assumptions

We make a number of simplifying assumptions in our complexity analysis.

We assume that all targets have been assigned to a QBX center, so that all evaluation at targets is done in Stage 9. This is the primary usage pattern for on-surface evaluation of a layer potential.

We assume, consistent with the hypotheses of Theorem 6, that $t_f < 2$. An assumption of this nature is useful for the analysis of List 4 (Lemma 15).

We assume the use of spherical harmonic expansions throughout the algorithm. The cost of translation is modeled using 'point-and-shoot' translation operators (see Section C.1 for details). We also assume that $p_{\text{qbx}} \leq p_{\text{fmm}}$.

5.5.2 Summary

Theorem 7 summarizes the complexity of Algorithm 4. The cost of the tree build phase (Stage 1) and the evaluation phase of the algorithm (Stages 2–9) are treated separately. Under broadly applicable assumptions, the evaluation phase can be shown to run in time that is proportional to the number of particles. Nevertheless, the proportionality constant is affected by the details of the particle distribution in two ways. First, the average size of the 'near neighborhoods' of QBX centers affects the number of direct interactions in the algorithm. This is measured by the parameter M_C. Second, the number of boxes in the tree, N_B, affects the number of intermediate expansions that are formed by the algorithm. This parameter is also determined by the details of the particle distribution.

In the final statement of the complexity analysis, we make the simplifying assumption that $M_C = O(1)$ and $N_B = O(N)$.

Theorem 7 (Complexity estimate for GIGAQBX algorithm)**.** *(a) The cost in modeled flops of the tree build phase of the GIGAQBX FMM is $O(NL)$.*

(b) Assume that $p_{\text{fmm}} = O(|\log \varepsilon|)$. For a fixed value of $n_{\max}$, the cost in modeled flops of the evaluation stage of the GIGAQBX FMM is $O((N_C + N_S + N_B)|\log\varepsilon|^3 + (N_C M_C + NL)|\log\varepsilon|^2 + N_T|\log\varepsilon|^2)$. With a

Table 3: ℓ^∞ error in Green's formula $\mathcal{S}(\partial_n u) - \mathcal{D}(u) = u/2$, scaled by $1/\|u\|_\infty$, for the 'urchin' γ_8, using the GIGAQBX algorithm. p_{fmm} denotes the FMM order and p_{qbx} the QBX order. The geometry was discretized with 4.85×10^4 triangles for the stage-1 discretization, and 2.78×10^5 triangles for the stage-2 discretization, with 2.95×10^2 nodes per element. An idealized a-priori estimate for the 1-away point FMM error [45] is included in the first column for comparison. Entries in bold indicate that the FMM error is negligible compared to the other error contributions.

$(3/4)^{p_{\text{fmm}}+1}$	p_{fmm}	$p_{\text{qbx}} = 3$	$p_{\text{qbx}} = 5$	$p_{\text{qbx}} = 7$	$p_{\text{qbx}} = 9$
3.16×10^{-1}	3	8.29×10^{-3}	9.68×10^{-3}	9.15×10^{-3}	9.18×10^{-3}
1.78×10^{-1}	5	1.43×10^{-3}	2.67×10^{-3}	2.85×10^{-3}	2.78×10^{-3}
4.22×10^{-2}	10	$\mathbf{6.08 \times 10^{-5}}$	6.44×10^{-5}	1.27×10^{-4}	1.47×10^{-4}
1.00×10^{-2}	15	$\mathbf{6.08 \times 10^{-5}}$	$\mathbf{6.38 \times 10^{-6}}$	3.24×10^{-6}	7.07×10^{-6}
2.38×10^{-3}	20	$\mathbf{6.08 \times 10^{-5}}$	$\mathbf{6.38 \times 10^{-6}}$	1.41×10^{-6}	2.51×10^{-7}

level-restricted octree and $t_f < \sqrt{3} - 1$, the modeled cost is $O((N_C + N_S + N_B)|\log \varepsilon|^3 + N_C M_C |\log \varepsilon|^2 + N_T |\log \varepsilon|^2)$. Assuming that the particle distribution satisfies $N_B = O(N)$ and $M_C = O(1)$, the worst-case modeled cost using a level-restricted octree and $t_f < \sqrt{3} - 1$ is linear in N (with a constant dependent on the particle distribution and the desired accuracy ε).

Proof. The estimate for (a) follows from the cost of Stage 1 as listed in Table 1. This estimate for (b) follows from adding up the costs of Stages 2–9 as found in Table 1. The linear running time in the case of a level-restricted octree and $t_f < \sqrt{3} - 1$ follows from Remark 14. □

6 Numerical Experiments

We use a family of smooth 'urchin' test geometries γ_k given analytically in spherical coordinates (r_k, θ, ϕ) by prescribing r_k as a function of (θ, ϕ), where

$$r_k(\theta, \phi) = 0.2 + \frac{\operatorname{Re} Y_k^{\lfloor k/2 \rfloor}(\theta, \phi) - m_k}{M_k - m_k}, \tag{31}$$
$$\begin{matrix} M_k \\ m_k \end{matrix} = \begin{matrix} \max \\ \min \theta \in [0,\pi], \phi \in [0, 2\pi] \end{matrix} \operatorname{Re} Y_k^{\lfloor k/2 \rfloor}(\theta, \phi),$$

using the definition of spherical harmonics from (4). Figure 3 gives a visual impression of γ_8.

To obtain an accurate unstructured triangular mesh of γ_k, we use an icosahedron as a starting point. Each of the icosahedron's faces is equipped with a mapping $\Psi_k \in (P^8)^3$ and the expansion of Ψ_k in orthogonal polynomials on the triangle [15, 38] is computed. While the ℓ^2-norm of the coefficients of the mapping corresponding to the polynomials of the two highest total degrees exceeds 10^{-10} times the ℓ^2-norm of all coefficients of the mapping, the element is bisected, and the warping function (31) is (nodally) reevaluated.

6.1 Accuracy

We use an analogous procedure to the one from [57] to test the accuracy of the algorithm of this paper through a sequence of experiments. With u a harmonic function defined inside γ_8 and extending smoothly to the boundary, we make use of *Green's formula*. Because of smoothness, u has a well-defined normal derivative $\partial_n u$ at the boundary. Then Green's formula (e.g. [39, Theorem 6.5]) states that for $x \in \Gamma$,

$$\mathcal{S}(\partial_n u)(x) - \mathcal{D}(u)(x) = \frac{u(x)}{2}.$$

We use the residual in this identity as a measure for the accuracy that our scheme achieves in the evaluation of layer potential evaluations. The achieved accuracy in Green's formula is predictive of the accuracy one

might achieve in the solution of boundary value problems. Data to support this assertion (in two dimensions) is presented in [57].

Letting u be the potential due to a charge located outside Γ at $(3,1,2)^T$, we evaluate $\mathcal{S}(\partial_n u) - \mathcal{D}(u)$ using our scheme and report the error in the discrete ℓ^∞-norm. The error reported is the absolute error scaled by $1/\|u\|_\infty$. We use the urchin geometry γ_8 and test with various combinations of QBX order p_{qbx} and FMM order p_{fmm}. γ_8 was discretized with 48 500 triangles for the stage-1 discretization, and 277 712 triangles for the stage-2 discretization, with 295 nodes per element, to eliminate the influence of quadrature error as a confounding factor in this experiment.

To 'balance' the algorithm, we compute a modeled flop count (cf. Sections 5.5 and 6.3) and chose the value of $n_{\max}$ that minimizes this modeled cost. In our experiments, $n_{\max} = 512$ was the approximate minimizer. We choose the target confinement factor as $t_f = 0.9$.

Table 3 shows the results of these experiments for the GIGAQBX FMM, scaled by the norm of the test function u and varying p_{qbx} across columns and p_{fmm} across rows. We show table entries in bold if no decrease in error is observed for at least one subsequent value of p_{fmm}.

The error in each entry of the table may be interpreted as the additive contribution of truncation error (Lemma 1), quadrature error (Lemma 2), and FMM acceleration error (Theorem 6). The latter two sources of error are present even without FMM acceleration. Thus, reading down a given column (with p_{qbx} held fixed and p_{fmm} varying), we do not expect the error to decrease below a fixed amount, which we term the 'unaccelerated QBX error.' This quantity empirically corresponds to the error shown in bold.

The large decrease in the error as p_{fmm} increases suggests that as long as the error still decreases with p_{fmm}, we can assume that the acceleration error is the dominant error component. Reading across a row of the table (with p_{fmm} held fixed and p_{qbx} varying), we observe that if an entry appears to be dominated by acceleration error (i.e. is not in bold), the errors in the row are very roughly of the same order of magnitude. This is consistent with Theorem 6, which implies an FMM acceleration error bound that is independent of the QBX order.

For our chosen value of t_f, Theorem 6 roughly establishes $\|u\|_\infty (3/4)^{p_{\text{fmm}}+1}$ as a bound on the absolute error incurred by acceleration, neglecting a number of other factors given in the precise statement of the theorem. We show $(3/4)^{p_{\text{fmm}}+1}$ in the left column of the table. Importantly, this value is an upper bound for the values in its row, and thus also a bound on the acceleration error. The strict obedience to this bound also confirms that the algorithm does not require an FMM order increase to maintain accuracy (cf. [48] and the discussion in Section 2.4).

The actual acceleration component of the error in Table 3 in fact appears to decrease more rapidly than the first column. A similar phenomenon was observed for the two-dimensional case in [57]. This is not entirely unexpected, as the error in the potential is a weighted average of the individual errors due to the source particles, which are likely to be separated more generously from the target than the worst-case estimates in Section 4 assume.

In summary, the results in this table empirically confirm the validity of Theorem 6 as well as of an additive error model:

$$|\text{total error}| \leq |\text{unaccelerated QBX error}| + \|u\|_\infty (3/4)^{p_{\text{fmm}}+1}. \tag{32}$$

6.2 A BVP with Complex Geometry for the Helmholtz Equation

To support the assertion that our algorithm is broadly applicable and robust, we demonstrate its use on a challenging, moderate-frequency boundary value problem for the Helmholtz equation. While we have discussed a version of the algorithm for the Laplace equation, a direct analog of our algorithm is applicable for the Helmholtz and many other related elliptic PDEs, assuming the availability of translation operators with suitable complexity. We expect our complexity and accuracy analysis to carry over to the case of the Helmholtz equation with only minor changes. Empirical evidence suggests that this is the case.

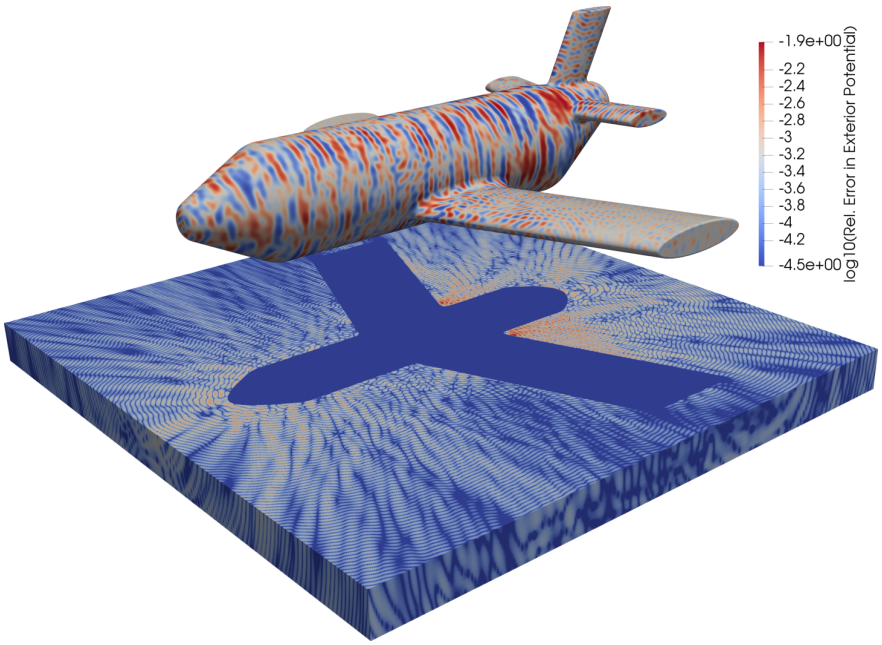


Figure 11: An exterior Dirichlet boundary value problem for the Helmholtz equation solved on a 'toy plane' geometry. The shading on the geometry itself reflects the obtained density μ using a Brakhage-Werner representation [4] $-\mathcal{D}\mu + i\mathcal{S}\mu$. The volume visualization illustrates the logarithm of error in the computed exterior potential. The top of the volume visualization represents a cut roughly at 'wing' level of the source geometry. The maximal relative ℓ^∞ error observed anywhere in the exterior computational domain (which extends to cover the entire geometry, including at points on or near the surface) was 1.38×10^{-2}. Section 6.2 describes the computational setup in more detail. The potential is evaluated at 104 947 200 targets in the volume, with 11 482 688 source points, 1 230 288 QBX centers, and 615 144 on-surface targets.

We solve an exterior Dirichlet boundary value problem

$$\begin{aligned} \left(\triangle + k^2\right) u &= 0 \quad \text{in } \mathbb{R}^3 \setminus \Omega, \\ u &= f \quad \text{on } \partial\Omega, \\ \lim_{r\to\infty} r\left(\frac{\partial}{\partial r} - ik\right) u &= 0 \end{aligned}$$

where $\Omega \subset \mathbb{R}^3$ is a closed, bounded region with smooth boundary $\Gamma = \partial\Omega$. Ω is given by the geometry `surface-3d/betterplane.brep` from [36]. We obtain a surface mesh consisting of triangles with second-order polynomial mapping functions for the geometry using Gmsh [18]. The geometry ('nose' to 'tail') is approximately 19 units long and 20 units wide ('wingtip' to 'wingtip'). The original geometry has 37 244 triangles, the stage-1 mesh has 45 755 order 2 triangles (with 6 nodes per element), and the stage-2 mesh has 102 524 order-2 triangles; the triangles of the stage-2 quadrature discretization each have 112 nodes. The Helmholtz parameter was set to $k = 20$.

We use a Brakhage-Werner representation [4] to solve for boundary values obtained from a point potential emanating from a number of sources in the 'tail' of the geometry. Using L^2-weighted degrees of freedom [5], GMRES [50] attained a decrease in the residual norm by a factor of 10^{-5} in 79 iterations. The calculation took around two days on a dual-socket Intel Xeon E5-2650 v4 machine. The Helmholtz translation operators used in the algorithm were those from FMMLIB3D [20, 19].

We verify that the potential obtained from the boundary value solve matches the point potential through point evaluations in the volume, obtaining roughly two digits of accuracy. Details of the relative error in the potential evaluation in the volume can be found in Figure 11.

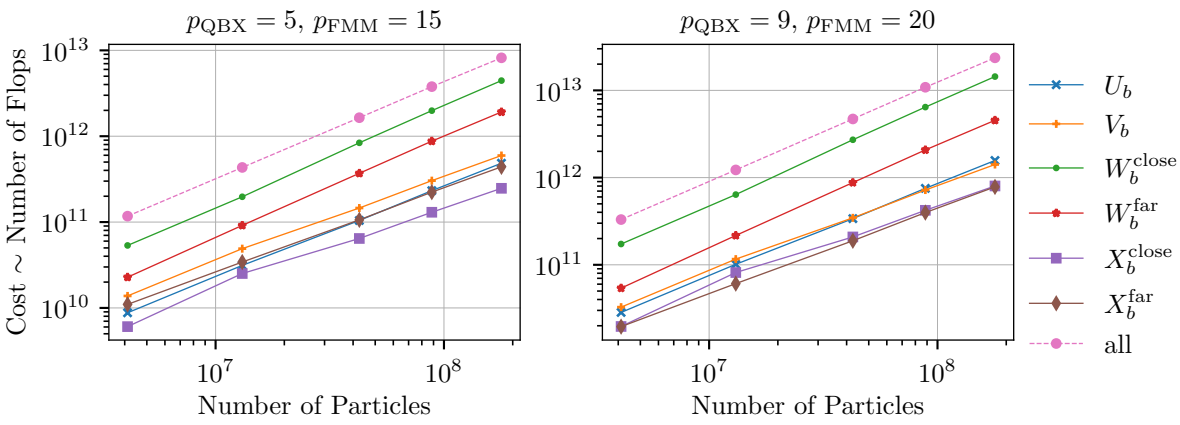


Figure 12: Modeled operation counts for the GIGAQBX FMM for evaluating the single-layer potential on a sequence of 'urchin' geometries of increasing particle count. The operations are counted according to the model presented in Table 4. Here, $n_{\max} = 512$ and $t_f = 0.9$. The scaling test used the 'urchin' geometries $\gamma_2, \gamma_4, \ldots, \gamma_{10}$.

6.3 Cost and Scaling

It remains to examine both the computational cost and the scaling thereof that the algorithm achieves on geometries of varying size. Rather than relying on wall time (which is sensitive to machine details as well as varying levels of optimization and code quality), we present an abstract operation count intended to asymptotically match the number of floating point operations, similarly to the approach of Section C.1. We account for each entry in the interaction lists of Section C.3 with the counts shown in Table 4. We use the 'urchin' test geometries $\gamma_2, \gamma_4, \ldots, \gamma_{10}$ for this computational experiment.

We show cost data for two pairs of $(p_{\mathrm{qbx}}, p_{\mathrm{fmm}})$, corresponding to different accuracies. The first, $(p_{\mathrm{qbx}}, p_{\mathrm{fmm}}) = (5, 15)$, corresponds to roughly five digits of accuracy following Table 3, whereas the second, $(p_{\mathrm{qbx}}, p_{\mathrm{fmm}}) = (9, 20)$, corresponds to around seven digits of accuracy.

As in point FMMs, the main tuning parameter that may be used to balance various cost contributions and minimize computational cost is $n_{\max}$, the maximal number of particles per box. We chose $n_{\max}$ to minimize the modeled computational cost, obtaining a value of $n_{\max} = 512$. The TCF t_f mainly trades off cost and accuracy, we choose $t_f = 0.9$. We show graphs of computational cost across geometry sizes in Figure 12.

Unlike in two dimensions, we observe that W^{close} and, to a lesser extent, W^{far} dominate the run time of the algorithm. This is not entirely unexpected, as the size of the TCR, naturally larger by particle count in three dimensions, makes its influence felt. A further factor in the large contribution of W^{far} is the high cost of translations even when the target (QBX) expansion is of comparatively low order, cf. Section C.1.

In accordance with the results of Section 5.5, the experiments support the conclusion that the algorithm exhibits linear scaling in the number of source and target particles, with one decade of geometry growth (indicated by the vertical grid lines) leading to one decade of cost growth (indicated by horizontal grid lines).

6.4 Cost Implications of the ℓ^2-Based Target Confinement Region

Next, we seek to understand the impact of the change in the shape of the TCR, which was box-shaped and defined by the ℓ^∞-norm in the earlier version of our algorithm [57], but which now is spherical and measured by an ℓ^2-norm to better match the actual region of convergence of the obtained local expansions. Table 5 summarizes the results of an experiment determining the comparative cost of both approaches. Both versions of the algorithm were balanced individually before conducting the experiments, in both cases $n_{\max} = 512$ turned out to be near-optimal. First, we observe that the algorithmic change has led to a reduction of (modeled) computational cost by around 25 per cent. We note a marked increase in the cost contribution of the V list, as well as marked decreases in the cost of the U and W^{far} lists, all of which are indicative of the

Table 4: Cost per interaction list entry modeled in Figure 12, i.e. for a single (source box, target box) interaction list pair. p_{fmm} = FMM order and p_{qbx} = QBX order. n_s = number of sources in the source box and n_t = number of QBX centers in the target box.

List	Cost
U_b	${p_{\text{qbx}}}^2 n_s n_t$
V_b	${p_{\text{fmm}}}^3$
W_b^{close}	${p_{\text{qbx}}}^2 n_s n_t$
W_b^{far}	${p_{\text{fmm}}}^3 n_t$
X_b^{close}	${p_{\text{qbx}}}^2 n_s n_t$
X_b^{far}	${p_{\text{fmm}}}^2 n_s$

Table 5: A comparison of (modeled, cf. Table 4) cost for the GIGAQBX FMM between the use of the conventional ℓ^∞ box extent norms (analogous to [57]) and the ℓ^2 box extent norms introduced in this article. Entries in the table show modeled floating point operations in the sense of Section 5.5. The experiment used the 'urchin' geometry γ_8 discretized with 48,500 stage-1 elements and 277,712 stage-2 elements, where each of the latter had 295 nodes. The evaluated columns use $p_{\text{fmm}} = 15$ and $p_{\text{qbx}} = 5$, corresponding to around five digits of accuracy following Table 3. For brevity, we let $p = p_{\text{fmm}}$ and $q = p_{\text{qbx}}$. Note that the rows shown do *not* add up to the shown total. The latter includes minor contributions to the overall cost (such as the upward and downward passes) that we have omitted.

	ℓ^∞ (sym.)	ℓ^∞	ℓ^2 (sym.)	ℓ^2
U_b	$2.77 \times 10^{10} q^2$	6.92×10^{11}	$9.29 \times 10^{9} q^2$	2.32×10^{11}
V_b	$6.14 \times 10^{7} p^3$	2.07×10^{11}	$9.00 \times 10^{7} p^3$	3.04×10^{11}
W_b^{close}	$8.88 \times 10^{10} q^2$	2.22×10^{12}	$7.94 \times 10^{10} q^2$	1.99×10^{12}
W_b^{far}	$4.56 \times 10^{8} p^3$	1.54×10^{12}	$2.59 \times 10^{8} p^3$	8.76×10^{11}
X_b^{close}	$2.62 \times 10^{9} q^2$	6.55×10^{10}	$5.20 \times 10^{9} q^2$	1.30×10^{11}
X_b^{far}	$9.51 \times 10^{8} p^2$	2.14×10^{11}	$9.92 \times 10^{8} p^2$	2.23×10^{11}
Total		4.97×10^{12}		3.78×10^{12}

higher efficiency of the method with the ℓ^2 TCR.

7 Conclusion

This paper introduces a fast algorithm for the accurate evaluation of layer potentials in three dimensions using Quadrature by Expansion (QBX).

Our work builds on and extends the GIGAQBX algorithm in two dimensions [57]. Many features of the algorithm carry over broadly unchanged from the two dimensional setting. However, some parts have required careful redevelopment. A practical QBX implementation must provide a mechanism to control for truncation error, quadrature error, and error introduced by FMM acceleration. To address these challenges in three dimensions, our work combines new error estimates for FMM translations in three dimensions, a new local refinement criterion for truncation error control based on scaled-curvature, and a novel adaptive refinement scheme for achieving source quadrature resolution. In a series of numerical experiments, we demonstrate that this combination can achieve high accuracy for layer potential evaluation on complicated geometries. In particular, we show how the FMM acceleration recovers similar levels of acceleration error as the point FMM. The numerical evidence for the usefulness of our error control strategies is robust. A rigorous mathematical treatment of these error control strategies appears eminently feasible in some cases, such as for translation operators. We leave this as a subject of future investigation.

Additionally, we describe a benign set of sufficient conditions on the geometry under which the running time of the algorithm has linear complexity. By counting modeled flops on large scale geometries, the scaling of the algorithm is shown to be linear in practice. The most expensive part of the algorithm is the QBX near field evaluations. Fortunately, strategies such as changing the shape of the target confinement region are available to reduce this expense. Further optimizations to the scheme are the subject of ongoing investigation.

Acknowledgments

The authors' research was supported by the National Science Foundation under award numbers DMS-1418961 and DMS-1654756. Any opinions, findings, and conclusions, or recommendations expressed in this article are those of the authors and do not necessarily reflect the views of the National Science Foundation; NSF has not approved or endorsed its content. Part of the work was performed while the authors were participating in the HKUST-ICERM workshop 'Integral Equation Methods, Fast Algorithms and Their Applications to Fluid Dynamics and Materials Science' held in 2017. The authors would like to thank Hao Gao for identifying a software issue in a prior version of this work.

A Area Queries

Area queries were introduced in [48] in two dimensions. We describe their (largely straightforward) three-dimensional generalization in this section. They form the core mechanism on which the many of the geometric operations in this article are performed. Given a center c and a radius r, the area query computes the set of leaf boxes that intersect $\overline{B_\infty}(r, c)$. It is assumed that c falls inside the computational domain and that r is at most the radius of the tree.

Since the area query as a primitive is used to retrieve sets of sources and targets by way of their containing boxes, it may appear flawed that the area query only considers leaf boxes, when the fast algorithm of Section 5 permits targets (specifically, those with extent, i.e. QBX expansion balls) to occur in non-leaf boxes. This is not an issue since all application scenarios of the area query (Sections 3.3, 3.4, and 3.5) use it to find point-shape objects (i.e. objects without extent) which are necessarily found in leaf boxes of the tree.

An area query proceeds by descending the tree towards the query center c until the descent has reached a box whose size is commensurate with the size of the query box $\overline{B_\infty}(r, c)$. This box is referred to as the *guiding box*. Specifically, the guiding box is the smallest box whose 1-near neighborhood contains $\overline{B_\infty}(r, c)$. Once this box has been found, only the leaf descendants of the 1-near neighborhood of b need to be checked for intersection with the query box. The full procedure to carry out an area query is given in Algorithm 5.

Recall that a *colleague* of a box (what we also refer to as a 1-colleague) is a box on the same level as b that is adjacent to b. In a non-adaptive tree, the 1-near neighborhood is specified by a box b and its set of colleagues. In an adaptive tree, this is no longer the case as some colleagues may be missing. The notion of a *peer box* is a generalization of a colleague which allows for larger boxes to stand in as colleagues if necessary. This makes it useful for the area query.

Definition 10 (Peer box [48])**.** *Let b be a box in an octree. A box c is a* peer box *of b if (a) c is b or adjacent to b, (b) the size of c is at least the size of b, and (c) no child of c satisfies the previous criteria.*

Algorithm 5: Area Query

Require: A center c and a query radius r.
Ensure: Computes the set of leaf boxes which intersect $\overline{B_\infty}(r, c)$.

Find the guiding box b
$b \leftarrow$ the root box.
loop
  if $|b| < r \leq 2|b|$ **or** b has no child containing c **then**
    break
  end if
  $b \leftarrow$ the child of b containing c.
end loop

Parameter	Choice
R	$\{0.1, 1, 10\}$
ρ	$\{0.1, 1, 10\}$
r	$\{0.25\rho, 0.5\rho, 0.75\rho\}$
q	$\{3, 5, 10, 15, 20\}$
p	$\{3, 5, 10, 15, 20\}$

Table 6: Summary of parameters chosen for the numerical experiments used to obtain the results in Sections 4.3.1, 4.3.2, and 4.3.3.

Check leaf descendants of b's peers
for all peers p of b **do**
  for all leaf descendants l of p **do**
    if $\overline{B_\infty}(r, c) \cap l$ is nonempty **then**
      Add l to the output set.
    end if
  end for
end for

B Numerical Experiments in Support of FMM Translation Error Estimates

The code used for the numerical experiments performed to obtain the results of Section 4 is available at [56]. In this appendix, we describe the procedure the code uses.

B.1 Multipole and Multipole-to-Local Accuracy

We use the notation of Section 4.3.1. Hypothesis 1 pertains to the accuracy of approximating a local expansion using an intermediate multipole expansion.

As a numerical experiment, we test the truth of this hypothesis at selected values of the parameters (R, r, ρ, p, q). For a given value of these parameters, estimate of the value $E_M(q)$ (25) is produced by sampling this value at $42^2 \cdot 57$ tuples of the form (s, c, c', t). The values of the parameters (R, r, ρ, p, q) are chosen according to Table 6. In total, 675 parameter tuples are tested.

The details of the sampling procedure are as follows. A multipole expansion center is placed at $c = (0, 0, R + \rho)$. First, 42 sources s are selected from the sphere of radius r centered at c. Second, 57 centers c' are selected in the ball of radius R centered at the origin. Third, for each center c', 42 targets t are selected from the sphere of radius $R - |c'|$ centered at c'.

The points selected on the sphere are selected to be approximately equispaced [13] and included the poles of the sphere. The points in the ball are selected from concentric spheres inside the ball.

For each sampled value (s, c, c', t), the multipole expansion due to s is formed at c and translated to a local expansion centered at c'. The quantity $E_M(q)$ is evaluated at the target t. We save the largest observed value of $E_M(q)$ and take this as an upper bound on that quantity for the given set of parameters (R, r, ρ, p, q).

Hypothesis 3 in Section 4.3.3 pertains to the accuracy of a local expansion approximated by an intermediate multipole and local expansion. This hypothesis is the same geometric setting as Hypothesis 1. We follow an identical sampling procedure to obtain a numerical estimate of the quantity $E_{M2L}(q)$ (27).

B.2 Local Accuracy

Section 4.3.2 uses a different geometrical scenario from the multipole case. Hypothesis 2 in this section pertains to the accuracy of approximating a local expansion using an intermediate local expansion.

This hypothesis is tested at selected values of the parameters (r, ρ, p, q). For a given value of these parameters, we obtain an estimate of the quantity $E_L(q)$ (26) by sampling a set of $42 \cdot 57$ tuples of the form (s, c, t) using a sampling procedure. The parameters are taken according to Table 6, with a total of 225 parameter tuples tested.

The details of the sampling procedure are as follows. First, the source s is placed at $(0, 0, \rho)$. Second, 57 centers c are chosen from inside the ball of radius r centered at the origin. Third, each center c, 42 targets t is chosen from the sphere of radius $r - |c|$ centered at c.

For each value of the tuple (s, c, t), a p-th order local expansion due to s is formed at the origin. The local expansion is subsequently translated to a q-th order local expansion at c. The quantity $E_L(q)$ is evaluated at t. The largest observed value of $E_L(q)$ is taken as an estimate of the upper bound on the quantity.

C Detailed Complexity Analysis

This section provides the details of the complexity analysis from Section 5.5, under the assumptions highlighted in Section 5.5.1. In Section C.1 and C.2 we review the complexity of translations and the effect of the particle distribution. Section C.3 provides the details supporting the analysis in Table 1.

We use the parameters from Table 2 throughout this section. In addition to the parameters from this table, we make use of the following notation. First, we let S and C denote, respectively the sets of sources and QBX centers in the tree. Second, given a particle p in the tree, we let b_p denote the box owning p.

C.1 Complexity of Translation Operators

A p-th order multipole/local expansion requires $(p+1)^2$ expansion coefficients. The $(p+1)^2$ corresponding basis functions for the coefficients may be evaluated in $O(p^2)$ time using well-known recurrences [14, Ch. 14]. As a result, we model the cost of forming or evaluating a p-th order multipole/local expansion in spherical harmonics as p^2 operations, which is correct to leading order.

The cost of translations of spherical harmonic expansions (multipole/local $\to$ local) may be modeled as follows. With a simple extension to the commonly used 'point and shoot' translation scheme [25], a p-th order expansion can be translated into a q-th order expansion in the following three steps:

1. Rotate the coordinate system so that the translation direction is along the z-axis at a cost of $O(p^3)$ operations.
2. Translate the expansion at a cost of $O(pq^2)$ operations.
3. Rotate the coordinate system back into the original at a cost of $O(q^3)$ operations.

It follows that the complexity of a translation operator is $O(\max(p, q)^3)$. If we assume that $p_{\text{qbx}} \leq p_{\text{fmm}}$, the cost of translating a p_{fmm}-th order expansion into a p_{qbx}-th order expansion can be modeled as ${p_{\text{fmm}}}^3$ operations.

C.2 Effect of Particle Distribution

The running time of the GIGAQBX FMM cannot be entirely independent of the particle distribution. Unlike the point FMM, the algorithm may place more than $n_{\max}$ particles in a box. This occurs due to clustering of QBX centers in boxes because of the target confinement rule. This phenomenon cannot be disregarded as it occurs in practice even with smooth geometries.

To handle this in the complexity analysis, we find it useful to distinguish between QBX centers that have 'settled' in a leaf box and those that are 'suspended', i.e. that cannot be placed in a lower box due to target confinement restrictions. By definition, only the latter kind of centers can cluster in the tree beyond $n_{\max}$ particles per box.

Definition 11 (Leaf-settled / suspended centers)**.** *Suppose the tree has been constructed following Algorithm 4. Call a center c owned by a box b a* suspended center *if c cannot be placed in any hypothetical child box of b due to target confinement restrictions. A center that is not suspended in any box is called* leaf-settled*.*

To bound the number of algorithmic operations involving suspended QBX centers, we introduce a parameter into the complexity analysis that corresponds to the average size of a 'neighborhood' of a suspended QBX center—in other words, the average number of sources with which a QBX center must interact directly. While this quantity may at first seem to be tied to the tree structure imposed on the geometry, it is possible to bound this quantity independently of the tree. A tree-independent bound on this quantity follows from the following lemma.

Lemma 8 (Size of a suspended QBX expansion ball relative to box neighborhood)**.** *Let c be a suspended QBX center of radius r_c owned by the box b_c. Then the closed cube $\overline{B_\infty}(4\sqrt{3}r_c/t_f, c)$ is, geometrically, a superset of the 2-near neighborhood of b_c.*

Proof. For any box b, the Euclidean distance from a point x inside b to a point on the boundary of $\mathsf{TCR}(b)$ is at least $\sqrt{3}|b|t_f$. This distance is minimized when x is a box corner.

Since c is suspended, c cannot fit in any hypothetical child box of b_c, which has radius $|b_c|/2$. It follows from the previous observation that

$$r_c > \sqrt{3}|b_c|t_f/2.$$

Regardless of where c is located in b_c, $\overline{B_\infty}(c, 6|b_c|)$ is a superset of the 2-near neighborhood of b. The claim follows by observing $4\sqrt{3}r_c/t_f > 6|b_c|$. □

The quantity M_C is then defined as follows:

$$M_C \coloneqq \frac{1}{N_C}\sum_{c\in C}\left|S \cap \overline{B_\infty}(4\sqrt{3}r_c/t_f, c)\right|$$

where r_c is the radius of the center c. The following proposition is an immediate consequence of the definition of M_C and the previous lemma.

Proposition 9 (Bound on 2-near-neighborhood interactions for QBX expansions)**.** *The number of source-center pairs (s, c), such that c is a suspended QBX center and s is a source particle in the 2-neighborhood of the box owning c, is at most $N_C M_C$.*

C.3 Complexity of Algorithmic Stages Associated with Interaction Lists

Proposition 10 (Number of larger leaf boxes in the 1-neighborhood of a box)**.** *Let b be a box. There are at most 27 leaf boxes at least at large as b intersecting the 1-near neighborhood of b.*

Proof. Let l be such a leaf box. If $l \neq b$, choose a box c_l which is a colleague of b that is geometrically contained inside l. The mapping $l \mapsto c_l$ is injective, and b has at most $3^3 - 1 = 26$ colleagues. □

Lemma 11 (List 1 complexity)**.** *The amount of work done in Stage 3 (direct evaluation of the potential from adjacent source boxes) is at most*

$$(27(N_C + N_S)n_{\max} + N_C M_C){p_{\text{qbx}}}^2.$$

Proof. Define the set U as $U = \{(s, c) \mid b_s \in U_{b_c}\}$. Each source-center interaction costs ${p_{\text{qbx}}}^2$ operations. The number of Stage 3 interactions is $|U|$, so the cost of Stage 3 is at most ${p_{\text{qbx}}}^2|U|$. U may be written as a disjoint union $U = U_{\text{big}} \cup U_{\text{small}}$, where U_{big} contains all pairs (s, c) such that $|b_s| \geq |b_c|$.

For any center c, Proposition 10 implies that there may be at most $27n_{\max}$ sources in leaf boxes at least as large as b_c contributing to the potential via U_{b_c}. Thus $|U_{\text{big}}| \leq 27N_C n_{\max}$.

In a similar way, one can show that there are at most $27N_S n_{\max}$ pairs $(s, c) \in U_{\text{small}}$ such that c is a leaf-settled center. By Proposition 9, there are at most $N_C M_C$ pairs $(s, c) \in U_{\text{small}}$ such that c is a suspended center. It follows that $|U_{\text{small}}| \leq N_C M_C + 27N_S n_{\max}$. □

Lemma 12 (List 2 complexity)**.** *The amount of work done in Stage 4 (translation of multipole to local expansions) is at most* $875 N_B {p_{\text{fmm}}}^3$.

Proof. There are at most N_B boxes. The size of List 2 for a box b is at most $10^3 - 5^3 = 875$, since there are at most 10^3 descendants of 2-colleagues of the parent of b, of which 5^3 are 2-colleagues of b itself so they cannot be in List 2 of b. Each multipole-to-local translation costs ${p_{\text{fmm}}}^3$ operations. □

Lemma 13 (List 3 complexity)**.** *The amount of work done in Stage 5 (evaluation of List 3 close and far) is at most*

$$N_C M_C {p_{\text{qbx}}}^2 + 124 L N_S n_{\max} {p_{\text{qbx}}}^2.$$

Proof. We make the simplifying assumption that *all* Stage 5 interactions are mediated by List 3 close. This assumption *will not lead to an undercount* of the cost of the Stage 5 interactions, if the optimization in Remark 5 has been applied. Recall that a List 3 far interaction is a multipole-to-target interaction, and a List 3 close interaction is a source-to-target interaction. The only way the above assumption could undercount the cost of Stage 5 is if a List 3 far interaction were more expensive than the equivalent interactions that occur with the leaf descendants using List 3 close. But if a List 3 far interaction is more expensive than List 3 close interaction, the interaction of the former type may be converted to the latter type with no loss in accuracy.

Under this assumption, Proposition 9 implies that the cost of all Stage 5 interactions aimed at suspended centers is at most $N_C M_C {p_{\text{qbx}}}^2$.

Now, consider the Stage 5 interactions aimed at leaf-settled centers. Let s be a source owned by a box b_s. If s interacts via List 3 close with a leaf-settled center c, then c must be owned by a box that is a 2-colleague of either b_s or an ancestor of b_s. A box has at most $5^3 - 1 = 124$ boxes that are 2-colleagues. Since each source-center interaction costs ${p_{\text{qbx}}}^2$ and there are at most $n_{\max}$ leaf-settled centers per box, the cost of all interactions aimed at leaf-settled centers is at most $124 L N_S n_{\max} {p_{\text{qbx}}}^2$. □

Remark 14 (Effect of level-restriction on List 3 complexity)**.** *The factor of L in Lemma 13 suggests that Stage 5 has a worst-case superlinear scaling. A number of modifications to the algorithm are available that can provably remove the asymptotic factor of L. For instance, a cost estimate for Stage 5 that is independent of L may be derived assuming that the tree is level-restricted, meaning that adjacent leaves differ by at most one level, and that $t_f < \sqrt{3} - 1 \approx 0.73$.*

For $t_f < \sqrt{3} - 1$, $\mathsf{TCR}(b)$ is contained strictly inside the 1-neighborhood of b. If the tree is level-restricted, this implies that if b is a leaf box, any box in $W_b^{\text{close}} \cup W_b^{\text{far}}$ cannot be more than a constant factor smaller than b. This further implies that for a leaf box b, the quantity $|W_b^{\text{close}} \cup W_b^{\text{far}}|$ is at most a constant that depends on t_f and on the dimension. This can be used to construct a cost estimate independent of L. We leave the details of the derivation to the reader.

Using a level-restricted tree does not impact the asymptotic scaling of any other stage of the algorithm. Any octree tree may be converted to be level-restricted by repeatedly subdividing the larger of the leaf boxes that violate the level-restriction criterion. The level-restricted tree that results has a constant factor as many boxes. See [44] for details.

Lemma 15 (List 4 complexity)**.** *The cost of all Stage 6 interactions (evaluation of the potential due to List 4 close and far) is at most*

$$375 N_B n_{\max} {p_{\text{fmm}}}^2 + 250 N_C n_{\max} {p_{\text{qbx}}}^2.$$

Proof. First, we show $|X_b| \leq 125$. Every box in X_b is a leaf that is either a 2-colleague of b not adjacent to b, or adjacent to the parent of b and at least as large as the parent of b. There are at most $5^3 - 3^3 = 98$ boxes that fall into the first category and, by Proposition 10, at most 27 boxes that fall into the second category.

Next, we show that $X_b^{\text{close}} \subseteq X_b \cup X_{\mathsf{Parent}(b)}$, which implies that $|X_b^{\text{close}}| \leq 250$. Recall that X_b^{close} must be a subset of the List 4's of the ancestors of b. If b' is an ancestor of b, then b must be separated by an ℓ^∞ distance of $2^{k+1}|b|$ from any box in $X_{b'}$. In particular, consider a box $e \in X_g$, where g is $k \geq 2$ levels above b. Then e will be separated by an ℓ^∞ distance of at least $8|b|$ from b. It follows that the ℓ^∞ distance from the center of b to the boundary of e is at least $9|b|$, which is at least $3(1+t_f)|b|$ (since $t_f < 2$ by assumption). Thus $\mathsf{TCR}(b) \prec e$ from the definition of '$\prec$'. It follows that X_b^{close} is disjoint from the List 4 of a grandparent of b or above.

Finally, $|X_b^{\text{far}}| \leq 375$ follows since, by definition, $X_b^{\text{far}} \subseteq X_b \cup X_{\mathsf{Parent}(b)}^{\text{close}}$.

The cost estimate follows since, for List 4 close, each center will interact directly with at most $250n_{\max}$ source particles, with each interaction costing ${p_{\text{qbx}}}^2$. For List 4 far, each box will interact with at most $375n_{\max}$ source particles, at a cost of ${p_{\text{fmm}}}^2$ per interaction. □

References

[1] K. E. Atkinson and D. Chien, *Piecewise polynomial collocation for boundary integral equations*, SIAM Journal on Scientific Computing 16.3 (1995), pp. 651–681, DOI: 10.1137/0916040.

[2] A. H. Barnett, *Evaluation of Layer Potentials Close to the Boundary for Laplace and Helmholtz Problems on Analytic Planar Domains*, SIAM Journal on Scientific Computing 36.2 (2014), A427–A451, DOI: 10.1137/120900253.

[3] J. T. Beale and M.-C. Lai, *A method for computing nearly singular integrals*, SIAM Journal on Scientific Computing 38.6 (2001), pp. 1902–1925, DOI: 10.1137/S0036142999362845.

[4] H. Brakhage and P. Werner, *Über das Dirichletsche Außenraumproblem für die Helmholtzsche Schwingungsgleichung*, Archiv der Mathematik 16.1 (1965), pp. 325–329, ISSN: 0003-889X, DOI: 10.1007/BF01220037.

[5] J. Bremer, *On the Nyström discretization of integral equations on planar curves with corners*, Applied and Computational Harmonic Analysis (2011), DOI: 10.1016/j.acha.2011.03.002.

[6] J. Bremer, Z. Gimbutas, and V. Rokhlin, *A nonlinear optimization procedure for generalized Gaussian quadratures*, SIAM Journal on Scientific Computing 32 (2010), pp. 1761–1788, DOI: 10.1137/080737046.

[7] J. Bremer, *A fast direct solver for the integral equations of scattering theory on planar curves with corners*, Journal of Computational Physics 231.4 (2012), pp. 1879–1899, ISSN: 0021-9991, DOI: 10.1016/j.jcp.2011.11.015.

[8] J. Bremer and Z. Gimbutas, *A Nyström method for weakly singular integral operators on surfaces*, J. Comput. Phys. 231.14 (2012), pp. 4885–4903, ISSN: 0021-9991, DOI: 10.1016/j.jcp.2012.04.003.

[9] O. P. Bruno and L. A. Kunyansky, *A fast, high-order algorithm for the solution of surface scattering problems: basic implementation, tests, and applications*, Journal of Computational Physics 169 (2001), pp. 80–110, DOI: 10.1006/jcph.2001.6714.

[10] M. Carley, *Numerical quadratures for singular and hypersingular integrals in boundary element methods*, SIAM Journal on Scientific Computing 29.3 (2007), pp. 1207–1216.

[11] R. Chapko, R. Kress, and L. Mönch, *On the numerical solution of a hypersingular integral equation for elastic scattering from a planar crack*, en, IMA J Numer Anal 20.4 (Oct. 2000), pp. 601–619, DOI: 10.1093/imanum/20.4.601.

[12] P. J. Davis and P. Rabinowitz, *Methods of Numerical Integration*, Academic Press, San Diego, 1984.

[13] M. Deserno, *How to generate equidistributed points on the surface of a sphere*, https://www.cmu.edu/biolphys/deserno/pdf/sphere_equi.pdf, accessed March 2018, 2004.

[14] *NIST Digital Library of Mathematical Functions*, https://dlmf.nist.gov/, Release 1.0.18 of 2018-03-27, F. W. J. Olver, A. B. Olde Daalhuis, D. W. Lozier, B. I. Schneider, R. F. Boisvert, C. W. Clark, B. R. Miller and B. V. Saunders, eds., URL: https://dlmf.nist.gov/.

[15] M. Dubiner, *Spectral methods on triangles and other domains*, Journal of Scientific Computing 6 (4 Dec. 1991), pp. 345–390, DOI: 10.1007/BF01060030.

[16] C. L. Epstein, L. Greengard, and A. Klöckner, *On the convergence of local expansions of layer potentials*, SIAM J. Numer. Anal. 51.5 (2013), pp. 2660–2679, DOI: 10.1137/120902859.

[17] L. Farina, *Evaluation of single layer potentials over curved surfaces*, SIAM Journal on Scientific Computing 23.1 (2001), pp. 81–91, DOI: 10.1137/S1064827599363393.

[18] C. Geuzaine and J. Remacle, *Gmsh: A 3-D finite element mesh generator with built-in pre- and post-processing facilities*, International Journal for Numerical Methods in Engineering 79.11 (2009), pp. 1309–1331, DOI: 10.1002/nme.2579.

[19] Z. Gimbutas and L. Greengard, *A fast and stable method for rotating spherical harmonic expansions*, Journal of Computational Physics 228.16 (Sept. 2009), pp. 5621–5627, ISSN: 0021-9991, DOI: 10.1016/j.jcp.2009.05.014.

[20] Z. Gimbutas and L. Greengard, *FMMLIB3D*, Accessed May 2018, URL: https://github.com/zgimbutas/fmmlib3d.

[21] J. Goodman, T. Y. Hou, and J. Lowengrub, *Convergence of the point vortex method for the 2-D Euler equations*, Communications on Pure and Applied Mathematics 43 (1990), pp. 415–430, DOI: 10.1002/cpa.3160430305.

[22] R. D. Graglia and G. Lombardi, *Machine Precision Evaluation of Singular and Nearly Singular Potential Integrals by Use of Gauss Quadrature Formulas for Rational Functions*, IEEE Transactions on Antennas and Propagation 56.4 (2008), pp. 981–998, DOI: 10.1109/TAP.2008.919181.

[23] L. Greengard, *The rapid evaluation of potential fields in particle systems*, ACM Distinguished Dissertations, MIT Press, Cambridge, MA, 1988, pp. xiv+91, ISBN: 0-262-07110-X.

[24] L. Greengard and V. Rokhlin, *A fast algorithm for particle simulations*, J. Comput. Phys. 73.2 (1987), pp. 325–348, DOI: 10.1016/0021-9991(87)90140-9.

[25] N. A. Gumerov and R. Duraiswami, *Comparison of the efficiency of translation operators used in the fast multipole method for the 3D Laplace equation*, tech. rep., University of Maryland Institute of Advanced Computer Studies, 2005.

[26] W. Hackbusch and S. A. Sauter, *On numerical cubatures of nearly singular surface integrals arising in BEM collocation*, Computing 52.2 (1994), pp. 139–159, DOI: 10.1007/BF02238073.

[27] S. Hao et al., *High-order accurate methods for Nyström discretization of integral equations on smooth curves in the plane*, en, Advances in Computational Mathematics 40.1 (Feb. 2014), pp. 245–272, ISSN: 1019-7168, 1572-9044, DOI: 10.1007/s10444-013-9306-3.

[28] D. J. Haroldsen and D. I. Meiron, *Numerical Calculation of Three-dimensional Interfacial Potential Flows using the Point Vortex Method*, Communications on Pure and Applied Mathematics 43 (1990), pp. 415–430, DOI: 10.1137/S1064827596302060.

[29] J. Helsing and R. Ojala, *Corner singularities for elliptic problems: integral equations, graded meshes, quadrature, and compressed inverse preconditioning*, Journal of Computational Physics 227 (2008), pp. 8820–8840, DOI: 10.1016/j.jcp.2008.06.022.

[30] S. Jarvenpää, M. Taskinen, and P. Yla-Oijala, *Singularity extraction technique for integral equation methods with higher order basis functions on plane triangles and tetrahedra*, International Journal for Numerical Methods in Engineering 58 (2003), pp. 1149–1165, DOI: 10.1002/nme.810.

[31] C. G. L. Johnson and L. R. Scott, *An Analysis of Quadrature Errors in Second-Kind Boundary Integral Methods*, SIAM Journal on Numerical Analysis 26.6 (1989), pp. 1356–1382, DOI: 10.1137/0726079.

[32] M. A. Khayat and D. R. Wilton, *Numerical Evaluation of Singular and Near-Singular Potential Integrals*, IEEE Transactions on Antennas and Propagation 53.10 (2005), pp. 3180–3190, DOI: 10.1109/TAP.2005.856342.

[33] L. af Klinteberg and A.-K. Tornberg, *A fast integral equation method for solid particles in viscous flow using quadrature by expansion*, Journal of Computational Physics 326 (2016), pp. 420–445, ISSN: 0021-9991, DOI: 10.1016/j.jcp.2016.09.006.

[34] L. af Klinteberg and A.-K. Tornberg, *Adaptive Quadrature by Expansion for Layer Potential Evaluation in Two Dimensions*, SIAM Journal on Scientific Computing 40.3 (2018), A1225–A1249, DOI: 10.1137/17M1121615.

[35] L. af Klinteberg and A.-K. Tornberg, *Error estimation for quadrature by expansion in layer potential evaluation*, Adv. Comput. Math. 43.1 (2017), pp. 195–234, DOI: 10.1007/s10444-016-9484-x.

[36] A. Klöckner, *A Repository of Sample Geometries*, Retrieved at revision a869fc3ad, URL: https://github.com/inducer/geometries/.

[37] A. Klöckner et al., *Quadrature by expansion: a new method for the evaluation of layer potentials*, J. Comput. Phys. 252 (2013), pp. 332–349, DOI: 10.1016/j.jcp.2013.06.027.

[38] T. Koornwinder, *Two-variable analogues of the classical orthogonal polynomials*, Theory and Applications of Special Functions (1975), ed. by R. Askey, pp. 435–495.

[39] R. Kress, *Linear integral equations*, Third, vol. 82, Applied Mathematical Sciences, Springer, New York, 2014, pp. xvi+412, ISBN: 978-1-4614-9592-5; 978-1-4614-9593-2, DOI: 10.1007/978-1-4614-9593-2.

[40] J. Lowengrub, M. Shelley, and B. Merriman, *High-order and efficient methods for the vorticity formulation of the Euler equations*, SIAM Journal on Scientific Computing 14 (1993), pp. 1107–1142, DOI: 10.1137/0914067.

[41] J. N. Lyness and L. M. Delves, *On Numerical Contour Integration Round a Closed Contour*, Math. Comp. 21.100 (Oct. 1967), pp. 561–577, DOI: 10.2307/2005000.

[42] A. Mayo, *Fast, High-order accurate solution of Laplace's equation on Irregular Regions*, Communications on Pure and Applied Mathematics 20 (1998), pp. 648–683, DOI: 10.1137/0906012.

[43] A. Mayo and A. Greenbaum, *Fourth order accurate evaluation of integrals in potential theory on exterior 3D regions*, Journal of Computational Physics 220.2 (2007), pp. 900–914, ISSN: 0021-9991, DOI: 10.1016/j.jcp.2006.05.042.

[44] D. Moore, *The Cost of Balancing Generalized Quadtrees*, in: *Proceedings of the Third ACM Symposium on Solid Modeling and Applications*, SMA '95, Salt Lake City, Utah, USA: ACM, 1995, pp. 305–312, ISBN: 0-89791-672-7, DOI: 10.1145/218013.218078.

[45] H. G. Petersen, E. R. Smith, and D. Soelvason, *Error estimates for the fast multipole method. II. The three-dimensional case*, en, Proc. R. Soc. Lond. A 448.1934 (Mar. 1995), pp. 401–418, ISSN: 0962-8444, 2053-9177, DOI: 10.1098/rspa.1995.0024.

[46] H. G. Petersen et al., *Error Estimates for the Fast Multipole Method. I. The Two-Dimensional Case*, Proceedings: Mathematical and Physical Sciences 448.1934 (1995), pp. 389–400, DOI: 10.1098/rspa.1995.0023.

[47] M. Rachh, *Integral equation methods for problems in electrostatics, elastostatics and viscous flow*, PhD thesis, New York University, 2015.

[48] M. Rachh, A. Klöckner, and M. O'Neil, *Fast algorithms for Quadrature by Expansion I: Globally valid expansions*, J. Comput. Phys. 345 (2017), pp. 706–731, DOI: 10.1016/j.jcp.2017.04.062.

[49] A. Rahimian, A. Barnett, and D. Zorin, *Ubiquitous evaluation of layer potentials using Quadrature by Kernel-Independent Expansion*, BIT Numerical Mathematics (Nov. 2017), ISSN: 1572-9125, DOI: 10.1007/s10543-017-0689-2.

[50] Y. Saad and M. H. Schultz, *GMRES: A Generalized Minimal Residual Algorithm for Solving Nonsymmetric Linear Systems*, SIAM Journal on Scientific and Statistical Computing 7.3 (July 1986), pp. 856–869, DOI: 10.1137/0907058.

[51] C. Schwab and W. L. Wendland, *On numerical cubatures of singular surface integrals in boundary element methods*, Numerische Mathematik 62 (1992), pp. 342–369, DOI: 10.1007/BF01396234.

[52] A. Sidi and M. Israeli, *Quadrature methods for periodic singular Fredholm integral equations*, Journal of Scientific Computing 3 (1988), pp. 201–231, DOI: 10.1007/BF01061258.

[53] M. Siegel and A.-K. Tornberg, *A local target specific quadrature by expansion method for evaluation of layer potentials in 3D*, Journal of Computational Physics 364 (2018), pp. 365–392, ISSN: 0021-9991, DOI: 10.1016/j.jcp.2018.03.006.

[54] J. Strain, *Locally-corrected multidimensional quadrature rules for singular functions*, SIAM Journal on Scientific Computing 16.4 (1995), pp. 992–1017, DOI: 10.1137/0916058.

[55] B. Vioreanu and V. Rokhlin, *Spectra of Multiplication Operators as a Numerical Tool*, SIAM Journal on Scientific Computing 36.1 (Jan. 2014), A267–A288, ISSN: 1064-8275, DOI: 10.1137/110860082.

[56] M. Wala, *GIGAQBX Accuracy Experiments*, Retrieved at revision 2314b0d83, URL: https://github.com/mattwala/gigaqbx-accuracy-experiments/.

[57] M. Wala and A. Klöckner, *A fast algorithm with error bounds for Quadrature by Expansion*, Journal of Computational Physics 374 (2018), pp. 135–162, DOI: 10.1016/j.jcp.2018.05.006.

[58] H. Xiao and Z. Gimbutas, *A numerical algorithm for the construction of efficient quadrature rules in two and higher dimensions*, Computers & Mathematics with Applications 59.2 (Jan. 2010), pp. 663–676, ISSN: 08981221, DOI: 10.1016/j.camwa.2009.10.027.

[59] L. Ying, G. Biros, and D. Zorin, *A high-order 3D boundary integral equation solver for elliptic PDEs in smooth domains*, Journal of Computational Physics 219 (2006), pp. 247–275, DOI: 10.1016/j.jcp.2006.03.021.

[60] H. Zhao et al., *A spectral boundary integral method for flowing blood cells*, Journal of Computational Physics 229.10 (2010), pp. 3726–3744, ISSN: 0021-9991, DOI: 10.1016/j.jcp.2010.01.024.