MATHEMATICS OF COMPUTATION
https://doi.org/10.1090/mcom/3422
Article electronically published on March 14, 2019

A WALK OUTSIDE SPHERES FOR THE FRACTIONAL LAPLACIAN: FIELDS AND FIRST EIGENVALUE

TONY SHARDLOW

Abstract. The solution of the exterior-value problem for the fractional Laplacian can be computed by a Walk Outside Spheres algorithm. This involves sampling α-stable Levy processes on their exit from maximally inscribed balls and sampling their occupation distribution. Kyprianou, Osojnik, and Shardlow (2018) developed this algorithm, providing a complexity analysis and an implementation, for approximating the solution at a single point in the domain. This paper shows how to efficiently sample the whole field by generating an approximation in $L^2(D)$ for a domain D. The method takes advantage of a hierarchy of triangular meshes and uses the multilevel Monte Carlo method for Hilbert space-valued quantities of interest. We derive complexity bounds in terms of the fractional parameter α and demonstrate that the method gives accurate results for two problems with exact solutions. Finally, we show how to couple the method with the variable-accuracy Arnoldi iteration to compute the smallest eigenvalue of the fractional Laplacian. A criteria is derived for the variable accuracy and a comparison is given with analytical results of Dyda (2012).

1. Introduction

Walk On Spheres is a classical method for solving the Poisson problem

$$-\Delta u = f \text{ on } D, \qquad u = g \text{ on } \partial D$$

for a domain $D \subset \mathbb{R}^d$, boundary data $g\colon \partial D \to \mathbb{R}$, and source term $f\colon D \to \mathbb{R}$. In [KOS18], the algorithm was extended to a Walk Outside Spheres (WOS) algorithm for the following problem for the fractional Laplacian: find $u\colon D \to \mathbb{R}$ such that

$$-\Delta^{\alpha/2} u = f \text{ on } D, \qquad u = g \text{ on } D^{\mathsf{c}} \tag{1.1}$$

for $\alpha \in (0,2)$ and exterior data $g\colon D^{\mathsf{c}} \to \mathbb{R}$. A review of the fractional Laplacian and its importance in the applied sciences is given by [LPG+18], which includes a description of WOS and other numerical approaches. The fractional Laplacian in (1.1) is the generator of the α-stable Levy process on $\mathbb{R}^d$. This observation leads to the following identity, on which the WOS algorithm is based:

$$u(\boldsymbol{x}) = \mathbb{E}\left[g(\boldsymbol{X}(\tau)) + \sum_{n=0}^{N^{\boldsymbol{x}}-1} r_n^{\alpha} \int_{B(\mathbf{0},1)} f(\boldsymbol{x}_n + r_n \boldsymbol{y})\, V_1(\boldsymbol{y})\, d\boldsymbol{y} \right], \qquad \boldsymbol{x} \in D. \tag{1.2}$$

Received by the editor April 25, 2018, and, in revised form, November 28, 2018.
2010 *Mathematics Subject Classification.* Primary 65C05, 34A08, 60J75, 34B09.
Key words and phrases. Fractional Laplacian, walk on spheres, Levy processes, Arnoldi algorithm, exterior-value problems, multilevel Monte Carlo, numerical solution of PDEs, eigenvalue problems.

To define the terms here, consider an α-stable Levy process $\boldsymbol{X}(t)$, $t \geq 0$, starting at $\boldsymbol{X}(0) = \boldsymbol{x}$, and let $\tau > 0$ be the first-exit time of $\boldsymbol{X}(t)$ from D. Define the discrete-time WOS process starting at $\boldsymbol{x}_0 = \boldsymbol{x}$ by $\boldsymbol{x}_n = \boldsymbol{X}(\tau_n)$, where $\tau_0 = 0$ and $\tau_n = \min\{t > \tau_{n-1} \colon \boldsymbol{X}(t) \notin B(\boldsymbol{x}_{n-1}, r_n)\}$ and where $B(\boldsymbol{x}_{n-1}, r_n)$ is the ball of maximum radius r_n contained in D and centred at $\boldsymbol{x}_{n-1}$. Further, $N^{\boldsymbol{x}}$ is the exit time for the WOS process $\boldsymbol{x}_n$ from D (so $\tau_{N^{\boldsymbol{x}}} = \tau$ and $\boldsymbol{X}(\tau) = \boldsymbol{x}_{N^{\boldsymbol{x}}}$) and V_1 is the expected occupation density of $\boldsymbol{X}(t)$ before exiting a unit ball after starting at the origin, and given by [BGR61]

(1.3)
$$V_1(\boldsymbol{y}) = 2^{-\alpha}\,\pi^{-d/2}\,\frac{\Gamma(d/2)}{\Gamma(\alpha/2)^2}\,\|\boldsymbol{y}\|^{\alpha-d}\int_0^{\|\boldsymbol{y}\|^{-2}-1}(u+1)^{-d/2}u^{\alpha/2-1}\,du, \quad \|\boldsymbol{y}\| \leq 1.$$

[KOS18] provides a complete Monte Carlo algorithm for approximating $u(\boldsymbol{x})$ for a single $\boldsymbol{x} \in D$. It works by sampling the WOS process $\boldsymbol{x}_n$ and the occupation density V_1, and computing $u(\boldsymbol{x})$ as a sample average. [KOS18] includes an implementation [OS17] and a study of the mean number of WOS steps required for the sampling of $\boldsymbol{X}(\tau)$. This paper addresses the problem of calculating the whole field $u \colon D \to \mathbb{R}$ as an element of $L^2(D)$ and the leading eigenvalue of the fractional Laplacian. In Section 2, we review the existence and regularity theory for (1.1) as well as the WOS algorithm from [KOS18]. Section 3 develops a more efficient algorithm for computing the solution $u \in L^2(D)$ using multilevel Monte Carlo. The key step is the coupling between WOS solves on nested triangular meshes, which leads to a bound on the complexity of the method. Numerical experiments show that the method gives accurate results for two problems with known solutions. In Section 4, we turn to the computation of the smallest eigenvalue of the fractional Laplacian by using the field solve as part of an Arnoldi iteration. To make the process more efficient, we show how the accuracy of the field solve should be varied during the Arnoldi iteration. Computations are given comparing the method to analytical results of [Dyd12]. The Julia code used for the computations is available for download.[1] In Section 5, we include a comparison of the proposed method to the adaptive finite element method of [AG17].

2. Review

We will make use of the following bounds on the unique solution to (1.1). We use $C^r(D)$ to denote the Banach space of r-times differentiable functions with the supremum norm on derivatives up to order $r \in \mathbb{N}$. For $s \in (0,1)$, $C^{r+s}(D)$ denotes the Hölder-continuous subspace of $u \in C^r(D)$ with norm $\|u\|_{C^r} + \sup_{|\boldsymbol{r}|=r}[\mathcal{D}^{\boldsymbol{r}}u]_s < \infty$ for
$$[u]_s \coloneqq \sup_{\boldsymbol{x},\boldsymbol{y}\in D}\frac{|u(\boldsymbol{x})-u(\boldsymbol{y})|}{\|\boldsymbol{x}-\boldsymbol{y}\|^s},$$
and $\mathcal{D}^{\boldsymbol{r}}$ is the partial derivative defined by the multi-index $\boldsymbol{r}$.

Theorem 2.1. *For a bounded Lipschitz domain D, suppose that $g \colon D^{\mathsf{c}} \to \mathbb{R}$ is continuous and satisfies*
$$\int_{D^{\mathsf{c}}}\frac{|g(\boldsymbol{x})|}{1+\|\boldsymbol{x}\|^{\alpha+d}}\,d\boldsymbol{x} < \infty,$$

[1] `https://github.com/tonyshardlow/julia_wos`

and that $f \in C^r(\overline{D})$ for some $r \in \mathbb{N}$ with $r > \alpha$. Then, there exists a unique continuous solution to (1.1)*. Further, if g is uniformly bounded, $\|u\|_{C^{r+\alpha}} \le C(\alpha, D)(\|g\|_\infty + \|f\|_{C^r})$.*

Proof. The existence and uniqueness is provided by [KOS18, Theorem 6.1]. This result also gives that $\|u\|_\infty \le C\left(\|f\|_\infty + \|g\|_\infty\right)$ for a constant C independent of u. Now apply [ROS16, Corollary 3.5] on any ball contained in D to gain the $C^{r+\alpha}$ regularity. □

2.1. **Point WOS.** The WOS method for approximating the solution $u(\boldsymbol{x})$ of (1.1) at a single $\boldsymbol{x} \in D$ is based on the probabilistic representation (1.2). The WOS process $\boldsymbol{x}_n$ can be efficiently sampled as follows: let $\mathrm{Beta}(\alpha, \beta)$ denote the family of Beta distributions, S^{d-1} denote the unit sphere in $\mathbb{R}^d$, and $\mathrm{U}(S)$ denote the uniform distribution on a bounded measurable subset S of $\mathbb{R}^d$. Choose independent samples $\beta_n \sim \mathrm{Beta}(\alpha/2, (2-\alpha)/2)$ and $\boldsymbol{\Theta}_n \sim \mathrm{U}(S^{d-1})$, and define the iteration

$$\boldsymbol{x}_{n+1} = \boldsymbol{x}_n + \boldsymbol{\Theta}_n \frac{1}{\sqrt{\beta_n}}\, d(\boldsymbol{x}_n), \qquad \boldsymbol{x}_0 = \boldsymbol{x}, \tag{2.1}$$

where $d(\boldsymbol{y})$ denotes the distance from $\boldsymbol{y}$ to the boundary of D. Define the exit time $N^{\boldsymbol{x}} = \min\{n \ge 0 \colon \boldsymbol{x}_n \notin D\}$. For independent samples $S_k \sim \mathrm{U}([0,1])$ and $\boldsymbol{\Phi}_k \sim \mathrm{U}(S^{d-1})$, let

$$v(\boldsymbol{x}) \coloneqq g(\boldsymbol{x}_{N^{\boldsymbol{x}}}) + \sum_{k=0}^{N^{\boldsymbol{x}}-1} F(\boldsymbol{x}_k; S_k, \boldsymbol{\Phi}_k), \tag{2.2}$$

where

$$F(\boldsymbol{x}_k; S_k, \boldsymbol{\Phi}_k) \coloneqq A_1\, d(\boldsymbol{x}_k)^\alpha \Big[f(\boldsymbol{x}_k + d(\boldsymbol{x}_k) S_k^{1/\alpha} \boldsymbol{\Phi}_k) - f(\boldsymbol{x}_k) \mathbb{P}\Big(\beta_k < 1 - S_k^{2/\alpha}\Big) + A_2\, f(\boldsymbol{x}_k) \Big]$$

for

$$A_1 \coloneqq \frac{1}{\alpha} 2^{-\alpha+1} \Gamma(\alpha/2)^{-2} B((2-\alpha)/2, \alpha/2), \qquad A_2 \coloneqq \int_0^1 \mathbb{P}\big(\beta_n < 1 - z^{2/\alpha}\big)\, dz.$$

To understand the relation to (1.2), note that $\boldsymbol{\Theta}/\sqrt{\beta}$ has the same law as the exit distribution of the α-stable Levy process from a unit ball for $\beta \sim \mathrm{Beta}(\alpha/2, (2-\alpha)/2)$ and $\boldsymbol{\Theta} \sim \mathrm{U}(S^{d-1})$. The expectation of $A_1\, d(\boldsymbol{x}_k)^\alpha \left(f(\boldsymbol{x}_k + d(\boldsymbol{x}_k)\, S_k^{1/\alpha}\, \boldsymbol{\Phi}_k)\right)$ equals

$$\mathbb{E}\left[r_n^\alpha \int_0^{\|\boldsymbol{y}\|^{-2}-1} f(\boldsymbol{x}_n + r_n \boldsymbol{y})\, V_1(\boldsymbol{y})\, d\boldsymbol{y} \right].$$

The adjustment in the definition of F is for computational efficiency: we may precompute A_2 (with quadrature methods), so that the term $f(\boldsymbol{x}_k + d(\boldsymbol{x}_k) S_k^{1/\alpha} \boldsymbol{\Phi}_k) - f(\boldsymbol{x}_k)$ has lower variance, thereby yielding more easily to Monte Carlo approximation. To approximate $u(\boldsymbol{x})$, we generate M iid samples $v^j(\boldsymbol{x})$ of $v(\boldsymbol{x})$ and evaluate the sample mean $\frac{1}{M}\sum_{j=1}^M v^j(\boldsymbol{x})$. The method is unbiased so that the sample mean converges to the exact solution $u(\boldsymbol{x})$ as $M \to \infty$ and the error is described via the central limit theorem. See [KOS18] for further details.

3. WOS field solve

Rather than evaluate $u(\boldsymbol{x})$ at a single point, we are interested in the whole field $u \in L^2(D)$. The basic idea is to generate point estimates for a set of points and use interpolation to define an approximate $u \in L^2(D)$. We examine the error when independent samples are used at each point, and then look at multilevel Monte Carlo as a method for improving efficiency by coupling point samples.

3.1. **WOS field error: Independent samples.** What errors result when we generate WOS approximations to $u(\boldsymbol{z}^j)$ for a set of points $\{\boldsymbol{z}^j\} \subset D$ and use an interpolant to approximate u in $L^2(D)$? We answer this question for $\{\boldsymbol{z}^j\}$ given by the vertices of a triangular mesh for D for the root-mean-square error (the $L^2(\Omega, L^2(D))$ norm).

We work in two dimensions ($d = 2$) and consider a shape-regular triangulation $\mathcal{T}^h$ of a domain $D \subset \mathbb{R}^2$ with mesh-width h. That is, D is the union of non-intersecting triangles $\tau \in \mathcal{T}^h$ and the mesh-width $h = \max h_\tau$ for h_τ equal to the longest edge length of $\tau \in \mathcal{T}^h$. Here, the shape-regular condition says there exists C such that $\text{Area}(\tau) \geq C h_\tau^d$ for all $\tau \in \mathcal{T}^h$. Let I_h denote the interpolation operator taking values at the vertices of the triangulation $\mathcal{T}^h$ to the piecewise-linear interpolant. Denote the vertices of $\mathcal{T}^h$ by $\boldsymbol{z}^1, \dots, \boldsymbol{z}^N$. The next lemma gives a bound on the $L^2(D)$ bias error in approximating u by an interpolant with vertex data given by a random vector $\boldsymbol{a}$.

Lemma 3.1. *Suppose that D is a bounded polygonal domain in $\mathbb{R}^2$. Suppose that $g\colon D^{\mathsf{c}} \to \mathbb{R}$ is bounded and that $f \in C^r(D)$ for $r = 2 - \alpha$. Let $\boldsymbol{a}$ be an $\mathbb{R}^N$-valued random variable such that $\mathbb{E}[\boldsymbol{a}] = [u(\boldsymbol{z}^1), \dots, u(\boldsymbol{z}^N)]$, where $\boldsymbol{z}^i$ are the vertices of a shape-regular triangulation with mesh-width h. Then, for some constant $C(\alpha, D) > 0$,*

$$\big\|u - \mathbb{E}\big[I_h \boldsymbol{a}\big]\big\|_{L^2(D)} \leq C(\alpha, D)\, h^2 \big\|g\big\|_\infty + \big\|f\big\|_{C^r}.$$

Proof. As $\mathbb{E}[\boldsymbol{a}] = [u(\boldsymbol{z}^1), \dots, u(\boldsymbol{z}^N)]$ and $\mathbb{E}[I_h \boldsymbol{a}] = I_h \mathbb{E}[\boldsymbol{a}]$, the functions u and $\mathbb{E}[I_h \boldsymbol{a}]$ agree at the vertices $\boldsymbol{z}^j$. The Bramble–Hilbert lemma (e.g., [BS08, Theorem 4.4.20]) together with the regularity given by Theorem 2.1 (for $r + \alpha = 2$ so that $u \in H^2(D)$) imply the result. □

Denoting $s_h \coloneqq I_h \boldsymbol{a}$, the lemma says that the expectation of the interpolant s_h is close to u in the $L^2(D)$ sense for a fine triangulation (small mesh-width h). We now look at the sample errors due to the WOS Monte Carlo method. Let $\|\cdot\|$ denote the Euclidean distance and $\|\cdot\|_{\text{F}}$ denote the matrix Frobenius norm.

Lemma 3.2. *Let a_i equal the sample average of M iid WOS samples of $v(\boldsymbol{z}^i)$ (defined in (2.2)), and let $\boldsymbol{s}_h \coloneqq [s_h(\boldsymbol{z}^1), \dots, s_h(\boldsymbol{z}^N)]$ for $s_h = I_h \boldsymbol{a}$. Suppose that $\text{Var}(v(\boldsymbol{z}_i)) \leq C$ for a constant C. Then*

$$\mathbb{E}\left[\frac{1}{N}\big\|\boldsymbol{s}_h - \mathbb{E}\big[\boldsymbol{s}_h\big]\big\|^2\right] \leq C\frac{1}{M}.$$

Proof. $\mathbb{E}\big[\|\boldsymbol{s}_h - \mathbb{E}[\boldsymbol{s}_h]\|^2\big] = \|\text{Cov}(\boldsymbol{a})\|_{\text{F}}$, where $\text{Cov}(\boldsymbol{a})$ is diagonal with entries $\text{Var}(v(\boldsymbol{z}^i))/M$. As there are N points $\boldsymbol{z}^i$ and $\text{Var}(v(\boldsymbol{z}^i)) \leq C$, the result holds. □

We combine the two estimates to give an L^2-bound in physical and probability space on the approximation error.

Theorem 3.3. *Suppose that D is a bounded polygonal domain in two dimensions. Suppose that $g\colon D^{\mathsf{c}} \to \mathbb{R}$ is bounded, that*

$$\int_{D^{\mathsf{c}}} \frac{g(\boldsymbol{x})^2}{1+\|\boldsymbol{x}\|^{\alpha+d}}\,d\boldsymbol{x} < \infty,$$

and that $f \in C^r(D)$ for some $r > \alpha$. There exists $C(\alpha, D) > 0$ such that

$$\big\|u - s_h\big\|_{L^2(\Omega,L^2(D))} \le C(\alpha,D)\bigg[h^2\big\|g\big\|_\infty + \big\|f\big\|_{C^r} + \frac{1}{\sqrt{M}}\bigg],$$

where s_h is the piecewise-linear interpolant on a shape-regular triangulation $\mathcal{T}^h$ of D of averages of M iid WOS samples at the vertices.

Proof. Write $u - s_h = (\mathbb{E}[s_h] - s_h) + (u - \mathbb{E}[s_h])$. The term $(\mathbb{E}[s_h] - s_h)$ is the deviation of the interpolant s_h from its mean. For linear interpolants,

$$\begin{aligned}\big\|\mathbb{E}\big[s_h\big] - s_h\big\|^2_{L^2(\Omega,L^2(D))} &= \sum_{\tau\in\mathcal{T}^h} \mathbb{E}\Big[\big\|\mathbb{E}\big[s_h\big] - s_h\big\|^2_{L^2(\tau)}\Big]\\ &\le \max_{\tau\in\mathcal{T}^h} \operatorname{Area}(\tau)\big\|\mathbb{E}\big[\boldsymbol{s}_h\big] - \boldsymbol{s}_h\big\|^2.\end{aligned}$$

The area $\operatorname{Area}(\tau)$ of each triangle is uniformly bounded by h^d, as h is the maximum edge length of any τ. Further, [KOS18, Corollary 6.5] implies that the variance of $v(\boldsymbol{z}^i)$ is bounded uniformly over $\boldsymbol{z}^i \in D$. Hence, for a constant C, Lemma 3.2 implies that

$$\big\|\mathbb{E}\big[s_h\big] - s_h\big\|^2_{L^2(\Omega,L^2(D))} \le C\,h^d\,\frac{N}{M}.$$

As the number of vertices $N \le 3\operatorname{Area}(D)/\min_\tau \operatorname{Area}(\tau) = \mathcal{O}\big(1/h^d\big)$ by the shape-regular condition, then, for a possibly larger constant C,

$$\mathbb{E}\Big[\big\|\mathbb{E}\big[s_h\big] - s_h\big\|^2_{L^2(D)}\Big] \le C\frac{1}{M}.$$

The second term $u - \mathbb{E}[s_h]$ is described by Lemma 3.1 and

$$\big\|u - \mathbb{E}\big[s_h\big]\big\|_{L^2(D)} \le C(\alpha,D)\,h^2\,\big\|g\big\|_\infty + \big\|f\big\|_{C^r}.$$

Together the two inequalities give that, for a possibly larger constant $C(\alpha, D)$,

$$\big\|u - s_h\big\|_{L^2(\Omega,L^2(D))} \le C(\alpha,D)\,h^2\,\big\|g\big\|_\infty + \big\|f\big\|_{C^r} + \frac{1}{\sqrt{M}}. \qquad \square$$

For accuracy ε, the required number of samples grows like $1/\varepsilon^2$. The required mesh-width $h = \mathcal{O}(\sqrt{\varepsilon})$ and the number of vertices grows like $1/h^d$. Hence, the required number of vertices grows like $\varepsilon^{-d/2}$ and the total work required is $\mathcal{O}(\varepsilon^{-(2+d/2)})$. Thus, even for $d = 2$, the WOS method with independent samples requires $\mathcal{O}(\varepsilon^{-3})$ work.

3.2. **Multilevel Monte Carlo.** The multilevel Monte Carlo (MLMC) method offers a practical way to reduce computational effort in Monte Carlo runs. The idea is to introduce nested triangular meshes with vertices $\boldsymbol{z}^i_\ell$ for $i = 1, \dots, N_\ell$ on level ℓ, and define $v_\ell \in L^2(D)$ based on WOS approximations at $\boldsymbol{z}^1_\ell, \dots, \boldsymbol{z}^{N_\ell}_\ell$. Then,

$$\mathbb{E}\big[v_\ell\big] = \mathbb{E}\big[v_1\big] + \sum_{j=1}^{\ell-1} \mathbb{E}\big[v_{j+1} - v_j\big].$$

We show how to choose the nested triangular meshes and coupling between samples so that the $v_{j+1} - v_j$ have small variances, which allows reduced computation time for a given accuracy level. Giles' complexity theorem describes the relationship between the work required and the coupling, number of samples, and errors.

Theorem 3.4 (MLMC complexity theorem)**.** *Let $u \in L^2(D)$ and let v_ℓ be $L^2(D)$-valued random variables for $\ell = \ell_0, \ell_0+1, \ell_0+2, \dots$. For constants c_i and a, β, γ, suppose that the following hold for $\ell = \ell_0, \ell_0+1, \dots$.*

Consistency condition: $\|\mathbb{E}[v_\ell] - u\|_{L^2(D)} \le c_1\, 2^{-a\ell}$.

Coupling condition: *We have $L^2(D)$-valued random variables $v_{\ell+1}^{\mathsf{f}}$ and v_ℓ^{c} (the superscripts denote fine and coarse) equal in distribution to v_ℓ that are coupled in the sense that*

$$V_\ell \coloneqq \mathbb{E}\Big[\big\|v_{\ell+1}^{\mathsf{f}} - v_\ell^{\mathsf{c}} - \mathbb{E}\big[v_{\ell+1}^{\mathsf{f}} - v_\ell^{\mathsf{c}}\big]\big\|_{L^2(D)}^2\Big] \le c_2\, 2^{-\beta\ell}.$$

The random variables v_ℓ^{f} and v_ℓ^{c} are independent.

Complexity condition: *The cost C_ℓ of computing one level-ℓ sample (i.e., of v_1 or v_ℓ^{c} or v_ℓ^{f}) is bounded by $c_3\, 2^{\gamma\ell}$.*

For tolerance ε, choose the smallest number of levels L such that $c_1\, 2^{-aL} \le \varepsilon/2$ and choose the number of samples on level ℓ as

$$M_\ell = \left\lceil 2\epsilon^{-2}\sqrt{C_\ell/V_\ell}\sum_{\ell=\ell_0}^{L}\sqrt{V_\ell C_\ell}\right\rceil, \qquad \ell = \ell_0, \ell_0+1, \dots, L-1.$$

For iid samples $(v_1^i, v_\ell^{\mathsf{f},i}, v_\ell^{\mathsf{c},i})$ of $(v_1, v_\ell^{\mathsf{f}}, v_\ell^{\mathsf{c}})$, define the level-$L$ MLMC approximation

$$v^{\mathrm{ML}} = \frac{1}{M_0}\sum_{i=1}^{M_0} v_{\ell_0}^i + \sum_{j=\ell_0}^{L-1}\frac{1}{M_j}\sum_{i=1}^{M_j}(v_{j+1}^{\mathsf{f},i} - v_j^{\mathsf{c},i}). \tag{3.1}$$

Suppose that $a > 1/2$ and $\beta < \gamma$. We achieve $\mathbb{E}\Big[\|u - v^{\mathrm{ML}}\|_{L^2(D)}^2\Big] \le \varepsilon^2$ with $\mathcal{O}(\epsilon^{-2-(\gamma-\beta)/a})$ work.

Proof. It is key for this application of the complexity theorem that the quantities of interest are Hilbert-space valued, which is described in [Gil15, Section 2.5]. We have stated the result only for the case $a > 1/2$ and $\beta < \gamma$, which is most relevant for our application. □

3.3. **Nested triangulations.** We set up the triangulations for MLMC WOS sampling. Consider a polygonal domain $D \subset \mathbb{R}^2$ and a triangulation $\{\tau_1^k\}$ of D with mesh-width h. Let $\boldsymbol{z}_1^j$ for $j = 1, \dots, N_1$ denote the vertices of the triangles. Define a refinement by dividing each triangle into four equal parts to define a new triangulation $\{\tau_2^k\}$ and set of vertices $\boldsymbol{z}_2^j$. Continue recursively to define the vertices $\boldsymbol{z}_\ell^j$ at level ℓ of the shape-regular triangulations $\{\tau_\ell^k\}$ with mesh-width $2^{1-k}h$. Denote the number of vertices on level ℓ by N_ℓ. Let $I_\ell \colon \mathbb{R}^{N_\ell} \to L^2(D)$ denote the piecewise-linear interpolation operator, which sends values at vertices $\boldsymbol{z}_\ell^j$ to the piecewise-linear interpolant on the level-ℓ triangulation.

In the following, we will be interested in evaluating the $L^2(D)$ norm of piecewise-linear functions. This can be done exactly by choosing a cubature rule $Q_\tau(\phi) \approx \int_\tau \phi(\boldsymbol{x})\,d\boldsymbol{x}$ of degree-of-precision two on the triangle τ, and computing

$$\big\|\phi\big\|_{L^2(D)} = \sum_k Q_{\tau_\ell^k}(\phi^2)^{1/2},$$

which is exact for piecewise-linear functions on the level-ℓ triangulation. In two dimensions, a suitable Q_τ is defined by

$$Q_\tau(\phi) = \frac{1}{3}\operatorname{Area}(\tau)\Big[\phi(\boldsymbol{m}_1) + \phi(\boldsymbol{m}_2) + \phi(\boldsymbol{m}_3)\Big],$$

where $\boldsymbol{m}_i$ are the midpoints of the edges of τ. This means, in particular, we can write, for a piecewise-linear function ϕ at level ℓ,

$$\big\|\phi\big\|_{L^2(D)} = \sum_k \operatorname{Area}(\tau_\ell^k) \sum_{i,j=1}^{3} a_{ij}\phi(\boldsymbol{z}_i)\phi(\boldsymbol{z}_j)\Bigg)^{1/2}, \tag{3.2}$$

where $\boldsymbol{z}_i$ are vertices of τ_ℓ^k and the coefficients $a_{ij} > 0$ (independent of level ℓ and triangle τ_ℓ^k).

To apply Theorem 3.4, we require values for a, β, γ. At each level, the number of points increases by a factor of 2^d (in dimension d) at most and $\gamma = d$. From Lemma 3.1, it is clear that $a = 2$. In the next two sections, we describe precisely the coupling and give a lower bound on the value of β, which describes the strength of the coupling.

3.4. **Coupling WOS processes.** In preparation for describing the coupling in the MLMC WOS field solver, we study the behaviour of two WOS processes $\boldsymbol{x}_n$ and $\boldsymbol{y}_n$ generated by the same random variables. That is,

$$\boldsymbol{x}_{n+1} = \boldsymbol{x}_n + d(\boldsymbol{x}_n)\,\boldsymbol{\Theta}_n\,\frac{1}{\sqrt{\beta_n}}, \qquad \boldsymbol{y}_{n+1} = \boldsymbol{y}_n + d(\boldsymbol{y}_n)\,\boldsymbol{\Theta}_n\,\frac{1}{\sqrt{\beta_n}} \tag{3.3}$$

for independent samples $\boldsymbol{\Theta}_n \sim \mathrm{U}(S^{d-1})$ and $\beta_n \sim \mathrm{Beta}(\alpha/2, 1-\alpha/2)$. The first result describes how the distance $r_n \coloneqq \|\boldsymbol{x}_n - \boldsymbol{y}_n\|$ depends on n and relies on the following assumption.

Assumption 3.5. Suppose, for some $\lambda < 1$ and $\mu > 0$, that

$$\mathbb{E}\left[\frac{\|\boldsymbol{x}_1 - \boldsymbol{y}_1\|}{\|\boldsymbol{x}_0 - \boldsymbol{y}_0\|}^{\mu\alpha} 1_{\{\boldsymbol{x}_1, \boldsymbol{y}_1 \in D\}} \,|\, \boldsymbol{x}_0, \boldsymbol{y}_0\right] \le \lambda \qquad \text{for all } \boldsymbol{x}_0, \boldsymbol{y}_0 \in D,$$

and for $\boldsymbol{x}_1 = \boldsymbol{x}_0 + d(\boldsymbol{x}_0)\,\boldsymbol{\Theta}/\sqrt{\beta}$ and $\boldsymbol{y}_1 = \boldsymbol{y}_0 + d(\boldsymbol{y}_0)\,\boldsymbol{\Theta}/\sqrt{\beta}$, where $\beta \sim \mathrm{Beta}(\alpha/2, 1-\alpha/2)$ and $\boldsymbol{\Theta} \sim \mathrm{U}(S^{d-1})$.

We verify the assumption holds for a large range of α in Appendix A, by computing the expectation numerically for $B = [0,1]^2$; it is seen to hold with $\mu = 1$ for $\alpha \le 1$ and with $\mu \approx 0.5$ for $\alpha \le 1.8$.

Lemma 3.6. *Consider a domain D in two dimensions. Let $\boldsymbol{x}_n, \boldsymbol{y}_n$ be coupled WOS processes (as defined by (3.3)) and write $\|\boldsymbol{x}_n - \boldsymbol{y}_n\| =: r_n$. If Assumption 3.5 holds for some $\mu \le 1$, there exists $C > 0$ such that, for $n \ge 1$,*

$$\mathbb{E}\big[r_{n+1}^{\mu\alpha}\, 1_{\{\boldsymbol{x}_k, \boldsymbol{y}_k \in D:\ k=1,\dots,n\}} \,|\, \boldsymbol{x}_0, \boldsymbol{y}_0\big] \le C\,\lambda^n\, r_0^{\mu\alpha}, \qquad \boldsymbol{x}_0, \boldsymbol{y}_0 \in D. \tag{3.4}$$

Proof. In two dimensions, we may write

$$(\boldsymbol{x}_{n+1}-\boldsymbol{y}_{n+1}) = (\boldsymbol{x}_n-\boldsymbol{y}_n) + \frac{d(\boldsymbol{x}_n)-d(\boldsymbol{y}_n)}{\sqrt{\beta_n}}\boldsymbol{\Theta}, \qquad \boldsymbol{\Theta} = \hat{\boldsymbol{r}}_n\cos\theta + \hat{\boldsymbol{r}}_n^{\perp}\sin\theta,$$

for some $\theta\in(-\pi,\pi]$, where $\hat{\boldsymbol{r}}$ is the unit vector in the direction $\boldsymbol{x}_n-\boldsymbol{y}_n$ and $\hat{\boldsymbol{r}}^{\perp}$ is a unit vector orthogonal to $\hat{\boldsymbol{r}}$. Note that $d(\boldsymbol{x}) = d(\boldsymbol{y}+(\boldsymbol{x}-\boldsymbol{y})) \le \|\boldsymbol{y}-\boldsymbol{z}+(\boldsymbol{x}-\boldsymbol{y})\| \le d(\boldsymbol{y})+\|\boldsymbol{x}-\boldsymbol{y}\|$, by choosing $\boldsymbol{z}\in\partial D$ such that $\|\boldsymbol{y}-\boldsymbol{z}\|=d(\boldsymbol{y})$. Then, $|d(\boldsymbol{x}_n)-d(\boldsymbol{y}_n)| \le r_n$ and

$$r_{n+1} \le r_n 1+\beta_n^{-1/2}\cos\theta^2+\beta_n^{-1}\sin^2\theta^{1/2} = r_n(1+2\beta_n^{-1/2}\cos\theta+\beta_n^{-1})^{1/2}.$$

The pdf of the beta distribution $\mathrm{Beta}(\alpha/2,(2-\alpha)/2)$ is

$$\begin{aligned} p(\beta) &= \frac{1}{B(\alpha/2,(2-\alpha)/2)}\,\beta^{\alpha/2-1}\,(1-\beta)^{-\alpha/2} \\ &= \frac{1}{\pi}\sin(\pi\alpha/2)\,\beta^{\alpha/2-1}\,(1-\beta)^{-\alpha/2}, \qquad \beta\in(0,1). \end{aligned}$$

The term $(1+2\beta_n^{-1/2}\cos\theta+\beta_n^{-1})^{\mu\alpha/2}p(\beta_n)$ grows like $\beta_n^{-\mu\alpha/2}\beta_n^{\alpha/2-1}$ as $\beta_n\downarrow 0$, and hence $\mathbb{E}\Big[(1+2\beta_n^{-1/2}\cos\theta+\beta_n^{-1})^{\mu\alpha/2}\Big]=\infty$ for $\mu>1$. This explains the restriction to $\mu\le 1$.

For $\mu\le 1$, let

$$C \coloneqq \mathbb{E}\big[(1+2\beta_n^{-1/2}\cos\theta+\beta_n^{-1})^{\mu\alpha/2}\,p(\beta_n)\big] < \infty.$$

Consequently,

$$\mathbb{E}\big[r_{n+1}^{\mu\alpha}\,|\,\boldsymbol{x}_n,\boldsymbol{y}_n\big] \le C\,r_n^{\mu\alpha}\,1_{\{\boldsymbol{x}_k,\boldsymbol{y}_k\in D:\ k=1,\dots,n\}}.$$

To estimate the right-hand side in terms of the initial separation r_0, we use Assumption 3.5, which provides a uniform bound on the $\mu\alpha$-moment of r_{n+1}/r_n given $\boldsymbol{x}_{n+1},\boldsymbol{y}_{n+1}\in D$, uniformly over any choice of $\boldsymbol{x}_n,\boldsymbol{y}_n\in D$ with separation r_n. This implies that

$$\mathbb{E}\big[r_n^{\mu\alpha}\,1_{\{\boldsymbol{x}_k,\boldsymbol{y}_k\in D:\ k=1,\dots,n\}}\big] \le \lambda\,\mathbb{E}\big[r_{n-1}^{\mu\alpha}\,1_{\{\boldsymbol{x}_k,\boldsymbol{y}_k\in D:\ k=1,\dots,n-1\}}\big].$$

Iterating this, we achieve the result. □

We make precise the difficulty with the case $\mu>1$.

Lemma 3.7. *Let D be the unit ball in two dimensions. For all $\mu>1$ and $K>1$, there exists $\boldsymbol{x}_0,\boldsymbol{y}_0\in D$ such that*

$$\mathbb{E}\big[r_1^{\mu\alpha}\,1_{\{\boldsymbol{x}_1,\boldsymbol{y}_1\in D\}}\big] \ge K\,r_0^{\mu\alpha}.$$

Proof. Again, write

$$(\boldsymbol{x}_1-\boldsymbol{y}_1) = (\boldsymbol{x}_0-\boldsymbol{y}_0) + \frac{d(\boldsymbol{x}_0)-d(\boldsymbol{y}_0)}{\sqrt{\beta_0}}\hat{\boldsymbol{r}}_0\cos\theta+\hat{\boldsymbol{r}}_0^{\perp}\sin\theta.$$

Choose $\boldsymbol{z}\in\partial D$ with inward-pointing normal $\boldsymbol{n}$ and let L be the distance from $\boldsymbol{z}$ to the far boundary in the direction $\boldsymbol{n}$. Choose $\delta>0$ and $\boldsymbol{x}_0,\boldsymbol{y}_0\in D$ with $r_0=\|\boldsymbol{x}_0-\boldsymbol{y}_0\|\le\delta$ on the line through $\boldsymbol{z}$ in direction $\boldsymbol{n}$, such that $d(\boldsymbol{x}_0)=\delta$ and $d(\boldsymbol{y}_0)=d(\boldsymbol{x}_0)+r_0$. By taking $\theta_0>0$ small, we can choose an interval $(-\theta_0,\theta_0)$ so all directions $\boldsymbol{\Theta}$ from $\boldsymbol{x}_0,\boldsymbol{y}_0$ that are angle $\theta\in(-\theta_0,\theta_0)$ from $\boldsymbol{n}$ are at least $L/2$ from the far boundary (uniformly in δ small). All jumps from $\boldsymbol{x}_0,\boldsymbol{y}_0$ in the

direction $\theta \in (-\theta_0, \theta_0)$ of length $L/2$ or less remain in D. When $d(\boldsymbol{x}_0)/\sqrt{\beta_0} \le L/2$, we have $\beta_0 \in ((2\delta/L)^2, 1)$. Then,

$$\mathbb{E}\big[r_1^{\mu\alpha}\big] \ge r_0^{\mu\alpha} \int_{(2\delta/L)^2}^1 1 + \frac{1}{\beta_0^{1/2}} \min_{\theta\in(-\theta_0,\theta_0)} \cos\theta^{\mu\alpha} p(\beta_0)\, d\beta_0.$$

The pdf of the beta distribution $\text{Beta}(\alpha/2, (2-\alpha)/2)$ is

$$p(\beta) = \frac{1}{B(\alpha/2,(2-\alpha)/2)}\, \beta^{\alpha/2-1}\,(1-\beta)^{-\alpha/2} = \frac{1}{\pi}\sin(\pi\alpha/2)\,\beta^{\alpha/2-1}\,(1-\beta)^{-\alpha/2}.$$

Consequently, for $\mu > 1$,

$$\int_0^1 \beta^{-\mu\alpha/2} p(\beta)\, d\beta = \infty.$$

Hence,

$$\int_{(2\delta/L)^2}^1 1 + \frac{1}{\beta_0^{1/2}} \min_{\theta\in(-\theta_0,\theta_0)} \cos\theta^{\mu\alpha} p(\beta_0)\, d\beta_0 \to \infty \qquad \text{as } \delta \downarrow 0.$$

Thus, for any $K>1$, we can choose δ, r_0 (and hence $\boldsymbol{x}_0, \boldsymbol{y}_0$) such that $\mathbb{E}\big[r_1^{\mu\alpha}\, 1_{\{\boldsymbol{x}_1,\boldsymbol{y}_1\in D\}}\big] \ge K r_0^{\alpha}$. □

Via Chebyshev's inequality, Assumption 3.5 implies the following bound on the probability coupled WOS paths are separated.

Corollary 3.8. *Under the assumptions of Lemma* 3.6*, there exist a constant* $C > 0$ *and* $\lambda \in (0,1)$ *such that, for any* $\boldsymbol{x}_0, \boldsymbol{y}_0 \in D$*,*

$$\mathbb{P}\big(r_n > \epsilon;\ \boldsymbol{x}_k, \boldsymbol{y}_k \in D \text{ for } k = 1,\dots,n-1 \,|\, \boldsymbol{x}_0, \boldsymbol{y}_0\big) \le C\, \frac{r_0^{\mu\alpha}}{\epsilon^{\mu\alpha}}\, \lambda^n.$$

Proof. As $r_n \le (1 + 1/\sqrt{\beta})\, r_{n-1}$, the event that $r_n > \epsilon$ given r_{n-1} is contained in $\sqrt{\beta} \le 1/(\epsilon/r_{n-1} - 1) \le 2(r_{n-1}/\epsilon)$ for $r_{n-1} < \epsilon/2$. The density $p(\beta)$ of $\text{Beta}(\alpha/2, 1-\alpha/2)$ satisfies $p(\beta) = \mathcal{O}(\beta^{\theta/2-1})$ as $\beta \downarrow 0$. Hence, for some $C > 0$, $p(\beta) \le C\beta^{\alpha/2-1}$ for $\beta < 1/2$. Hence, for larger constant C,

$$\begin{aligned}\mathbb{P}\left(\beta \le \frac{1}{(\epsilon/r_{n-1}-1)^2} \mid r_{n-1}\right) &\le \mathbb{P}\left(\beta \le 4\frac{r_{n-1}^2}{\epsilon^2} \mid r_{n-1}\right) \\ &\le C\, \frac{r_{n-1}^{\alpha}}{\epsilon^{\alpha}} \qquad \text{if } \frac{2r_{n-1}}{\varepsilon} \le \frac{1}{2}.\end{aligned}$$

This inequality also holds for $r_{n-1} \ge \epsilon/2$ by making sure $C \ge 2^{\alpha}$. As $\mu \le 1$, to complete the proof we apply Lemma 3.6. □

We require the following assumption regarding the WOS process near to the boundary of D, to control how WOS particles accumulate near the boundary. It is shown to hold for a specific domain in Appendix A by numerically evaluating integrals. For $B = [0,1]^2$, it is seen to hold for $t \approx 1$ for $\alpha \le 0.5$ and for $t \approx 0.9$ for $\alpha \le 1.8$.

Assumption 3.9. Suppose, for some $\lambda, t \in (0,1)$ and $A > 0$, that

$$\mathbb{E}\big[\Phi(\boldsymbol{x}_1) \,|\, \boldsymbol{x}_0\big] \le \lambda\, \Phi(\boldsymbol{x}_0),$$

where

$$\Phi(\boldsymbol{x}) \coloneqq 1_{\{\boldsymbol{x}\in D\}}\, \max\left\{A, \frac{1}{d(\boldsymbol{x})^t}\right\}. \tag{3.5}$$

The initial set of vertices has the following moment property with respect to $\Phi(\boldsymbol{x})$.

Lemma 3.10. *Let $\boldsymbol{z}_\ell^k$ for $i = 1, \dots, N_\ell$ be the vertices of the triangulation at level ℓ defined in Section 3.3. For the Φ in Assumption 3.9, there exists $C > 0$ independent of ℓ such that*

$$\frac{1}{N} \sum_{k=1}^{N_\ell} \Phi(\boldsymbol{z}_\ell^k) < C.$$

Proof. Let $\bar{d} = \sup\{\|\boldsymbol{x} - \boldsymbol{y}\| \colon \boldsymbol{x}, \boldsymbol{y} \in D\}$ denote the diameter of D. The vertices are distributed uniformly and the proportion of the N_ℓ vertices $\boldsymbol{z}_\ell^k$ such that $d(\boldsymbol{z}_\ell^k) \in ((j-1)\epsilon, j\epsilon)$ for $j = 1, \dots, \lceil \bar{d}/2\varepsilon \rceil$ is less than $C\epsilon$ for a constant C (that depends on the geometry of D; $C = 4/\bar{d}$ when D is a ball of diameter $\bar{d}$). Hence, for $t \in (0,1)$,

$$\frac{1}{N_\ell} \sum_{k=1}^{N_\ell} \frac{1}{d(\boldsymbol{z}_\ell^k)^t} \leq \sum_{j=1}^{\lceil \bar{d}/2\varepsilon \rceil} \frac{C\epsilon}{(j\epsilon)^t} \leq C \int_0^{\bar{d}/2} \frac{1}{x^t}\, dx = \frac{C}{1-t} \Big[x^{1-t} \Big]_0^{\bar{d}/2} = \frac{C}{1-t} \frac{\bar{d}}{2}^{1-t}.$$

□

3.5. **Main theorem on coupling.** The MLMC estimator v^{ML} defined in (3.1) is defined in terms of $v_\ell^{\mathsf{c}}, v_\ell^{\mathsf{f}}, v_{\ell_0}$. Key to the success of the estimator is the coupling between the fine-level estimator $v_{\ell+1}^{\mathsf{f}}$ and coarse-level estimator v_ℓ^{c}, which must be generated by the same random variables. To write this down precisely, consider independent random variables $\beta_n \sim \mathrm{Beta}(\alpha/2, 1-\alpha/2)$, $\boldsymbol{\Theta}_n, \boldsymbol{\Phi}_n \sim \mathrm{U}(S^{d-1})$, and $S_k \sim \mathrm{U}([0,1])$. If $\boldsymbol{x}_n$ is the WOS process starting at $\boldsymbol{x}_0 = \boldsymbol{x}$ generated by these inputs, define

$$V(\boldsymbol{x}) = g(\boldsymbol{x}_{N^{\boldsymbol{x}}}) + \sum_{k=0}^{N^{\boldsymbol{x}}-1} F(\boldsymbol{x}_k), \tag{3.6}$$

using $F(\boldsymbol{x}_k) = F(\boldsymbol{x}_k; S_k, \boldsymbol{\Phi}_k)$ (see (2.2)). The field V is not easily evaluated as it requires a WOS process for every $\boldsymbol{x}$ and we evaluate it only at the vertices of the triangulations. Let V_ℓ be independent copies of V. The fields $v_{\ell+1}^{\mathsf{f}}$ and v_ℓ^{c} are defined as linear interpolants using vertices of the relevant triangulation as initial data. That is,

$$v_{\ell+1}^{\mathsf{f}} \coloneqq I_{\ell+1}\{V_\ell(\boldsymbol{z}_{\ell+1}^k) \colon k = 1, \dots, N_{\ell+1}\}, \qquad v_\ell^{\mathsf{c}} \coloneqq I_\ell\{V_\ell(\boldsymbol{z}_\ell^k) \colon k = 1, \dots, N_\ell\}.$$

In this way, $v_{\ell+1}^{\mathsf{f}}$ and v_ℓ^{c} are given by the linear interpolant of the same copy of V, based on vertices of the level-$(\ell+1)$ or ℓ triangulation. Note that v_ℓ^{f} and $v_{\ell+1}^{\mathsf{f}}$ are independent (and similarly for the coarse versions).

We now give the main result on coupling between v_ℓ^{c} and $v_{\ell+1}^{\mathsf{f}}$.

Theorem 3.11. *Suppose that Assumptions 3.5 and 3.9 hold with exponents $t, \mu < 1$. Suppose that f and g are uniformly $\mu\alpha$-Hölder continuous with Hölder constant L. Then, the coupling condition holds for the MLMC complexity theorem (Theorem 3.4) with $\beta = \min\{\mu\alpha, \mu\alpha t/(t+\mu\alpha)\}$. In particular, there exists $C > 0$ such that*

$$\mathbb{E}\Big[\big\| v_{\ell+1}^{\mathsf{f}} - v_\ell^{\mathsf{c}} - \mathbb{E}\big[v_{\ell+1}^{\mathsf{f}} - v_\ell^{\mathsf{c}}\big] \big\|^2 \Big] \leq C \Big[L\, h_\ell^{\mu\alpha} + h_\ell^{\mu t\alpha/(t+\mu\alpha)} \Big]. \tag{3.7}$$

Proof. The constant C in this proof is a generic constant independent of ℓ and h_ℓ and changes from line to line. Following (3.2), the key observation is that

$$\mathbb{E}\Big[\big\|v_{\ell+1}^{\mathsf{f}} - v_\ell^{\mathsf{c}} - \mathbb{E}\big[v_{\ell+1}^{\mathsf{f}} - v_\ell^{\mathsf{c}}\big]\big\|_{L^2(D)}^2\Big] = \mathbb{E}\left[\sum_k \operatorname{Area}(\tau_k^{\ell+1}) \sum_{i,j=1}^{3} a_{ij}\, \delta(\boldsymbol{z}^i)\, \delta(\boldsymbol{z}^j)\right],$$

where $\boldsymbol{z}^i$ are the vertices of $\tau_k^{\ell+1}$ and $\delta = v_{\ell+1}^{\mathsf{f}} - v_\ell^{\mathsf{c}} - \mathbb{E}\big[v_{\ell+1}^{\mathsf{f}} - v_\ell^{\mathsf{c}}\big]$. As $ab \le \frac{1}{2}(a^2+b^2)$, there exists $C > 0$ such that

$$\begin{aligned}\mathbb{E}\Big[\big\|v_{\ell+1}^{\mathsf{f}} - v_\ell^{\mathsf{c}} - \mathbb{E}\big[v_{\ell+1}^{\mathsf{f}} - v_\ell^{\mathsf{c}}\big]\big\|_{L^2(D)}^2\Big] &\le C \max_k \operatorname{Area}(\tau_k^{\ell+1})\, \mathbb{E}\left[\sum_i \delta(\boldsymbol{z}^i)^2\right] \\ &\le C\, \frac{1}{N^\ell}\, \mathbb{E}\left[\sum_i \delta(\boldsymbol{z}^i)^2\right], \qquad (3.8)\end{aligned}$$

where the sum is over all vertices $\boldsymbol{z}^i$ of the fine triangulation $\tau_k^{\ell+1}$. If $\boldsymbol{z}^i$ is also a vertex in the coarse triangulation τ_k^ℓ, $\delta(\boldsymbol{z}^i) = 0$. Otherwise, $\boldsymbol{z}^i$ is the midpoint of an edge $\boldsymbol{y}_a \leftrightarrow \boldsymbol{y}_b$ of the coarse triangulation. Then, $v_\ell^{\mathsf{c}}(\boldsymbol{z}^i)$ is defined by linear interpolation and

$$\delta(\boldsymbol{z}^i) = v_{\ell+1}^{\mathsf{f}}(\boldsymbol{z}^i) - \frac{1}{2} v_\ell^{\mathsf{c}}(\boldsymbol{y}_a) + v_\ell^{\mathsf{c}}(\boldsymbol{y}_b) - \mathbb{E}\left[v_{\ell+1}^{\mathsf{f}}(\boldsymbol{z}^i) - \frac{1}{2} v_\ell^{\mathsf{c}}(\boldsymbol{y}_a) + v_\ell^{\mathsf{c}}(\boldsymbol{y}_b)\right].$$

As $(a+b)^2 \le 2a^2 + 2b^2$,

$$\begin{aligned}\mathbb{E}\big[\delta(\boldsymbol{z}^i)^2\big] &\le \mathbb{E}\left[\big(v_{\ell+1}^{\mathsf{f}}(\boldsymbol{z}^i) - \frac{1}{2} v_\ell^{\mathsf{c}}(\boldsymbol{y}_a) + v_\ell^{\mathsf{c}}(\boldsymbol{y}_b)\big)^2\right] \\ &\le \frac{1}{2}\mathbb{E}\big[(v_{\ell+1}^{\mathsf{f}}(\boldsymbol{z}^i) - v_{\ell+1}^{\mathsf{f}}(\boldsymbol{y}_a))^2\big] + \mathbb{E}\big[(v_{\ell+1}^{\mathsf{f}}(\boldsymbol{z}^i) - v_{\ell+1}^{\mathsf{f}}(\boldsymbol{y}_b))^2\big]. \qquad (3.9)\end{aligned}$$

In the last line, we use the fact that $v_{\ell+1}^{\mathsf{f}}$ and v_ℓ^{c} agree on the level-ℓ triangulation. Putting together (3.8) and (3.9), this means we establish the result if we show, for a constant C, that, for any $\boldsymbol{z}_\ell^k + \boldsymbol{r}_0 \in D$,

$$\frac{1}{N_\ell} \sum_{k=1}^{N_\ell} \mathbb{E}\Big[v_{\ell+1}^{\mathsf{f}}(\boldsymbol{z}_\ell^k) - v_{\ell+1}^{\mathsf{f}}(\boldsymbol{z}_\ell^k + \boldsymbol{r}_0)^2\Big] \le C\big\|\boldsymbol{r}_0\big\|^\beta. \qquad (3.10)$$

We start by estimating the contribution to $v_{\ell+1}^{\mathsf{f}}$ from the exterior function g for paths starting at $\boldsymbol{x}_0$ and $\boldsymbol{y}_0 = \boldsymbol{x}_0 + \boldsymbol{r}_0$, first ignoring the contribution from the right-hand side term f. At the end of the proof, we briefly discuss what changes need to be made to account for a non-zero f.

Let $N^{\boldsymbol{x}}$ denote the WOS exit time of a path starting from $\boldsymbol{x} \in D$. Denote coupled WOS paths by $\boldsymbol{x}_n$ and $\boldsymbol{y}_n$ with $\boldsymbol{y}_0 = \boldsymbol{x}_0 + \boldsymbol{r}_0$. Write

$$\begin{aligned}&\mathbb{E}\big[g(\boldsymbol{x}_{N^{\boldsymbol{x}}}) - g(\boldsymbol{y}_{N^{\boldsymbol{y}}})^2\big] \\ &= \mathbb{E}\left[\sum_{k\ge1} g(\boldsymbol{x}_k) - g(\boldsymbol{y}_k)^2 1_{\{k=N^{\boldsymbol{x}}=N^{\boldsymbol{y}}\}}\right] + \mathbb{E}\left[\sum_{k\ge1} g(\boldsymbol{x}_{N_{\boldsymbol{x}}}) - g(\boldsymbol{y}_{N_{\boldsymbol{y}}})^2 1_{\{k=N^{\boldsymbol{x}}<N^{\boldsymbol{y}}\}}\right] \\ &+ \mathbb{E}\left[\sum_{k\ge1} g(\boldsymbol{x}_{N_{\boldsymbol{x}}}) - g(\boldsymbol{y}_{N_{\boldsymbol{y}}})^2 1_{\{k=N^{\boldsymbol{y}}<N^{\boldsymbol{x}}\}}\right].\end{aligned}$$

For the first term, using the Hölder regularity of g, note that

$$\mathbb{E}\left[\sum_{k\geq 1} g(\boldsymbol{x}_k)-g(\boldsymbol{y}_k)^2 1_{\{k=N^{\boldsymbol{x}}=N^{\boldsymbol{y}}\}}\right] \leq \sum_{k\geq 1}\mathbb{E}\big[(g(\boldsymbol{x}_k)-g(\boldsymbol{y}_k))^2 1_{\{k=N^{\boldsymbol{x}}=N^{\boldsymbol{y}}\}}\big] \leq L(2\|g\|_\infty)^{2-\mu\alpha}\sum_{k\geq 1}\mathbb{E}\big[r_k^{\mu\alpha}1_{\{k=N^{\boldsymbol{x}}=N^{\boldsymbol{y}}\}}\big].$$

By Lemma 3.6,

$$\mathbb{E}\big[r_k^{\mu\alpha}1_{\{k=N^{\boldsymbol{x}}=N^{\boldsymbol{y}}\}}\big]\leq C\,\lambda^k\,r_0^{\mu\alpha}.$$

Consequently,

$$\mathbb{E}\left[\sum_{k=N^{\boldsymbol{x}}=N^{\boldsymbol{y}}} g(\boldsymbol{x}_k)-g(\boldsymbol{y}_k)^2 1_{\{N^{\boldsymbol{x}}=N^{\boldsymbol{y}}=k\}}\right]\leq L\,(2\|g\|_\infty)^{2-\alpha}\,r_0^{\mu\alpha}\sum_{k\geq 1}\lambda^k.$$

As $\lambda<1$, we find a $C>0$ such that

$$\mathbb{E}\left[\sum_{k\geq 1} g(\boldsymbol{x}_k)-g(\boldsymbol{y}_k)^2\,1_{\{k=N^{\boldsymbol{x}}=N^{\boldsymbol{y}}\}}\right]\leq C\,L\,r_0^{\mu\alpha}.$$

It remains to consider the case where the paths exit at different times. The second and third terms are equivalent, and we consider only

$$\sum_{k\geq 1}\mathbb{E}\big[g(N^{\boldsymbol{x}})-g(N^{\boldsymbol{y}})^2 1_{\{k=N^{\boldsymbol{y}}<N^{\boldsymbol{x}}\}}\big]\leq 4\|g\|_\infty^2\sum_{k\geq 1}\mathbb{P}\big(k=N^{\boldsymbol{y}}<N^{\boldsymbol{x}}\big).$$

Now,

(3.11)
$$\mathbb{P}\big(k=N^{\boldsymbol{y}}<N^{\boldsymbol{x}}\big)=\mathbb{P}\big(d(\boldsymbol{x}_k)>\varepsilon,\ k=N^{\boldsymbol{y}}<N^{\boldsymbol{x}}\big)+\mathbb{P}\big(d(\boldsymbol{x}_k)\leq\varepsilon,\ k=N^{\boldsymbol{y}}<N^{\boldsymbol{x}}\big).$$

For the first term, note that $\mathbb{P}(d(\boldsymbol{x}_k)>\varepsilon,\ k=N^{\boldsymbol{y}}<N^{\boldsymbol{x}})\leq\mathbb{P}(r_k>\varepsilon,\ k=N^{\boldsymbol{y}}<N^{\boldsymbol{x}})$, so, by Corollary 3.8,

$$\mathbb{P}\big(d(\boldsymbol{x}_k)>\varepsilon,\,k=N^{\boldsymbol{y}}<N^{\boldsymbol{x}}\big)\leq C\,\lambda^n\,\frac{r_0^{\mu\alpha}}{\epsilon^{\mu\alpha}}.$$

As $\lambda<1$,

$$\sum_{k\geq 1}\mathbb{P}\big(d(\boldsymbol{x}_k)>\varepsilon,\,k=N^{\boldsymbol{y}}<N^{\boldsymbol{x}}\big)\leq C\,\frac{\lambda}{1-\lambda}\,r_0^{\mu\alpha}\,\frac{1}{\varepsilon^{\mu\alpha}}.$$

For the second term in (3.11), use Assumption 3.9 to see that

$$\mathbb{E}\big[\Phi(\boldsymbol{x}_1)\,|\,\boldsymbol{x}_0\big]\leq\lambda\,\Phi(\boldsymbol{x}_0),\qquad \boldsymbol{x}_0\in D.$$

Iterating, we have

$$\mathbb{E}\big[\Phi(\boldsymbol{x}_n)\,1_{\{\boldsymbol{x}_1,\ldots,\boldsymbol{x}_n\in D\}}\,|\,\boldsymbol{x}_0\big]\leq\lambda^n\,\Phi(\boldsymbol{x}_0).\tag{3.12}$$

Hence,

$$\sum_{n\geq 1}\mathbb{E}\big[\Phi(\boldsymbol{x}_n)1_{\{\boldsymbol{x}_1,\ldots,\boldsymbol{x}_n\in D\}}\,|\,\boldsymbol{x}_0\big]<\frac{1}{1-\lambda}\,\Phi(\boldsymbol{x}_0).$$

Note that $\Phi(\boldsymbol{x})\geq 1/d(\boldsymbol{x})^t$ if $d(\boldsymbol{x})^t\leq 1/A$. Hence $\mathbb{P}(d(\boldsymbol{x}_n)<\varepsilon)\leq C\,\varepsilon^t\,\mathbb{E}[\Phi(\boldsymbol{x}_n)]$ for a constant C. Hence,

$$\sum_{k\geq 1}\mathbb{P}\big(d(\boldsymbol{x}_k)\leq\varepsilon,\ N^{\boldsymbol{x}}\geq k\big)\leq C\,\varepsilon^t\,\Phi(\boldsymbol{x}_0).$$

To match the two probabilities and choose ε in terms of r_0, put $r_0^{\mu\alpha}/\varepsilon^{\mu\alpha} = \varepsilon^t$ so that $\varepsilon = r_0^{\mu\alpha/(t+\mu\alpha)}$. Thus,

$$\sum_{k\geq 1} \mathbb{E}\big[g(N_{\boldsymbol{x}}) - g(N_{\boldsymbol{y}})^2 1_{\{N^{\boldsymbol{x}} > N^{\boldsymbol{y}} = k\}}\big] \leq C\, r_0^{\mu\, t\, \alpha/(t+\mu\alpha)}\, \Phi(\boldsymbol{x}_0).$$

Then, adding the contributions for all terms, we conclude that

$$\mathbb{E}\big[g(\boldsymbol{x}_{N^{\boldsymbol{x}}}) - g(\boldsymbol{y}_{N^{\boldsymbol{y}}})^2\big] \leq CL\, r_0^{\mu\alpha} + \big\|g\big\|_\infty^2\, r_0^{\mu\, t\, \alpha/(t+\mu\alpha)}\, \Phi(\boldsymbol{x}_0). \tag{3.13}$$

By Lemma 3.10, $\frac{1}{N_\ell}\sum_{k=1}^{N} \Phi(\boldsymbol{z}_\ell^k)$ is less than a C independent of ℓ. This implies (3.10), as we can average over initial vertices $\boldsymbol{x}_0 = \boldsymbol{z}_\ell^k$ for $k = 1, \ldots, N_\ell$. Hence, in the case $f = 0$,

$$\big\|v_{\ell+1}^{\mathsf{f}} - v_\ell^{\mathsf{c}} - \mathbb{E}\big[v_{\ell+1}^{\mathsf{f}} - v_\ell^{\mathsf{c}}\big]\big\|_{L^2(D)} \leq CL\, r_0^{\mu\alpha} + \big\|g\big\|_\infty^2\, r_0^{\mu t\alpha/(t+\mu\alpha)}.$$

We now discuss extending the argument to $f \neq 0$. As f is $\mu\alpha$-Hölder continuous, from Section 2.1, we see that F is $\mu\alpha$-Hölder continuous. Now,

$$\begin{aligned}
I_1 &:= \mathbb{E}\left[\sum_{k\geq 0} (F(\boldsymbol{x}_k) - F(\boldsymbol{y}_k)) 1_{\{k < \min\{N^{\boldsymbol{x}}, N^{\boldsymbol{y}}\}\}}{}^2\right] \\
&\leq \mathbb{E}\left[\sum_{k=1}^{\infty} 1_{\{k=\min\{N^{\boldsymbol{x}}, N^{\boldsymbol{y}}\}\}} \sum_{j=0}^{k-1} (F(\boldsymbol{x}_j) - F(\boldsymbol{y}_j))^2\right] \\
&\leq \mathbb{E}\left[\sum_{k=1}^{\infty} 1_{\{k=\min\{N^{\boldsymbol{x}}, N^{\boldsymbol{y}}\}\}} \sum_{j=0}^{k-1} L\,(2\big\|F\big\|_\infty)^{2-\mu\alpha}\, r_0^{\mu\alpha}\, \lambda^j\right] \\
&\leq L\,(2\big\|F\big\|_\infty)^{2-\mu\alpha}\, r_0^{\mu\alpha} \frac{1}{1-\lambda}.
\end{aligned}$$

Hence, for a constant C,

$$I_1 \leq C\, L\, r_0^{\mu\alpha}. \tag{3.14}$$

Similarly,

$$\mathbb{E}\left[\sum_{N^{\boldsymbol{y}} \leq k < N^{\boldsymbol{x}}} F(\boldsymbol{x}_k)^2\right] \leq C\, \big\|F\big\|_\infty^2\, r_0^{t\mu\alpha/(t+\mu\alpha)} \Phi(\boldsymbol{x}_0). \tag{3.15}$$

To put everything together, recall that

$$v(\boldsymbol{x}) = g(\boldsymbol{x}_{N^{\boldsymbol{x}}}) + \sum_{k=0}^{N^{\boldsymbol{x}}-1} F(\boldsymbol{x}_k).$$

Hence,

$$\begin{aligned}
\big|v_{\ell+1}^{\mathsf{f}}(\boldsymbol{x}) - v_{\ell+1}^{\mathsf{f}}(\boldsymbol{y})\big| \leq \big|g(\boldsymbol{x}_{N^{\boldsymbol{x}}}) - g(\boldsymbol{y}_{N^{\boldsymbol{y}}})\big| &+ \sum_{k=0}^{\min\{N^{\boldsymbol{x}}, N^{\boldsymbol{y}}\}-1} \big|F(\boldsymbol{x}_k) - F(\boldsymbol{y}_k)\big| \\
&+ \sum_{k=N^{\boldsymbol{x}}}^{N^{\boldsymbol{y}}-1} \big|F(\boldsymbol{y}_k)\big| + \sum_{k=N^{\boldsymbol{y}}}^{N^{\boldsymbol{x}}-1} \big|F(\boldsymbol{x}_k)\big|.
\end{aligned}$$

The required bound (3.10) follows from (3.12)–(3.15). □

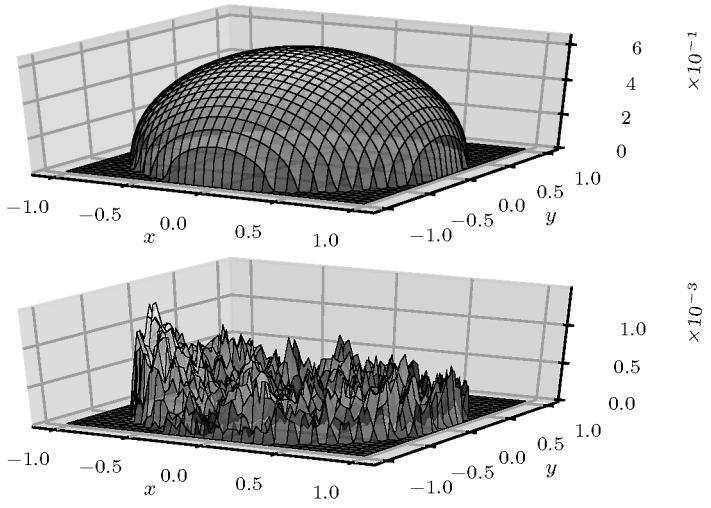

FIGURE 1. The upper plot shows the solution $u(\boldsymbol{x})$ defined in Example 3.13 for $\alpha = 1$. The lower plot shows the absolute errors for the computation with parameters with $\varepsilon = 1e-3$, $\ell_0 = 5$, $L = 7$. On the finest level, $h = 6.1 \times 10^{-5}$.

We now summarise the implications of the MLMC complexity theorem. We have designed an MLMC algorithm for solving the exterior-value problem for the fractional Laplacian with the following complexity.

Corollary 3.12. *Let the assumptions of Theorem* 3.11 *hold. Consider* $v^{\mathrm{ML}} \in L^2(D)$ *defined by* (3.1) *as an approximation to the solution of* u *of* (1.1)*. By choosing the number of samples* M_ℓ *and number of levels* L *as in Theorem* 3.4*, we achieve* $\|v^{\mathrm{ML}} - u\|_{L^2(\Omega, L^2(D))} \leq \varepsilon$ *with* $\mathcal{O}(\epsilon^{-3+\beta/2})$ *cost, where* $\beta = \min\{\alpha, t\mu\alpha/(t+\mu\alpha)\}$.

Proof. This is a consequence of Theorem 3.4 given values for a, β, γ. We have $a = 2$ from Lemma 3.1 and $\gamma = 2$ (due to the 2^d-factor increase in vertices with each level). For the coupling rate β, we have Theorem 3.11. □

This compares favourably to vanilla Monte Carlo, which has complexity ε^{-3}, reducing the computational cost by a factor $\varepsilon^{\beta/2}$.

3.6. **Numerical experiments.** We perform numerical experiments for $D = B(\mathbf{0}, 1)$. Define the triangulation $\{\tau_1^k\}$ of the square $[-1, 1]^2$ consisting of the four triangles given by drawing diagonals. Let $\{\tau_\ell^k\}$ be the level-ℓ triangulation given by recursively dividing triangles into four (using midpoints of edges). We apply WOS to the vertices within the unit ball $D = B(\mathbf{0}, 1)$ and define v^{ML} via multilevel Monte Carlo (see (3.1)), using piecewise-linear interpolants of sample averages to define v_ℓ on the triangulation $\{\tau_\ell^k\}$. The parameters for this algorithm are the coursest level ℓ_0, the finest level L, and the tolerance ε. As test cases, we recall two examples on the unit ball where exact solutions are known (e.g., [Buc16]).

Example 3.13. The problem with constant right-hand side and zero exterior condition on the unit ball,

$$(-\Delta u)^{\alpha/2} u = 1 \quad \text{on } D = B(\mathbf{0}, 1), \qquad u = 0 \quad \text{on } D^{\mathsf{c}},$$

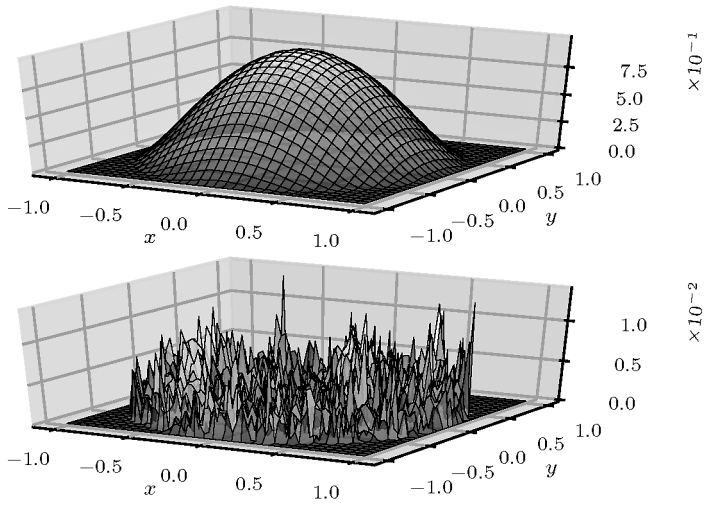

FIGURE 2. The upper plot shows the solution $u(\boldsymbol{x})$ of the problem in Example 3.14 for $\alpha = 1$. The lower plot shows the absolute errors for the computation with parameters with tol $= 1e-2$, $\ell_0 = 5$, and $L = 7$. On the finest level, $h = 6.1 \times 10^{-5}$.

has exact solution

$$u(\boldsymbol{x}) = \frac{\Gamma(1-\alpha/2)}{2^\alpha \Gamma(1+\alpha/2)(\alpha/2)B(\alpha/2, 1-\alpha/2)}(1 - \left\|\boldsymbol{x}\right\|^2)^{\alpha/2}, \qquad \boldsymbol{x} \in B(\mathbf{0}, 1).$$

The solution gives the mean first-exit time for an α-stable Levy process from $B(\mathbf{0}, 1)$. It is plotted in Figure 1 for $\alpha = 1$, along with the error from the WOS approximation with $\ell_0 = 5$, $L = 7$, and $\varepsilon = 1e-3$.

Example 3.14. The problem

$$(-\Delta u)^{\alpha/2} u = f \quad \text{on } D = B(\mathbf{0}, 1), \qquad u = 0 \quad \text{on } D^{\mathsf{c}}$$

for $f(\boldsymbol{x}) = (1 - (1+\alpha/2)\left\|\boldsymbol{x}\right\|^2),\ 2^\alpha\,\Gamma(2+\alpha/2)\,\Gamma(1+\alpha/2)$ has exact solution

$$u(\boldsymbol{x}) = (1 - \left\|\boldsymbol{x}\right\|^2)^{1+\alpha/2}. \tag{3.16}$$

For $\alpha = 1$, the solution is plotted in Figure 2 along with the error for $\varepsilon = 10^{-2}$, $\ell_0 = 5$, and $L = 7$.

The third example has less regular coefficients and there is no explicit exact solution.

Example 3.15. Let

$$g(\boldsymbol{x}) = \sin(\left\|\boldsymbol{x}\right\|^2), \qquad f(\boldsymbol{x}) = 2 + \left\|\boldsymbol{x}\right\|^2.$$

The numerical solution of (1.1) for $\alpha = 1$ with $D = B(\mathbf{0}, 1)$ is shown in Figure 3.

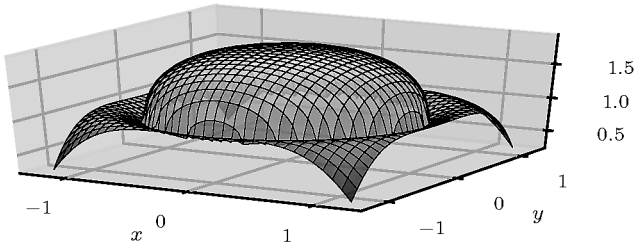

FIGURE 3. The solution $u(\boldsymbol{x})$ from Example 3.15 for $\alpha = 1$.

TABLE 1. Computed $L^2(D)$ relative and absolute errors and CPU execution time in seconds for a range of α.

Example 3.13				Example 3.14			
α	rel error	abs err	cpu time	α	rel error	abs err	cpu time
0.1	0.079	0.13	1.9	0.1	0.0093	0.0093	17.9
0.2	0.064	0.10	2.8	0.2	0.0057	0.0056	1.60
0.5	0.034	0.051	3.6	0.5	0.0057	0.0054	2.54
1.0	0.014	0.018	5.9	1.0	0.0058	0.0052	8.73
1.5	0.008	0.0090	17.0	1.5	0.0090	0.0075	17.79
1.8	0.00081	0.0085	36.6	1.8	0.0084	0.0068	40.47
1.9	0.0062	0.0065	60.0	1.9	0.0085	0.0068	68.37

Variance decay rates. We compute the variance decay rates for the coupling numerically, and calculate the variance at level ℓ defined by

$$V_\ell = \mathbb{E}\Big[\big\|v_{\ell+1}^{\mathsf{f}} - v_\ell^{\mathsf{c}} - \mathbb{E}\big[v_{\ell+1}^{\mathsf{f}}\big] - \mathbb{E}\big[v_\ell^{\mathsf{c}}\big]\big\|_{L^2(D)}^2\Big].$$

Figure 4 shows the variance V_ℓ against the mesh-width h_ℓ and indicates decay rates $\beta \approx 1$ for Example 3.13 independent of α, and rates $\beta = [0.25, 0.47, 0.67]$ for Example 3.14 for $\alpha = [0.5, 1.0, 1.5]$. Assuming $t = \mu = 1$, the theoretical rate is $\beta \approx \alpha/(1+\alpha)$, which yields $\beta \approx [0.33, 0.5, 0.6]$. The rate is a good prediction in Example 3.15. For Example 3.13, the right-hand side is constant and the exterior condition is zero, which is a much simpler scenario. Due to the constant right-hand side, zero coupling-error results on the interior and improved performance can be expected, as found with the variance coupling rates of $\beta \approx 1$.

As a demonstration of the relative efficiency of multilevel Monte Carlo and vanilla Monte Carlo methods, Figure 5 shows a plot of CPU time against tolerance. Multilevel Monte Carlo is clearly more efficient.

We can also compare the computed solution to the exact $u \in L^2(D)$ by computing an approximate $L^2(D)$ norm of the error. Table 1 shows errors for Examples 3.13 and 3.14 for seven values of α and the computational time for a tolerance 10^{-2}. Good accuracy results, especially for larger values of α, and all results are computed in less than a minute on a quad-core 3.2Ghz i5-6500 CPU with 8GB RAM (three cores are used in parallel for the WOS samples). Small values of α give poorer results. This is to be expected, due to the sharp gradient near the border (see (3.16)), which is explained by a larger constant $C(\alpha, D)$ in Theorem 3.3.

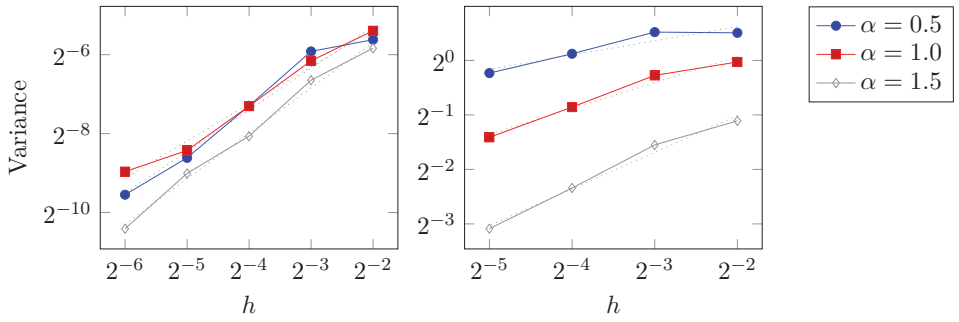

FIGURE 4. Plots of coupling variance V_ℓ against mesh-width h_ℓ for the multilevel Monte Carlo, with dotted lines showing best linear fits. The left-hand plot shows Example 3.13 with slopes $1.05, 0.94, 1.15$; the right-hand side shows Example 3.15 with slopes for $0.25, 0.47, 0.67$ for $\alpha = 0.5, 1.0, 1.5$, respectively.

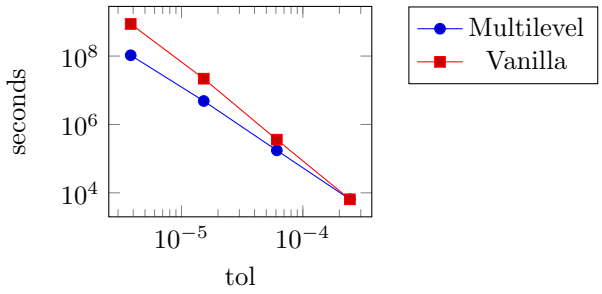

FIGURE 5. CPU execution time against tolerance ε for the multilevel Monte Carlo and vanilla (i.e., $\ell_0 = L$) Monte Carlo methods, based on Example 3.14. The multilevel parameter $\ell_0 = 6$ and $\varepsilon = 2^{-2L}$, where L is varied.

4. Leading eigenvalue using the Arnoldi algorithm

The Arnoldi algorithm is a well-known iterative method for computing the leading eigenvalues of a large sparse matrix, based on projecting the matrix onto a Krylov subspace. See [Arn51, TB97, BCS03, Saa11]. We show how to use Arnoldi to compute the smallest eigenvalue of the fractional Laplacian. That is, we seek the smallest $\lambda > 0$ such that

$$(-\Delta)^{\alpha/2} w = \lambda\, w \text{ on } D, \qquad w = 0 \text{ on } D^{\mathsf{c}}$$

for some non-trivial function $w\colon \mathbb{R}^d \to \mathbb{R}$. The first step is to discretise the fractional Laplacian and express the problem for a finite-dimensional linear operator. To do this, we consider a triangular mesh $\mathcal{T}$ with vertices $\boldsymbol{z}^1, \dots, \boldsymbol{z}^N$. We consider $\boldsymbol{v} \in \mathbb{R}^N$ corresponding to values of a function at vertices of the mesh and the solution operator for (1.1) with $g = 0$ and $f = I_h \boldsymbol{v}$, the piecewise-linear interpolant for $\boldsymbol{v}$ on the mesh $\mathcal{T}$. That is, for $\boldsymbol{v} \in \mathbb{R}^N$, let $u\colon D \to \mathbb{R}$ be the solution to $(-\Delta)^{\alpha/2} u = f$ on D, where $f = I_h \boldsymbol{v}$ and $u = 0$ on D^{c}. Then, we denote by A^{-1} the linear mapping from $\boldsymbol{v} \in \mathbb{R}^N$ to $\boldsymbol{u} = [u(\boldsymbol{z}^1), \dots, v(\boldsymbol{z}^N)] \in \mathbb{R}^N$. By applying

the Arnoldi algorithm to A^{-1}, we find the largest eigenvalue of A^{-1} and hence the smallest eigenvalue of A, which is the fractional Laplacian approximated on $\mathcal{T}$. We use this as our approximation to the smallest eigenvalue of the fractional Laplacian.

In practice, evaluating $A^{-1}\boldsymbol{v} = \boldsymbol{u}$ exactly is impossible and we will be using the WOS algorithm. This means we will be using the Arnoldi algorithm with inexact solves and exploiting the theory for variable-accuracy Arnoldi algorithms started by [BF00, Sim05] and developed further in [BMGS06, FS09]. It turns out that the accuracy of the solves can be reduced as the Arnoldi algorithm proceeds, without losing accuracy on the computed eigenvalue. This leads to significant speed ups. We develop the appropriate variable accuracy criterion for the WOS solve by establishing a criterion on the variance of the WOS solution necessary for a certain confidence interval in the computed eigenvalue.

We now describe the algorithm. Throughout, $\|\cdot\|$ denotes the Euclidean norm.

Inexact Arnoldi iteration. To determine the smallest eigenvalue of the fractional Laplacian $(-\Delta)^{\alpha/2}$ on a domain D, choose a triangulation $\mathcal{T}$ of the domain with N vertices and consider N-vectors of values at the vertices.

Algorithm 4.1.

(1) *Choose initial unit-vector $\boldsymbol{v}_1 \in \mathbb{R}^N$. Set $k = 1$.*
(2) *Evaluate $A^{-1}\boldsymbol{v}_k = \boldsymbol{u} + \boldsymbol{f}_k$, where $\boldsymbol{f}_k$ is the error resulting from the WOS solve.*
(3) *Gram–Schmidt: for $i = 1, \dots, k$, let $h_{ik} \coloneqq \boldsymbol{v}_i^{\mathsf{T}}\boldsymbol{u}$ and compute $\tilde{\boldsymbol{u}} \coloneqq \boldsymbol{u} - \sum_{i=1}^{k} h_{ik}\boldsymbol{v}_i$ to produce $\tilde{\boldsymbol{u}} \in \mathbb{R}^N$ orthogonal to the current Krylov space, given by span$\{\boldsymbol{v}_1, \dots, \boldsymbol{v}_k\}$ and coefficients h_{ik} for $i = 1, \dots, k$. Let $h_{k+1,k} \coloneqq \|\tilde{\boldsymbol{u}}\|$ and add the vector $\boldsymbol{v}_{k+1} \coloneqq \tilde{\boldsymbol{u}}/h_{k+1,k}$ to the Krylov space. Let V_k be the matrix with columns $\boldsymbol{v}_1, \dots, \boldsymbol{v}_k$.*
(4) *Compute eigenvectors $\boldsymbol{w}_j$ and eigenvalues θ_j (known as Ritz values) for the leading $k \times k$ submatrix H_k of the upper Hessenberg matrix (h_{ij}). The largest eigenpair $(\boldsymbol{w}, \theta)$ defines an approximate eigenvector for A^{-1} by $V_k\boldsymbol{w}$ and approximate eigenvalue θ^{-1}.*
(5) *Increase k and repeat.*

For finite-dimensional problems with exact solves ($\boldsymbol{f}_k = \boldsymbol{0}$ for all k), the algorithm is expected to converge as $k \to \infty$ to the leading eigenpair of A^{-1}. In our case, the algorithm introduces errors at several stages: first, we represent the WOS solutions on a triangular mesh and the WOS algorithm must evaluate the right-hand side function everywhere on the domain D. We use a piecewise-linear interpolant and this leads to an approximation error. Additionally, there is a Monte Carlo error on $\boldsymbol{u}$ due to the finite number of samples. We assume the error due to linear interpolation is negligible compared to Monte Carlo error; as the size of this error is quadratic in the mesh-width h_ℓ, this can be achieved by choosing the triangulation fine enough. For the Monte Carlo error, we develop a theory for the resulting error in the computed eigenvalue and a practical criteria for the tolerance for the WOS solve.

4.1. **Choosing the WOS tolerance.** To analyse the error due to the WOS solve at step k, we write

$$A^{-1}\boldsymbol{v}_k = \boldsymbol{u} + \boldsymbol{f}_k = \sum_{i=1}^{k} h_{ik}\boldsymbol{v}_k + h_{k+1,k}\boldsymbol{v}_{k+1} + \boldsymbol{f}_k,$$

where $\boldsymbol{f}_k$ represents the error due to the kth WOS solve. The right-hand side is given by the representation of $\boldsymbol{u}$ in the Krylov subspace (using the entries in the Hessenberg matrix H_k). We determine a criterion for relating the accuracy in the WOS solve (the size of $\boldsymbol{f}_k$) in terms of the desired eigenvalue accuracy. Stacking the expressions above, we have the well-known Arnoldi relation

$$A^{-1}V_m = V_m H_m + \boldsymbol{v}_{m+1} h_{m+1} \boldsymbol{e}_m^{\mathsf{T}} + F_m,$$

where $\boldsymbol{e}_m$ is the mth standard basis vector in $\mathbb{R}^N$ and F_m is the matrix of column vectors $[\boldsymbol{f}_1, \ldots, \boldsymbol{f}_m]$.

In the case of zero error $\boldsymbol{f}_k = \boldsymbol{0}$ for $k = 1, \ldots, m$, the size of the eigenvalue residual $\boldsymbol{r}_m = A^{-1}V_m\boldsymbol{w} - \theta^{-1}V_m\boldsymbol{w}$ is given by $\|\boldsymbol{r}_m\| = h_{m+1,m}|\boldsymbol{w}^{\mathsf{T}}\boldsymbol{e}_m|$, which follows easily from the Arnoldi relation

$$A^{-1}V_m = V_m H_m + \boldsymbol{v}_{m+1} h_{m+1} \boldsymbol{e}_m^{\mathsf{T}}$$

by taking the inner product with the eigenvector $\boldsymbol{w}$ of H_m. For the case of inexact solves, the residual itself is not readily available and the quantity $\|\boldsymbol{r}_m\| \coloneqq h_{m+1,m}|\boldsymbol{w}^{\mathsf{T}}\boldsymbol{e}_m|$ serves as a computationally convenient proxy.

For the next theorem, we will need the following non-degeneracy assumption [Sim05, (3.3)–(3.4)]. The notation $\sigma_{\min}(A)$ refers to the smallest singular value of A.

Assumption 4.2. For $k \leq m$, there exists an eigenpair $(\theta^{(k-1)}, \boldsymbol{w}^{(k-1)})$ of H_{k-1} sufficiently close to an eigenpair $(\theta, \boldsymbol{w})$ of H_m in the sense that

$$\left\|\boldsymbol{r}_{k-1}\right\| \leq \delta_{m,k-1}^2 \frac{1}{4\|\boldsymbol{s}_m\|}, \qquad \delta_{m,k} \coloneqq \sigma_{\min}(H_m - \theta^{(k)} I),$$

where $\boldsymbol{s}_m^{\mathsf{T}} \coloneqq [(\boldsymbol{w}^{(k-1)})^{\mathsf{T}}, \boldsymbol{0}^{\mathsf{T}}]H_m - \theta^{(k-1)}[(\boldsymbol{w}^{(k-1)})^{\mathsf{T}}, \boldsymbol{0}^{\mathsf{T}}] \in \mathbb{R}^m$, and

$$\left|\theta^{(k-1)} - \theta_j\right| > 2\frac{\|\boldsymbol{s}_m\| \, \|\boldsymbol{r}_{k-1}\|}{\delta_{m,k-1}} \text{ for all eigenvalues } \theta_j \neq \theta \text{ of } H_m.$$

In the following theorem, we develop the accuracy criterion for the WOS solves. We use the spectral gap $\delta^{(k-1)}$ for H_{k-1}: let $\Lambda(H_k)$ denote the set of eigenvalues of H_k and

$$\delta^{(k-1)} \coloneqq \min_{\theta \in \Lambda(H_{k-1}) - \theta^{(k-1)}} \left|\theta^{(k-1)} - \theta\right|, \tag{4.1}$$

where $\theta^{(k-1)}$ is the leading eigenvalue of H_{k-1}. The quantity depends on the spectrum of H_{k-1}, which is easily computable as k is generally small. We also use the σ-algebra $\mathcal{F}_m$ generated by the random variables used in the WOS solves up to step m, so that the residual $\boldsymbol{r}_m$ is $\mathcal{F}_m$-measurable. We will use the conditional expectation $\mathbb{E}[\cdot \,|\, \mathcal{F}_k]$ to average over input random variables (WOS samples from the field solves) for $m > k$.

Theorem 4.3. *Let $(\theta, \boldsymbol{w})$ be an eigenpair of the mth Hessenberg matrix H_m of an inexact Arnoldi iteration as described in Algorithm* 4.1 *such that*

$$\left\|\boldsymbol{r}_m\right\| \coloneqq h_{m+1,m}\left|\boldsymbol{e}_m^{\mathsf{T}}\boldsymbol{w}\right| \leq \text{tol}. \tag{4.2}$$

Suppose that the eigenvalue θ is simple and the eigenvector is normalised, $\|\boldsymbol{w}\| = 1$. Fix $\varepsilon > 0$ and suppose that the WOS error vector $\boldsymbol{f}_k$ at step k satisfies

$$\mathbb{E}\Big[\big\|\boldsymbol{f}_k\big\|^2 |\mathcal{F}_{k-1}\Big]^{1/2} \leq \frac{\varepsilon}{m} \times \begin{cases} \dfrac{\delta^{(k-1)}}{\|\boldsymbol{r}_{k-1}\|} & \text{if } k > 1 \text{ and Assumption 4.2 holds,} \\ 1 & \text{otherwise.} \end{cases}$$

Then, $V_m\boldsymbol{w}$ is an approximate eigenvector of A with eigenvalue θ^{-1}, in the sense that

$$\mathbb{P}\big(\big\|A^{-1}V_m\boldsymbol{w} - \theta^{-1}V_m\boldsymbol{w}\big\| > 2\,\mathrm{tol}\big) \leq \frac{\varepsilon^2}{\mathrm{tol}^2}\,\mathbb{E}\bigg[\frac{1}{m}\big\|\boldsymbol{\alpha}\big\|^2\bigg], \tag{4.3}$$

where $\boldsymbol{\alpha}$ has entries α_k satisfying

$$\big|\alpha_k\big| \leq 2\,\frac{\delta^{(k-1)}}{\delta_{m,k-1}} \text{ if Assumption 4.2 holds,} \quad \alpha_k = 1 \text{ otherwise.} \tag{4.4}$$

In particular, there exists an eigenvalue μ of A such that

$$\mathbb{P}\big(\big|\mu^{-1} - \theta^{-1}\big| > 2\,\mathrm{tol}\big) \leq \frac{\varepsilon^2}{\mathrm{tol}^2}\mathbb{E}\bigg[\frac{1}{m}\big\|\boldsymbol{\alpha}\big\|^2\bigg].$$

Proof. For inexact solves, the Arnoldi relationship is

$$A^{-1}V_m - F_m = V_mH_m + h_{m+1,m}\boldsymbol{v}_{m+1}\boldsymbol{e}_m^{\mathsf{T}},$$

where the columns of F_m are $\boldsymbol{f}_1, \ldots, \boldsymbol{f}_m$. If $(\theta, \boldsymbol{w})$ is an eigenpair of H_m, then

$$A^{-1}V_m\boldsymbol{w} - F_m\boldsymbol{w} = \theta\, V_m\boldsymbol{w} + h_{m+1,m}\boldsymbol{v}_{m+1}(\boldsymbol{e}_m^{\mathsf{T}}\boldsymbol{w}).$$

Hence, we have the eigenvalue residual

$$\big\|A^{-1}V_m\boldsymbol{w} - \theta V_m\boldsymbol{w}\big\| = h_{m+1,m}\big|\boldsymbol{e}_m^{\mathsf{T}}\boldsymbol{w}\big| + \big\|F_m\boldsymbol{w}\big\|. \tag{4.5}$$

The first residual corresponds to (4.2) and is monitored during the Arnoldi iteration. The second term, which we refer to as the extra residual, is due to the inexact solves and we analyse that now. Following [Sim05], it can be written as

$$F_m\boldsymbol{w} = \sum_{k=1}^m \boldsymbol{f}_k(\boldsymbol{e}_k^{\mathsf{T}}\boldsymbol{w}) = \sum_{k=1}^m \frac{\boldsymbol{f}_k\|\boldsymbol{r}_{k-1}\|}{\delta^{(k-1)}}(\boldsymbol{e}_k^{\mathsf{T}}\boldsymbol{w})\frac{\delta^{(k-1)}}{\|\boldsymbol{r}_{k-1}\|},$$

where $\delta^{(k-1)}$ is the spectral gap defined in (4.1). Write

$$F_m\boldsymbol{w} = \sum_{k=1}^m \alpha_k\,\beta_{k-1}\,\boldsymbol{f}_k$$

for $\beta_{k-1} = \|\boldsymbol{r}_{k-1}\|/\delta^{(k-1)}$ and $\alpha_k = \|\boldsymbol{r}_{k-1}\|^{-1}(\boldsymbol{e}_k^{\mathsf{T}}\boldsymbol{w})\,\delta^{(k-1)}$ if Assumption 4.2 holds, and $\beta_{k-1} = \alpha_k = 1$ otherwise. Here the subscripts for α_k and β_k indicate that they are $\mathcal{F}_k$-measurable. The Cauchy–Schwarz inequality provides that

$$\mathbb{E}\Big[\big\|F_m\boldsymbol{w}\big\|^2\Big] \leq \mathbb{E}\bigg[\sum_{k=1}^m m\,\beta_{k-1}^2\big\|\boldsymbol{f}_k\big\|^2\bigg]\,\mathbb{E}\bigg[\frac{1}{m}\sum_{k=1}^m \alpha_k^2\bigg].$$

By Chebyshev's inequality,

$$\mathbb{P}\big(\big\|F_m\boldsymbol{w}\big\| > \mathrm{tol}\big) \leq \frac{1}{\mathrm{tol}^2}\bigg\|\sum_{k=1}^m m\,\mathbb{E}\Big[\beta_{k-1}^2\big\|\boldsymbol{f}_k\big\|^2\Big]\bigg\|\,\mathbb{E}\bigg[\frac{1}{m}\sum_{k=1}^m \alpha_k^2\bigg].$$

If Assumption 4.2 holds, the condition on the WOS error $\boldsymbol{f}_k$ implies that

$$\mathbb{E}\Big[\big\|\boldsymbol{f}_k\big\|^2 \,|\, \mathcal{F}_{k-1}\Big] \le \frac{(\varepsilon\,\delta^{(k-1)})^2}{(m\,\|\boldsymbol{r}_{k-1}\|)^2} = \frac{\varepsilon^2}{m\,\beta_{k-1}^2}.$$

Otherwise, $\mathbb{E}\big[\|\boldsymbol{f}_k\|^2 \,|\, \mathcal{F}_{k-1}\big] \le \varepsilon^2/m^2$. Hence,

$$\mathbb{P}\big(\big\|F_m\boldsymbol{w}\big\| > \mathrm{tol}\big) \le \frac{\varepsilon^2}{\mathrm{tol}^2}\,\mathbb{E}\bigg[\frac{1}{m}\big\|\boldsymbol{\alpha}\big\|^2\bigg].$$

Using (4.5), this implies (4.3) under the condition (4.2).

Under Assumption 4.2, [Sim05, Proposition 2.2] implies that

$$\big|\boldsymbol{e}_k^{\mathsf{T}}\boldsymbol{w}\big| \le 2\frac{\|\boldsymbol{r}_{k-1}\|}{\delta_{m,k-1}}.$$

Here, it is important to note that $\delta_{m,k-1}$ is not $\mathcal{F}_k$-measurable in general and depends on the full Arnoldi run. Then,

$$\big|\alpha_k\big| \le 2\frac{\delta^{(k-1)}}{\delta_{m,k-1}},$$

which is (4.4). The final statement is a consequence of the Bauer–Fike theorem for normal matrices [GVL13], which says that the error in the eigenvalue is bounded by the eigenvalue residual. □

This theorem suggests a practical way of choosing the WOS tolerance at step k in dependence on a given eigenvalue-residual tolerance tol, parameter ε, and computed residual $\|\boldsymbol{r}_k\|$. For $B > 1$, to achieve

$$\mathbb{P}\big(\text{eigenvalue residual} > 2\,\mathrm{tol}\big) \le \frac{1}{B^2},$$

we assume that $\mathbb{E}\big[\frac{1}{m}\|\boldsymbol{\alpha}\|^2\big] \approx 1$, and choose $\varepsilon = \mathrm{tol}/B$ for $B > 1$. Then,

$$\mathbb{P}\big(\text{eigenvalue residual} > 2\,\mathrm{tol}\big) \le \frac{\varepsilon^2}{\mathrm{tol}^2}\mathbb{E}\bigg[\frac{1}{m}\big\|\boldsymbol{\alpha}\big\|^2\bigg] \le \frac{1}{B^2}.$$

For the kth WOS solve, from Theorem 4.3, we demand that

$$\mathbb{E}\Big[\big\|\boldsymbol{f}_k\big\|^2 | \mathcal{F}_{k-1}\Big]^{1/2} \le \frac{1}{B\,m}\,\mathrm{tol}\,\frac{\delta^{(k-1)}}{\|\boldsymbol{r}_{k-1}\|}.$$

This leads to a relaxed accuracy condition for the WOS calculation if the computed eigenvalue residual $\|\boldsymbol{r}_{k-1}\|$ is smaller than the spectral gap $\delta^{(k-1)}$. It is simple to implement and requires computing the spectrum of the $k \times k$ Hessenberg matrix H_k at each step (to determine $\delta^{(k-1)}$) and monitoring the variance in the WOS Monte Carlo calculation.

4.2. **Leading eigenvalue on the unit ball.** Dyda [Dyd12] provides upper and lower bounds on the leading eigenvalue for the fractional Laplacian on the unit ball, gained by rigorous analytical methods. We compare this to the eigenvalues computed by Algorithm 4.1. Table 2 shows the results of a computation with $\mathrm{tol} = 0.01$, $B = 3$, and WOS multilevel parameters $\ell_0 = 3$, and $L = 7$ with five Arnoldi iterations. The Arnoldi iteration produces estimates that are very close to Dyda's upper bound (the error relative to the upper bound are in the range 3×10^{-4} to 10^{-3}). The computations take between fifteen and thirty minutes for a Julia implementation on a quad-core 3.2Ghz i5-6500 CPU with 8GB RAM (three

TABLE 2. [Dyd12] provides the lower and upper bounds on the smallest eigenvalue of the fractional Laplacian on $B(\mathbf{0}, 1)$ shown. These are compared to the result of an Arnoldi computation with five iterations and the variable accuracy methods described.

	Dyda eigenvalue		Arnoldi			error relative
α	lower bound	upper bound	eigenvalue	residual	seconds	to upper bound
0.1	1.04874	1.05096	1.05187	0.0672	885	8×10^{-4}
0.2	1.10549	1.10993	1.11103	0.0896	933	9×10^{-4}
0.5	1.3313	1.34374	1.34464	0.8147	838	6×10^{-4}
1.0	1.96349	2.00612	2.00689	0.095	907	3×10^{-4}
1.5	3.13569	3.27594	3.27789	0.017	1008	5×10^{-4}
1.8	4.28394	4.56719	4.57029	0.014	1268	6×10^{-4}
1.9	4.77496	5.13213	5.13936	0.00075	1711	1×10^{-3}

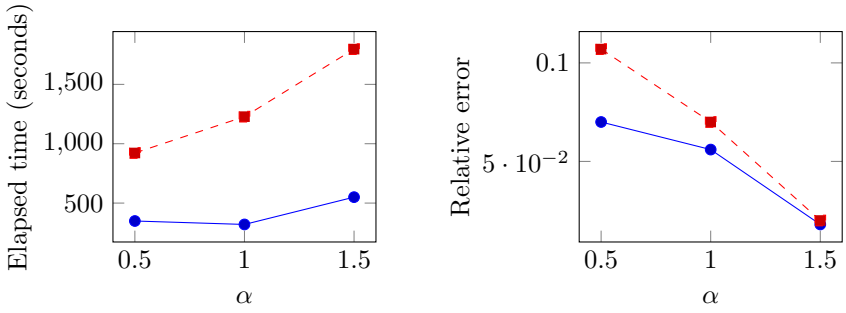

FIGURE 6. Comparison of run times with (solid) and without (dashed) variable accuracy, using tol $= 1e-2$, $L = 7$, $\ell_0 = 5$, $B = 3$. The left-hand plot shows run times in seconds against α. The right-hand plot shows the relative accuracy (relative to the upper bound of Dyda's interval).

cores are used in parallel for the WOS samples). The run times are compared in Figure 6 to the Arnoldi algorithm without variable accuracy, taking the same tolerance in both runs. The variable accuracy algorithm is twice as fast, for the same level of accuracy.

5. CONCLUSION

We have discussed Walk Outside Spheres for simulating the whole field rather than a point value of the solution $u\colon D \to \mathbb{R}$ of (1.1), extending the algorithm of [KOS18]. By using the multilevel Monte Carlo algorithm, we improved substantially on a naive method based on independent sampling at vertices. The improvement is demonstrated analytically (an improvement in the complexity for accuracy ε of factor ε^β, where $\beta \geq \min\{\alpha, t\mu\alpha/(t+\mu\alpha)\}$). The parameters μ, t lie in the range $(0, 1)$ and are generally unknown; numerical examination of the relevant assumptions shows that t can be chosen close to one and μ depends on α (μ may be chosen close to one for small α and must be reduced substantially as $\alpha \to 2$).

The improvement is also demonstrated numerically by looking at two problems with exact solutions and a third problem where variance estimates were made. Numerical experiments show the complexity bounds are pessimistic, even for $t = \mu = 1$. This is because the assumptions and analysis are based on a pair of coupled WOS processes, while the $L^2(D)$ error depends on an average over a WOS path for each initial vertex.

There are several deterministic approaches to the numerical approximation of the exterior-value problem for the fractional Laplacian [LPG+18]. We compare our results to the complexity and error analysis for the adaptive finite-element method (AFEM) of [AG17, AG18]. By using a posterioi error estimates and sparse approximations to the dense linear systems resulting from the global coupling in the fractional Laplacian, it converges in the $L^2(D)$ sense in two dimensions with $\mathcal{O}(n^{-1/2-\alpha/4})$, where n is the number of degrees of freedom. Due to the use of a clustering technique in assembly of the linear system, the solution can be computed in $\mathcal{O}\big(n\log^4(n)\big)$ operations. In terms of an $L^2(D)$ accuracy of ϵ, this method requires $\mathcal{O}(\epsilon^{-2+2\alpha/(2+\alpha)-\delta})$ operations on a polygonal domain in two dimensions (for any $\delta > 0$). In contrast, the WOS field method on a uniform mesh takes $\mathcal{O}(\epsilon^{-3+\beta/2})$ operations to achieve a root-mean-square $L^2(D)$ accuracy of ϵ, where β is given in Corollary 3.12.

Though AFEM is one order of magnitude faster, the field WOS method has significant potential. First, the WOS method is trivial to parallelise and this means the constant associated to the complexity analysis can be made small, given sufficient parallel resources. This can be very significant in practical situations. Second, the WOS method developed here is a classical Monte Carlo method, depending on a sequence of independent samples from certain probability distributions. The complexity of such methods depends on the $1/\sqrt{M}$ sampling error from M independent samples, even for multilevel Monte Carlo methods. Very often, the complexity is substantially improved by employing quasi-Monte Carlo techniques. In situations where the sample depends on an infinite number of random variables and the importance of these random variables decays suitably rapidly, the $\mathcal{O}(1/\sqrt{M})$ error can be replaced by $\mathcal{O}(1/M^{1-\delta})$ for $\delta > 0$ [HW00, GKN+11]. This sort of analysis has not been completed for the field WOS method and is beyond the scope of the present paper. It is worth noting that each $L^2(D)$ sample depends on a finite but unbounded number of random variables (depending on the number of steps for the WOS path to exit the domain), but only one set of random variables is used for each vertex in the mesh (due to coupling) and the importance of these random variables decays geometrically (as the exit time is geometrically distributed [KOS18]). It will be a subject of future research to develop a quasi-Monte Carlo-based field WOS method with improved complexity. If the $1/\sqrt{M}$ is replaced by $1/M$ in the complexity analysis, the overall complexity improves by a factor ϵ. There is a potential also to improve the method by using an adaptive mesh. At this point, the method becomes competitive with AFEM. If these issues can be overcome, a larger class of particle methods for solving PDEs can be coupled in the same way to give efficient field solvers.

Finally, we used the WOS algorithm to compute the leading eigenvalue of the fractional Laplacian. We developed a criterion for accuracy at the kth Arnoldi iteration based on the spectral gap of the Hessenberg matrix and the residual. The method is shown to give accurate results by comparing to analytical results

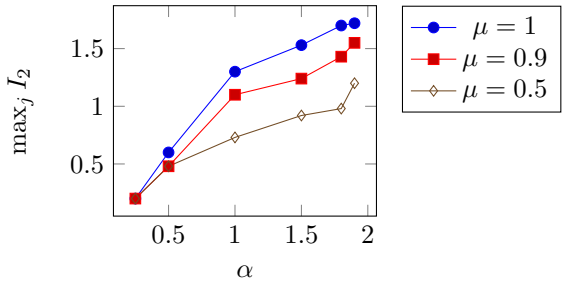

FIGURE 7. The maximum of $\max_{j=1,\dots,20} I_2(\boldsymbol{x}_0^j, \boldsymbol{y}_0^j)$ based on computing the expectation with 10^6 samples for $\mu = 0.5, 0.9, 1.0$.

of [Dyd12]. This algorithm can be developed further to get more leading eigenvalues and to incorporate the implicitly restarted Arnoldi method. Note the shift strategy (i.e., solving for $(A - sI)\boldsymbol{x} = \boldsymbol{b}$ for a shift value s) commonly used in eigenvalue solvers is not easy to apply with WOS; the implicitly restarted Arnoldi algorithm allows shifts to be introduced implicitly and only solves for the fractional Laplacian are required. Here again variable accuracy strategies are available [FS09] and could be adapted to the randomly inexact inner solves. This is the subject of future work.

APPENDIX A. ON ASSUMPTIONS 3.5 AND 3.9

We verify that Assumptions 3.5 and 3.9 are realistic, by computing the relevant quantities numerically for a square domain. For some $t, \lambda, \mu \in (0, 1)$ and $A > 0$, we wish to establish that

$$I_1(\boldsymbol{x}_0) := \mathbb{E}\left[\frac{\Phi(\boldsymbol{x}_1)}{\Phi(\boldsymbol{x}_0)} 1_{\{\boldsymbol{x}_1 \in D\}} \,|\, \boldsymbol{x}_0\right] \leq \lambda, \qquad \Phi(\boldsymbol{x}) := \max\left\{A, \frac{1}{d(\boldsymbol{x})^t}\right\}$$

and

$$I_2(\boldsymbol{x}_0, \boldsymbol{y}_0) := \mathbb{E}\left[\frac{\|\boldsymbol{x}_1 - \boldsymbol{y}_1\|}{\|\boldsymbol{x}_0 - \boldsymbol{y}_0\|}^{\mu\alpha} 1_{\{\boldsymbol{x}_1, \boldsymbol{y}_1 \in D\}}\right] \leq \lambda \qquad \text{for all } \boldsymbol{x}_0, \boldsymbol{y}_0 \in D.$$

If we replace A by $A/\bar{d}^t$ for the diameter $\bar{d}$ of D, it is clear both conditions are invariant to rescaling the domain. Hence, we focus on a box $D = [0, 1] \times [0, 1]$. We compute the values by a simple Monte Carlo method using $\boldsymbol{x}_1 = \boldsymbol{x}_0 + \boldsymbol{\Theta}\, d(\boldsymbol{x}_0)/\sqrt{\beta}$ for $\boldsymbol{\Theta} \sim \mathrm{U}(S^1)$ and $\beta \sim \mathrm{Beta}(\alpha/2, 1 - \alpha/2)$.

The expression for I_2 involves one parameter μ but must be checked for every pair of $\boldsymbol{x}_0, \boldsymbol{y}_0 \in D$. We draw twenty $\boldsymbol{x}_0^j, \boldsymbol{y}_0^j$ independently from $\mathrm{U}(D)$ and evaluate $I_2(\boldsymbol{x}_0^j, \boldsymbol{y}_0^j)$ using 10^6 samples of $\boldsymbol{x}_1, \boldsymbol{y}_1$ and record the maximum value. We show a plot of $\max_j I_2(\boldsymbol{x}_0^j, \boldsymbol{y}_0^j)$ against α for $\mu = 1, 0.9, 0.5$ in Figure 7. The condition $I_2(\boldsymbol{x}_0, \boldsymbol{y}_0) \leq \lambda < 1$ is satisfied with $\mu = 1$ for α small. For larger values of α, μ must be reduced; for example, for $\alpha = 1$, μ must be reduced to $\mu \approx 0.5$.

The expression for I_1 involves two parameters A, t, and we expect the appropriate choice of t to depend on α and for $A \to \infty$ as $t \to 1$. Again, we evaluate $I_1(\boldsymbol{x}_0)$ based on 10^6 samples and Table 3 shows the result of $\max_j I_1(\boldsymbol{x}_0^j)$ for different choice of A and t. By choice of A, we can always achieve $\max_j I_1(\boldsymbol{x}_0^j) \leq \lambda < 1$ for $t = 0.9$.

TABLE 3. The computation of $\max_{j=1,\dots,20} I_1(\boldsymbol{x}_0^j)$ is based on computing the expectation with 10^6 samples. The ⋮ symbol indicates the value is given by the entry above.

α	A	t	$\max_j I_1(\boldsymbol{x}_0^j)$
0.5	10^4	1.0	0.53
1.0	⋮	⋮	1.03
1.5	⋮	0.9	3.55
1.5	⋮	0.5	0.86
1.8	⋮	⋮	0.93
1.5	10^5	0.9	0.85
1.8	10^6	⋮	0.94

REFERENCES

[AG17] M. Ainsworth and C. Glusa, *Aspects of an adaptive finite element method for the fractional Laplacian: a priori and a posteriori error estimates, efficient implementation and multigrid solver*, Comput. Methods Appl. Mech. Engrg. **327** (2017), 4–35, DOI 10.1016/j.cma.2017.08.019. MR3725761

[AG18] M. Ainsworth and C. Glusa, *Towards an efficient finite element method for the integral fractional Laplacian on polygonal domains*, Contemporary Computational Mathematics - A Celebration of the 80th Birthday of Ian Sloan, Springer, 2018, pp. 17–57.

[Arn51] W. E. Arnoldi, *The principle of minimized iteration in the solution of the matrix eigenvalue problem*, Quart. Appl. Math. **9** (1951), 17–29, DOI 10.1090/qam/42792. MR0042792

[BMGS06] J. Berns-Müller, I. G. Graham, and A. Spence, *Inexact inverse iteration for symmetric matrices*, Linear Algebra Appl. **416** (2006), no. 2-3, 389–413, DOI 10.1016/j.laa.2005.11.019. MR2242738

[BGR61] R. M. Blumenthal, R. K. Getoor, and D. B. Ray, *On the distribution of first hits for the symmetric stable processes.*, Trans. Amer. Math. Soc. **99** (1961), 540–554, DOI 10.2307/1993561. MR0126885

[BF00] A. Bouras and V. Frayssé, *A relaxation strategy for inexact matrix-vector products for Krylov methods*, CERFACS TR0PA000015, European Centre for Research and Advanced Training in Scientific Computation (2000).

[BS08] S. C. Brenner and L. R. Scott, *The Mathematical Theory of Finite Element Methods*, 3rd ed., Texts in Applied Mathematics, vol. 15, Springer, New York, 2008. MR2373954

[Buc16] C. Bucur, *Some observations on the Green function for the ball in the fractional Laplace framework*, Commun. Pure Appl. Anal. **15** (2016), no. 2, 657–699, DOI 10.3934/cpaa.2016.15.657. MR3461641

[BCS03] J. F. Blowey, A. W. Craig, and T. Shardlow (eds.), *Frontiers in Numerical Analysis*, Universitext, Springer-Verlag, Berlin, 2003. Papers from the 10th LMS-EPSRC Numerical Analysis Summer School held at the University of Durham, Durham, July 7–19, 2002. MR2006323

[Dyd12] B. Dyda, *Fractional calculus for power functions and eigenvalues of the fractional Laplacian*, Fract. Calc. Appl. Anal. **15** (2012), no. 4, 536–555, DOI 10.2478/s13540-012-0038-8. MR2974318

[FS09] M. A. Freitag and A. Spence, *Shift-invert Arnoldi's method with preconditioned iterative solves*, SIAM J. Matrix Anal. Appl. **31** (2009), no. 3, 942–969, DOI 10.1137/080716281. MR2538660

[Gil15] M. B. Giles, *Multilevel Monte Carlo methods*, Acta Numer. **24** (2015), 259–328, DOI 10.1017/S096249291500001X. MR3349310

[GVL13] G. H. Golub and C. F. Van Loan, *Matrix Computations*, 4th ed., Johns Hopkins University Press, Baltimore, 2013.

[GKN+11] I. G. Graham, F. Y. Kuo, D. Nuyens, R. Scheichl, and I. H. Sloan, *Quasi-Monte Carlo methods for elliptic PDEs with random coefficients and applications*, J. Comput. Phys. **230** (2011), no. 10, 3668–3694, DOI 10.1016/j.jcp.2011.01.023. MR2783812

[HW00] F. J. Hickernell and H. Woźniakowski, *Integration and approximation in arbitrary dimensions*, Adv. Comput. Math. **12** (2000), no. 1, 25–58, DOI 10.1023/A:1018948631251. High dimensional integration. MR1758946

[KOS18] A. E. Kyprianou, A. Osojnik, and T. Shardlow, *Unbiased 'walk-on-spheres' Monte Carlo methods for the fractional Laplacian*, IMA J. Numer. Anal. **38** (2018), no. 3, 1550–1578, DOI 10.1093/imanum/drx042. MR3829169

[LPG+18] A. Lischke, G. Pang, M. Gulian, F. Song, C. Glusa, X. Zheng, Z. Mao, W. Cai, M. M. Meerschaert, M. Ainsworth, and G. E. Karniadakis, *What is the fractional Laplacian?* (January 2018), arXiv: 1801.09767.

[OS17] A. Osojnik and T. Shardlow, *MATLAB code for walk on spheres for fractional Laplacian*, posted on 2017, DOI 10.5281/zenodo.220877.

[ROS16] X. Ros-Oton and J. Serra, *Regularity theory for general stable operators*, J. Differential Equations **260** (2016), no. 12, 8675–8715, DOI 10.1016/j.jde.2016.02.033. MR3482695

[Saa11] Y. Saad, *Numerical Methods for Large Eigenvalue Problems*, Classics in Applied Mathematics, vol. 66, Society for Industrial and Applied Mathematics (SIAM), Philadelphia, PA, 2011. Revised edition of the 1992 original [1177405]. MR3396212

[Sim05] V. Simoncini, *Variable accuracy of matrix-vector products in projection methods for eigencomputation*, SIAM J. Numer. Anal. **43** (2005), no. 3, 1155–1174, DOI 10.1137/040605333. MR2177800

[TB97] L. N. Trefethen and D. Bau III, *Numerical Linear Algebra*, Society for Industrial and Applied Mathematics (SIAM), Philadelphia, PA, 1997. MR1444820

Department of Mathematical Sciences, University of Bath, Bath BA2 7AY, United Kingdom

Email address: t.shardlow@bath.ac.uk