



Multilevel Monte Carlo for the Feynman–Kac formula for the Laplace equation

Journal Article**Author(s):**

Pauli, Stefan; Gantner, Robert Nicholas; [Arbenz, Peter](#) ; [Adelmann, Andreas](#) 

Publication date:

2015-12

Permanent link:

<https://doi.org/10.3929/ethz-b-000107181>

Rights / license:

[In Copyright - Non-Commercial Use Permitted](#)

Originally published in:

BIT Numerical Mathematics 55(4), <https://doi.org/10.1007/s10543-014-0543-8>

Multilevel Monte Carlo for the Feynman–Kac formula for the Laplace equation

Stefan Pauli · Robert Nicholas Gantner ·
Peter Arbenz · Andreas Adelman

Received: 22 January 2014 / Accepted: 12 December 2014 / Published online: 20 January 2015
© Springer Science+Business Media Dordrecht 2015

Abstract Since its formulation in the late 1940s, the Feynman–Kac formula has proven to be an effective tool for both theoretical reformulations and practical simulations of differential equations. The link it establishes between such equations and stochastic processes can be exploited to develop Monte Carlo sampling methods that are effective, especially in high dimensions. There exist many techniques of improving standard Monte Carlo sampling methods, a relatively new development being the so-called Multilevel Monte Carlo method. This paper investigates the applicability of multilevel ideas to the stochastic representation of partial differential equations by the

Communicated by Desmond Higham.

The work of Stefan Pauli has been funded by the ETH interdisciplinary research Grant CH1-03 10-1.

S. Pauli (✉)
Computer Science Department and Seminar for Applied Mathematics,
ETH Zurich, Zurich, Switzerland
e-mail: stefan.pauli@inf.ethz.ch

R. N. Gantner · P. Arbenz
Computer Science Department, ETH Zurich, Zurich, Switzerland
e-mail: robert.gantner@sam.math.ethz.ch

P. Arbenz
e-mail: arbenz@inf.ethz.ch

Present address:
R. N. Gantner
Seminar for Applied Mathematics, ETH Zurich, Zurich, Switzerland

A. Adelman
Large Research Facilities (GFA), Paul Scherrer Institute (PSI), Villigen, Switzerland
e-mail: andreas.adelman@psi.ch

Feynman–Kac formula, using the Walk on Spheres algorithm to generate the required random paths. We focus on the Laplace equation, the simplest elliptic PDE, while mentioning some extension possibilities.

Keywords Multilevel Monte Carlo · Feynman–Kac · Walk on Spheres · Laplace equation

Mathematics Subject Classification 60H30 · 65C05 · 65N99

1 Introduction

The Monte Carlo (MC) error normally converges like $1/\sqrt{M}$, where M is the number of samples. If all samples are equally expensive, the error versus work convergence rate is $1/\sqrt{W}$, where W is the total work. Note that for a large class of problems the accuracy of the MC method is not only determined by the number of samples used, but also by the bias accepted in the computation of each sample. A small bias, using e.g. the Walk on Spheres (WoS) algorithm [15], comes with larger computation costs per sample. Hence the error versus work convergence rate falls short of $1/\sqrt{W}$. If applicable, Multilevel Monte Carlo (MLMC) methods may reach this optimal $1/\sqrt{W}$ convergence rate. This motivated us to evaluate the Feynman–Kac formula with MLMC methods instead of MC methods.

In this paper we present a procedure to evaluate the Feynman–Kac formula with MLMC using the WoS method. As a model problem we consider the Laplace equation, which is solved in high dimensions for instance for option pricing [12, Chap. 8] or in particle accelerator simulations [1]. We compute the MC convergence rate and show that it is suboptimal by a factor of $\log(W)$. We prove that for the MLMC method, the error versus work converges with the optimal $1/\sqrt{W}$ convergence rate. Our MLMC simulations, executed with an MPI parallel implementation, were up to twice as fast compared to the standard MC implementation.

First, in Sect. 2 we introduce the Feynman–Kac formula, using the WoS algorithm, and compute the error versus work convergence rate when using MC. Then, we derive a method to use MLMC in this setting, and evaluate its error versus work convergence rate in Sect. 3. In Sect. 4 we present numerical results, which quantify the advantage of using MLMC. Finally, we draw our conclusions in Sect. 5.

2 Standard walk on spheres (WoS)

2.1 Brownian motion

A Brownian motion [7] denoted by X_t , started at a point in a certain connected domain D , has several characteristic quantities. These include the first exit time and the first exit point, two important concepts in the application of the Feynman–Kac formula.

Definition 2.1 (*First exit time*) The time at which a realization of a Brownian motion X_t , started at some point $x \in D \subset \mathbb{R}^d$, first hits the domain boundary ∂D is called the first exit time $\tau := \inf\{t > 0 : X_t \in \partial D\}$.

Definition 2.2 (*First exit point*) The point at which a realization of a Brownian motion X_t , started at some point $x \in D \subset \mathbb{R}^d$, first hits the domain boundary ∂D is called the first exit point X_τ .

Of great importance for the derivation of the $W \circ S$ algorithm is the *distribution* of first exit points of a Brownian motion in a ball $B_x(R)$ of radius R centered at the point x . If the starting point of the Brownian motion is the center of the ball, this distribution is uniform on its surface, independent of the dimension [19, Theorem 3].

In evaluating functionals of stochastic processes, one is often interested in their expectations and variances. Throughout this paper, the L^2 -norm over the sample space Ω is defined as the expectation $\|X\|_{L^2(\Omega)} := \mathbb{E}[|X|^2]^{\frac{1}{2}}$.

2.2 Feynman–Kac formula for the Laplace equation

The Feynman–Kac formula, developed by Richard Feynman and Mark Kac [13], gives a probabilistic representation of the solution to certain PDEs at a single fixed evaluation point x . The Feynman–Kac formula allows us to write the solution of a very general elliptic PDE

$$Lu = \frac{1}{2} \Delta u + \sum_{i=1}^d b_i(x) \partial_i u + c(x)u = -g(x), \quad \text{in } D \subset \mathbb{R}^d,$$

$$u(x) = f(x), \quad \text{on } \partial D,$$

as an expectation over a stochastic process X_t beginning at the point x . This requires that b , c and g are smooth and satisfy a Lipschitz growth condition [6]. This process satisfies the stochastic differential equation $dX_t = b(x) dt + dW$, where dW is a Brownian increment, and is stopped as soon as it hits the boundary. Denoting the first exit time by τ and the condition that $X_0 = x$ by the subscript x in \mathbb{E}_x , the Feynman–Kac formula can be written as

$$u(x) = \mathbb{E}_x \left[f(X_\tau) \exp \left(\int_0^\tau c(X_s) ds \right) + \int_0^\tau g(X_t) \exp \left(\int_0^t c(X_s) ds \right) dt \right].$$

In this paper we are only interested in the Laplace equation $\Delta u = 0$ in d dimensions, $D \subset \mathbb{R}^d$. Thus, $b = c = g = 0$. This simplifies the development of an MLMC method, which can be used as a basis for more general equations. In this simple case the Feynman–Kac representation is given as the expectation of the exit point X_τ of a Brownian motion started at the point x . Denoting the condition $X_0 = x$ by the subscript x in \mathbb{E}_x and the boundary value at the exit point $Y := f(X_\tau)$, the solution is written as [6]

$$u(x) = \mathbb{E}_x[f(X_\tau)] =: \mathbb{E}_x[Y]. \tag{2.1}$$

2.3 Walk on spheres (WoS)

The WoS algorithm [15] can be viewed as an alternative to a conventional detailed simulation of the drift-free Brownian motion inside the domain D . For a Brownian path started at the center of a sphere, the exit point distribution is known to be an uniform distribution on the sphere [19]. This allows to simulate the Brownian motion using discrete jumps of a given size. Starting at $\tau = 0$ with $X_0 = x$ the algorithm measures the distance from the current position X_t to ∂D and jumps this distance in an uniformly random direction to the next position X_{t+1} . The algorithm terminates when it is ε -close to the boundary ∂D at the point X_N . The first exit point is approximated by the point $\bar{X}_N \in \partial D$ that is closest to X_N . This WoS algorithm may be used to solve the Laplace equation [15]. We recall that $\mathbb{E}[Y] = u(x)$ when Y is computed based on a processes started at x . $Y_\varepsilon := f(\bar{X}_N)$ is the estimator based on a WoS processes with the parameter ε .

The MC estimate uses M realizations $\{Y_\varepsilon^i\}_{i=1}^M$ of the random variable Y_ε obtained from independent paths of the WoS process to approximate the solution $\mathbb{E}[Y] \approx E_M[Y_\varepsilon] = \frac{1}{M} \sum_{i=1}^M Y_\varepsilon^i$. The errors associated with this approximation are twofold. One error results from estimating the expected solution $\mathbb{E}[Y_\varepsilon] \approx E_M[Y_\varepsilon]$ with either MC (see Sect. 2.4.1) or with MLMC (see Sect. 3). This is a purely statistical error, associated with the uncertainty in estimating the mean $\mathbb{E}[Y_\varepsilon]$ with finitely many realizations. This error can be reduced by increasing the number of realizations M . The other error is related to the stopping region of size ε as the individual paths are not computed exactly, but approximated with a WoS process. In fact the WoS paths stop prior to hitting the boundary, but at most ε away from it. This introduces a bias of $\mathcal{O}(\varepsilon)$ [14] as the stopping point depends on an ε perturbation of the boundary, resulting in

$$e_{\text{bias}} = \|\mathbb{E}[Y] - \mathbb{E}[Y_\varepsilon]\|_{L^2(\Omega)} \leq \mathcal{O}(\varepsilon). \quad (2.2)$$

Intuitively, we solve a different problem (up to the sampling error), with the boundary position perturbed by at most ε , resulting in the above bias.

Multiple upper bounds on the expected WoS path length $\mathbb{E}[N]$ are derived in [4]. For problems in dimensions $d \geq 3$ the upper bound $\mathbb{E}[N] = \mathcal{O}(\varepsilon^{4/d-2})$ holds. For domains fulfilling certain regularity conditions, tighter upper bounds of $\mathbb{E}[N] = \mathcal{O}(\log^p(\varepsilon^{-1}))$, $d \geq 2$ are proven for $p = 1$ or $p = 2$ depending on the domain. These estimates are relevant for bounding the expected work per sample.

2.4 Monte Carlo

2.4.1 Statistical error

Given M realizations $\{Y_\varepsilon^i\}_{i=1}^M$ of the random variable Y_ε , the value for $\mathbb{E}[Y_\varepsilon]$ can be approximated by the estimator $E_M[Y_\varepsilon] = \frac{1}{M} \sum_{i=1}^M Y_\varepsilon^i$. The error of this Monte Carlo estimator can be written in terms of the random variable Y_ε as

$$\begin{aligned}
 e_{\text{stat}} &= \|\mathbb{E}[Y_\varepsilon] - E_M[Y_\varepsilon]\|_{L^2(\Omega)} = \text{Var} \left[\frac{1}{M} \sum_{i=1}^M Y_\varepsilon^i \right]^{\frac{1}{2}} \\
 &= \frac{1}{\sqrt{M}} \sqrt{\text{Var}[Y_\varepsilon]} \leq \frac{1}{\sqrt{M}} \|Y_\varepsilon\|_{L^2(\Omega)},
 \end{aligned}
 \tag{2.3}$$

where we use the fact that the samples Y_ε^i are independent and identically distributed (i.i.d.) realizations of Y_ε .

2.4.2 Total error

The total error of an estimation $E_M[Y_\varepsilon]$ of $\mathbb{E}[Y]$ by the $W\circ S$ algorithm can be bounded using the triangle inequality,

$$\|\mathbb{E}[Y] - E_M[Y_\varepsilon]\|_{L^2(\Omega)} \leq \|\mathbb{E}[Y] - \mathbb{E}[Y_\varepsilon]\|_{L^2(\Omega)} + \|\mathbb{E}[Y_\varepsilon] - E_M[Y_\varepsilon]\|_{L^2(\Omega)}.$$

Using (2.2) and (2.3), we can bound the total error by

$$e_{\text{tot}} := \|\mathbb{E}[Y] - E_M[Y_\varepsilon]\|_{L^2(\Omega)} \leq \mathcal{O}(\varepsilon) + \sqrt{\frac{\text{Var}[Y_\varepsilon]}{M}}.
 \tag{2.4}$$

2.4.3 Error equilibration

One way of choosing the sample size is by equilibrating the statistical and discretization errors in (2.4), $\mathcal{O}(\varepsilon) = \sqrt{\text{Var}[Y_\varepsilon]}/M$. This yields the relationship

$$M = \mathcal{O}(\varepsilon^{-2}),
 \tag{2.5}$$

giving a total error behavior of $\mathcal{O}(\varepsilon)$.

2.5 Error versus expected work

The expected total work $\mathbb{E}[W]$ of a $W\circ S$ simulation is the number of paths times the expected length of a path:

$$\mathbb{E}[W] = M \cdot \mathbb{E}[N].$$

Each bound for the expected path length $\mathbb{E}[N]$ shown in Sect. 2.3 is multiplied with the M in (2.5), and the resulting expected work is solved for ε . Equation (2.6a) is valid for $d \geq 3$, Eq. (2.6b) for well-behaved domains in $d \geq 2$ dimensions. To derive (2.6b) we make use of the Lambert W -function, $\mathcal{W}_{\text{lam}}(\cdot)$, defined as the inverse of the map $w \mapsto w \exp(w)$. It can be approximated with the truncated expansion $\mathcal{W}_{\text{lam}}(x) \approx \log(x) - \log \log(x)$ [5]. This yields the following relationships:

$$\mathbb{E}[W] = \mathcal{O}(\varepsilon^{4/d-4}) \qquad e_{\text{tot}} = \mathcal{O}(\varepsilon) = \mathcal{O}(\mathbb{E}[W]^{\frac{1}{4} \frac{d}{1-d}})
 \tag{2.6a}$$

$$\mathbb{E}[W] = \mathcal{O}(\varepsilon^{-2} \log^2(\varepsilon^{-1})) \qquad e_{\text{tot}} = \mathcal{O}(\varepsilon) = \mathcal{O}(\mathbb{E}[W]^{-\frac{1}{2}} \log(\mathbb{E}[W]))
 \tag{2.6b}$$

The total error is linear in ε by the choice in (2.5). In both cases the WoS algorithm performs worse than the optimum possible in a Monte Carlo setting, namely $W^{-\frac{1}{2}}$. This motivates the formulation of a multilevel version of the WoS algorithm, with the hopes of achieving the optimal convergence rate.

3 Multilevel Walk on Spheres (MLWoS)

In this section, a multilevel version of the WoS algorithm is formulated and its error behavior analyzed. The main idea is to execute the WoS algorithm on different “discretization levels”, meaning for different values of the discretization parameter ε . The subscript $\ell \in \{0, \dots, L\}$ is used to denote a certain discretization level, where $\ell = 0$ is the coarsest discretization level, corresponding to a WoS simulation with ε_0 , and $\ell = L$ is the finest discretization level.

3.1 Multilevel formulation

On each discretization level $\ell = 0, \dots, L$, we define the discretization parameter $\varepsilon_\ell := \eta^{-\ell} \varepsilon_0$, where $\eta > 1$, i.e. the width of the stopping region is divided by η between successive levels.

In the following definition, a single WoS process with ε_ℓ is used to define the multilevel process in order to incorporate the fact that stopping points on higher levels are continuations of previously stopped processes.

Definition 3.1 (*Multilevel Walk on Spheres (MLWoS) process*) Given a domain $D \subset \mathbb{R}^d$, a point $x \in D$, a discretization parameter ε_0 and an $\eta > 1$. Consider a WoS process $\{X_i\}_{i=0}^{N_\ell}$ started at x with $\varepsilon_\ell = \eta^{-\ell} \varepsilon_0$ and the pair $(X_{N_{\ell-1}}, X_{N_\ell})$ obtained by $N_{\ell-1} = \min\{N \in \mathbb{N} : d_{\partial D}(X_N) < \varepsilon_{\ell-1}\}$, where $d_{\partial D}(X_N) = \min_{x' \in \partial D} |X_N - x'|$ is the distance to the boundary ∂D . The MLWoS on the level ℓ is the set of all such pairs.

In the context of multilevel Monte Carlo, we have multiple estimators Y_{ε_ℓ} , one for each discretization level with discretization parameter ε_ℓ . These are obtained in the same way as for the non-multilevel case, e.g. $Y_{\varepsilon_\ell} = f(\bar{X}_{N_\ell})$. The expectation of the estimator Y_{ε_L} is written in multilevel form as

$$\mathbb{E}[Y_{\varepsilon_L}] = \mathbb{E}[Y_{\varepsilon_0}] + \sum_{\ell=1}^L \mathbb{E}[Y_{\varepsilon_\ell} - Y_{\varepsilon_{\ell-1}}].$$

Replacing the expectation with the average over M_ℓ realizations on each discretization level ℓ , we get the MLMC estimate

$$E[Y_{\varepsilon_L}] = E_{M_0}[Y_{\varepsilon_0}] + \sum_{\ell=1}^L E_{M_\ell}[Y_{\varepsilon_\ell} - Y_{\varepsilon_{\ell-1}}].$$

The “sample” on level $\ell > 0$ is now $(Y_{\varepsilon_\ell} - Y_{\varepsilon_{\ell-1}})$, for which it is assumed that the two values on discretization levels ℓ and $\ell - 1$ come from the same WoS path. The

implication is that they should be simulated with the MLWOS process from Definition 3.1. The remaining estimator on discretization level 0 is computed with an ordinary WOS simulation.

Note that the individual Y_{ε_ℓ} are estimators on the *discretization level* ℓ , whereas $Y_{\varepsilon_\ell} - Y_{\varepsilon_{\ell-1}}$ are referred to as estimators on *level* ℓ .

3.2 Multilevel error bounds

The multilevel Monte Carlo error is defined as the difference between the expectation of the exact estimator $\mathbb{E}[Y]$ and the MLMC approximation involving all levels, $E[Y_{\varepsilon_L}]$, and is given by [2]

$$\begin{aligned} \|\mathbb{E}[Y] - E[Y_{\varepsilon_L}]\|_{L^2(\Omega)} &\leq \|\mathbb{E}[Y] - \mathbb{E}[Y_{\varepsilon_L}]\|_{L^2(\Omega)} \\ &\quad + M_0^{-\frac{1}{2}} \|Y_{\varepsilon_0}\|_{L^2(\Omega)} + \sum_{\ell=1}^L M_\ell^{-\frac{1}{2}} \|Y_{\varepsilon_\ell} - Y_{\varepsilon_{\ell-1}}\|_{L^2(\Omega)}. \end{aligned}$$

3.3 Asymptotic variance convergence rate

We want to bound the variance of the estimator on a level ℓ in the multilevel error bound. In more exact terms, we want to determine for which functions f and domains D the relationship

$$\|Y_{\varepsilon_\ell} - Y_{\varepsilon_{\ell-1}}\|_{L^2(\Omega)} = \mathcal{O}(\varepsilon_\ell^s) \quad \forall \ell > \ell_{min} \tag{3.1}$$

holds for some $\ell_{min} > 0, s > 0$. This behavior is desired since it allows a good multilevel performance. It ensures that the variance is small on the fine levels, such that only few fine-level realizations are required.

Conditioning the expectation on the current position $X_{N_{\ell-1}}$ incorporates the fact that the simulation continues with a given path from level $\ell - 1$, not generating a completely independent one on the finer level ℓ . Using $\bar{X}_{N_\ell} = \bar{X}_{N_{\ell-1}} + \Delta x$ yields

$$\|Y_{\varepsilon_\ell} - Y_{\varepsilon_{\ell-1}}\|_{L^2(\Omega)}^2 = \mathbb{E}\left[|f(\bar{X}_{N_\ell} + \Delta x) - f(\bar{X}_{N_{\ell-1}})|^2 \middle| X_{N_{\ell-1}}\right].$$

In order to write this expectation in terms of $|\Delta x|$, rather than in terms of the process $\bar{X}_{N_{\ell-1}}$, we make the following assumption:

Assumption 3.1 (*Hölder continuity*) There exist $C, \alpha > 0$ such that

$$|f(x) - f(y)| \leq C|x - y|^\alpha$$

for all $x, y \in \partial D$.

This assumption implies

$$\mathbb{E}\left[|f(\bar{X}_{N_{\ell-1}} + \Delta x) - f(\bar{X}_{N_{\ell-1}})|^2 \mid \bar{X}_{N_{\ell-1}}\right] \leq C^2 \mathbb{E}\left[|\Delta x|^{2\alpha} \mid \bar{X}_{N_{\ell-1}}\right]. \tag{3.2}$$

The conditioning can be omitted by bounding the right-hand side of (3.2) by the maximum over $\bar{X}_{N_{\ell-1}}$.

Splitting the expectation into a converging part where $|\Delta x| \leq R_\ell$ and a ‘‘diverging’’ part with $R_\ell < |\Delta x| \leq |D|$, where $|D|$ is the diameter of the circumsphere of D , we get

$$\begin{aligned} \mathbb{E}\left[|\Delta x|^{2\alpha}\right] &\leq \mathbb{P}[|\Delta x_\ell| \leq R_\ell] R_\ell^{2\alpha} + \mathbb{P}[|\Delta x_\ell| > R_\ell] |D|^{2\alpha} \\ &= (1 - p_{\text{div}}) R_\ell^{2\alpha} + p_{\text{div}} |D|^{2\alpha}. \end{aligned} \tag{3.3}$$

We denote by $p_{\text{div}} = \mathbb{P}\left[\bar{X}_{N_{\ell+1}} \notin B_{\bar{X}_{N_\ell}}(R_\ell)\right]$ the *divergence probability*, defined as the probability that the projected points on two consecutive levels, \bar{X}_{N_ℓ} and $\bar{X}_{N_{\ell+1}}$, are further than R_ℓ apart.

Since $|D|$ in the second term is a domain-dependent constant, we must find a bound for the divergence probability p_{div} in terms of the discretization parameter ε . As shown below, this is possible with the resulting behavior $p_{\text{div}} \propto \frac{\varepsilon_\ell}{R_\ell - \varepsilon_\ell}$.

In order for the expectation to converge with a certain rate in ε , both terms $(1 - p_{\text{div}}) R_\ell^{2\alpha}$ and p_{div} should converge. Solving $R_\ell^{2\alpha} = C \frac{\varepsilon_\ell}{R_\ell}$ results in $R_\ell = C \varepsilon_\ell^{\frac{1}{2\alpha+1}}$, which inserted in (3.3) yields

$$\begin{aligned} \mathbb{E}\left[|\Delta x|^{2\alpha}\right] &\leq R_\ell^{2\alpha} + p_{\text{div}} |D|^{2\alpha} \\ &\leq C^{2\alpha} \varepsilon_\ell^{\frac{2\alpha}{2\alpha+1}} + C' \frac{\varepsilon_\ell}{C \varepsilon_\ell^{\frac{1}{2\alpha+1}} - \varepsilon_\ell} |D|^{2\alpha} \\ &= C^{2\alpha} \varepsilon_\ell^{\frac{2\alpha}{2\alpha+1}} + C' \frac{\varepsilon_\ell}{\varepsilon_\ell^{\frac{1}{2\alpha+1}} (C - \varepsilon_\ell^{\frac{2\alpha}{2\alpha+1}})} |D|^{2\alpha} \\ &= C^{2\alpha} \varepsilon_\ell^{\frac{2\alpha}{2\alpha+1}} + C' \frac{\varepsilon_\ell^{\frac{2\alpha}{2\alpha+1}}}{C - \varepsilon_\ell^{\frac{2\alpha}{2\alpha+1}}} |D|^{2\alpha} \\ &= \mathcal{O}\left(\varepsilon_\ell^{\frac{2\alpha}{2\alpha+1}}\right) \end{aligned} \tag{3.4}$$

and hence

$$\|Y_{\varepsilon_\ell} - Y_{\varepsilon_{\ell-1}}\|_{L^2(\Omega)}^2 = \mathcal{O}\left(\varepsilon_\ell^{\frac{2\alpha}{2\alpha+1}}\right).$$

For differentiable f , we have $\alpha = 1$ and thus the convergence rate $\mathcal{O}(\varepsilon_\ell^{\frac{1}{3}})$ in (3.1).

3.3.1 Convergence of divergence probability

Our goal here is to obtain a bound on the divergence probability p_{div} on a certain level ℓ depending on ε_ℓ . Divergence here means that the process does not result in an estimate $\bar{X}_{N_{\ell+1}}$ that is located within a ball $B_{\bar{X}_{N_\ell}}(R_\ell)$ of radius R_ℓ around the current projected stopping point \bar{X}_{N_ℓ} .

In the derivation we use the so-called harmonic measure. There exist a few equivalent definitions of harmonic measure which are relevant to the current application. The first defines the harmonic measure as a harmonic function satisfying certain boundary conditions, the second is instructive in the context of Monte Carlo approximations of an integral and the third uses a definition relying on the distribution of the first exit point of a Brownian motion. Consider the closed set $D \subset \mathbb{R}^d$ with $d \geq 2$ and let $\mathcal{B}(\partial D)$ denote the σ -algebra of subsets of ∂D and define:

Definition 3.2 (*Harmonic measure—Dirichlet solution*) The harmonic measure $\omega_D : D \times \mathcal{B}(\partial D) \rightarrow [0, 1]$, viewed as a function $\omega_D(x, E)$ of x for every fixed D and E , is the unique harmonic function that satisfies the boundary condition [10, p. 39]

$$f(x) = \begin{cases} 1, & \text{if } x \in E, \\ 0, & \text{if } x \in \partial D \setminus E. \end{cases}$$

Definition 3.3 (*Harmonic measure—integral representation*) The harmonic measure $\omega_D : D \times \mathcal{B}(\partial D) \rightarrow [0, 1]$ is the unique function that satisfies [18, sec. 4.3]

- (a) for each $x \in D$, $E \mapsto \omega_D(x, E)$ is a probability measure on ∂D
- (b) if $f : \partial D \rightarrow \mathbb{R}$ is continuous, the solution of the Laplace equation in D is given by

$$u(x) = \int_{\partial D} f(x') \omega_D(x, dx'). \tag{3.5}$$

Definition 3.4 (*Harmonic measure—hitting probability*) Let X_t^x denote a Brownian motion started at x . The harmonic measure $\omega_D : D \times \mathcal{B}(\partial D) \rightarrow [0, 1]$ is given by

$$\omega_D(x, E) := \mathbb{P}[X_\tau^x \in E, \tau < \infty],$$

where $\tau = \inf\{t \geq 0 : X_t^x \in \partial D\}$ is the first exit time of X_t^x from D . [10, Theorem F.6.]

By (3.5), we can view the solution $u(x)$ of the Laplace equation as the expectation value of $f(X_\tau)$ under the distribution of first hitting points with probability measure given by ω_D . This agrees with the Feynman–Kac formula (2.1) applied to the Laplace equation. Simulating realizations of the stochastic process X_t thus amounts to approximating the integral (3.5) by Monte Carlo integration, i.e. generating realizations of the exit points distributed according to the harmonic measure.

Often, the harmonic measure in a certain domain can't be given analytically. Thus, it is important to be able to bound it from above and below. This can be accomplished using the general inequalities popularized by Carleman, illustrated in Fig. 1 and often

Fig. 1 An illustration of Carleman’s Principle in two dimensions. α is the unchanged part of the boundary, β is changed to β' . The extended domain with boundary $\alpha \cup \beta'$ is G' , G is the original one

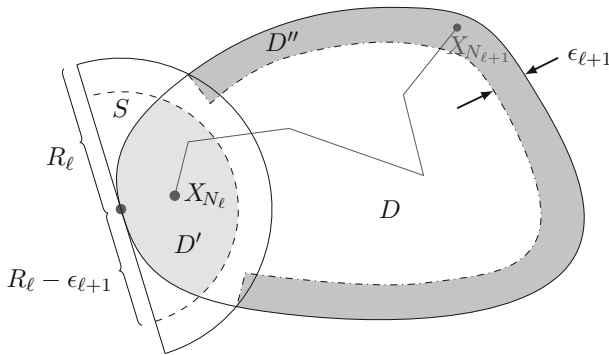
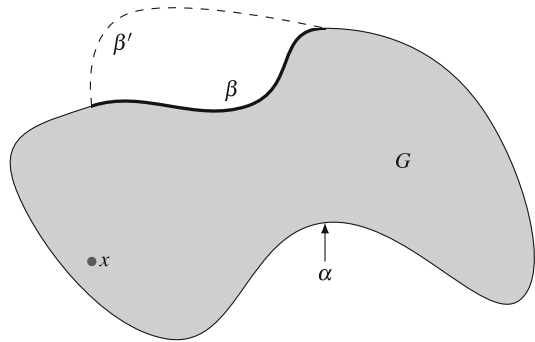


Fig. 2 Domain $D \subset \mathbb{R}^2$, with the semicircle S of radius $R_\ell - \epsilon_{\ell+1}$ used to bound the harmonic measure of a point X_{N_ℓ} located at most ϵ_ℓ from the boundary. The shaded region D' is the intersection of the domain D with the semicircle S . The path to $X_{N_{\ell+1}}$ is “divergent” with respect to X_{N_ℓ}

referred to as *Carleman’s Principle of Monotonicity* or the *Principle of Extension of Domain*.

Definition 3.5 (*Carleman’s principle*) The harmonic measure of $x \in G$ and a portion α of the boundary of $G \subset \mathbb{R}^d$ increases if the domain G is enlarged to $G' \supseteq G$ by modifying $\beta = \partial G \setminus \alpha$:

$$\omega_G(x, \alpha) \leq \omega_{G'}(x, \alpha). \tag{3.6}$$

Since $\omega_G(x, \alpha) + \omega_G(x, \beta) = 1$, it follows that the harmonic measure on the extended portion of the boundary decreases with respect to the new domain: [17, p. 68], [8, p. 131]

$$\omega_G(x, \beta) \geq \omega_{G'}(x, \beta'). \tag{3.7}$$

We begin by sketching the situation in Fig. 2. Two portions of the domain D are of relevance: D' , the intersection of D with a semi-ball S of radius $R_\ell - \epsilon_{\ell+1} > \epsilon_\ell$, and D'' , the portion of the stopping region of D with discretization parameter $\epsilon_{\ell+1}$ leading to a divergent path.

Since the two domains D'' and D' do not overlap, a MLWoS path related to a Brownian motion which does not enter $D \setminus D'$ does not diverge. This bounds the divergence

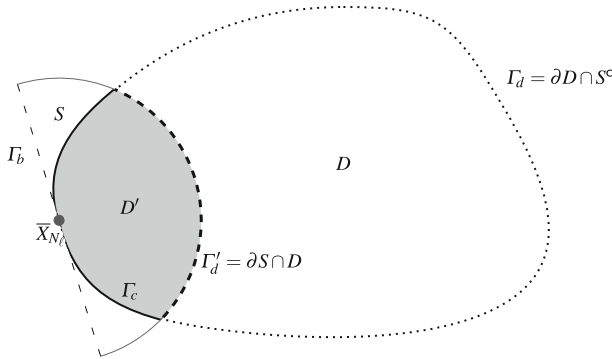


Fig. 3 The boundary of the domain D is split into two disjoint portions, the “convergent” boundary Γ_c and the “divergent” boundary Γ_d . The portion of the *semicircle boundary* ∂S inside of D is denoted by Γ'_d and the base of the *semicircle* is Γ_b

probability p_{div} of a MLWOS path by the harmonic measure $p_{\text{div}} \leq \omega_{D'}(X_{N_\ell}, \Gamma'_d)$, where Γ'_d is the portion of the semicircle boundary ∂S inside of D (see Fig. 3).

In Fig. 3, the boundaries are split up such that Carleman’s method can be applied to the harmonic measure $\omega_{D'}(X_{N_\ell}, \Gamma'_d)$.

We keep the portion Γ'_d fixed and extend the remaining portion Γ_c of the boundary of D' to the semicircle boundary $\partial D \setminus \Gamma'_d$. This lets us use the first inequality (3.6), where Γ'_d corresponds to α and Γ_c corresponds to β in the definition, yielding

$$\omega_{D'}(x, \Gamma'_d) \leq \omega_S(x, \Gamma'_d) \leq \omega_S(x, \partial S \setminus \Gamma_b),$$

where the second inequality comes from adding $\omega_S(x, \partial S \setminus (\Gamma'_d \cup \Gamma_b))$, the harmonic measure of the remaining portion of the arc, and using $\omega_S(x, \partial S) = 1$. Fortunately, $\omega_S(x, \Gamma_b)$ can be calculated analytically for any point $x \in D'$, giving us a bound for $\omega_D(x, \Gamma_d)$.

Define $f(x) := \frac{2\theta_x}{\pi} - 1$, where θ_x is the angle at some point x when viewing the two end-points of the diameter of S , see Fig. 4. For points on the base Γ_b of S , $\theta = \pi$ and $f(x) = 1$. On the arc, the value for θ is always $\frac{\pi}{2}$, yielding $f(x) = 0$. One can also show by inserting $\theta_x = \arctan\left(\frac{R-x}{y}\right) + \arctan\left(\frac{R+x}{y}\right)$ that f is harmonic. Since by (3.2) this function is unique, it must be the harmonic measure:

$$\omega_S(x, \Gamma_b) = \frac{2\theta_x}{\pi} - 1. \tag{3.8}$$

We now make the assumption that x lies on the plane perpendicular to the base of the hemisphere going through the midpoint of its base, as in Fig. 5. This allows the simplification $\theta = 2 \arctan((R_\ell - \varepsilon_{\ell+1})/d(x, \Gamma_b))$. By the construction of the WOS process, we know that $x = X_{N_\ell}$ is the stopping point on level ℓ and is at most ε_ℓ away from the boundary point \bar{X}_{N_ℓ} , which corresponds to the midpoint of the base of the hemisphere. Thus, we bound $d(x, \Gamma_b) \leq \varepsilon_\ell$, leading to $\theta \geq 2 \arctan((R_\ell -$

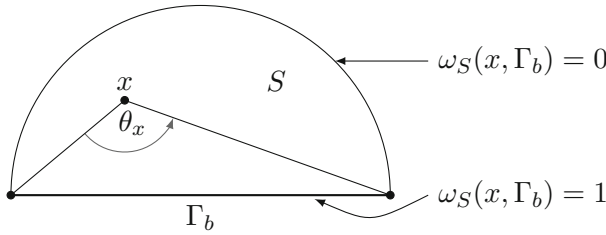


Fig. 4 A semicircle containing a point x with opening angle θ_x that determines the harmonic measure $\omega_S(x, \Gamma_b)$. It assumes the value 0 on the arc and 1 on Γ_b

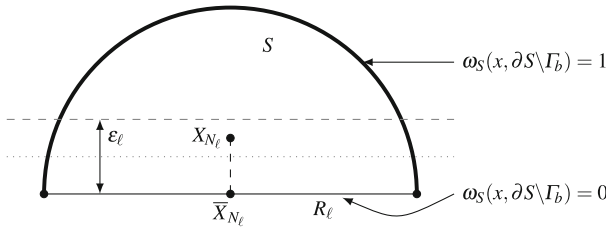


Fig. 5 More detailed illustration of the situation from Fig. 2

$\varepsilon_{\ell+1})/\varepsilon_\ell$) by monotonicity and obtain the result

$$1 - \omega_S(X_{N_\ell}, \Gamma_b) = 2 \left(1 - \frac{\theta}{\pi} \right) \leq 2 \left(1 - \frac{2}{\pi} \arctan \left(\frac{R_\ell - \varepsilon_{\ell+1}}{\varepsilon_\ell} \right) \right).$$

Remembering that $R_\ell - \varepsilon_{\ell+1} \gg \varepsilon_\ell$ we can use a series expansion of $\arctan(\xi)$ around $\xi = \infty$, resulting in a final bound for the divergence probability in terms of ε_ℓ

$$p_{\text{div}} \leq \frac{4C}{\pi} \frac{\varepsilon_\ell}{R_\ell - \varepsilon_{\ell+1}} \leq \frac{4C}{\pi} \frac{\varepsilon_\ell}{R_\ell - \varepsilon_\ell}, \tag{3.9}$$

for some constant C due to discarding the nonlinear terms of the series.

3.3.2 Optimal number of samples

To simplify notation, we define the estimator Y_ℓ to be

$$Y_\ell := \begin{cases} Y_{\varepsilon_0}, & \text{if } \ell = 0, \\ Y_{\varepsilon_\ell} - Y_{\varepsilon_{\ell-1}}, & \text{if } \ell > 0, \end{cases}$$

and w_ℓ the expected work performed to compute a sample on level ℓ .

Following [16], the optimization problem to solve is then

$$\min \mathbb{E}[W_{\text{tot}}] = \min \sum_{\ell=0}^L M_\ell w_\ell \quad \text{s.t.} \quad \varepsilon_L = \sqrt{\sum_{\ell=0}^L \frac{\text{Var}[Y_\ell]}{M_\ell}},$$

hence minimizing the expected work, while maintaining a sampling error proportional to the discretization error ($\mathcal{O}(\varepsilon_L)$). [16] provides the optimal number of samples per level

$$M_\ell = \frac{1}{\varepsilon_L^2} \sqrt{\frac{\text{Var}[Y_\ell]}{w_\ell}} \sum_{\ell=0}^L \sqrt{\text{Var}[Y_\ell] w_\ell}. \tag{3.10}$$

3.4 Error versus work

The following basic relations will be used in the sequel.

Refinement of ε_ℓ	$\varepsilon_\ell = \eta^{-\ell} \varepsilon_0 = \eta^{L-\ell} \varepsilon_L$
Scaling of work	$w_\ell = \begin{cases} \mathcal{O}(\varepsilon_\ell^{-\gamma}) \\ \mathcal{O}(\log^p(\varepsilon_\ell^{-1})) = \mathcal{O}(\ell^p) \end{cases}$
Scaling of variance	$\text{Var}[Y_\ell] = \mathcal{O}(\varepsilon_\ell^{2s})$

Determining M_ℓ (3.10) based on the estimates of variance and work leads to

$$M_\ell = \begin{cases} \varepsilon_L^{-2} \sqrt{\varepsilon_\ell^{2s+\gamma}} \sum_{\ell=0}^L \sqrt{\varepsilon_\ell^{2s-\gamma}} \\ \varepsilon_L^{-2} \sqrt{\varepsilon_\ell^{2s} \cdot \ell^{-p}} \sum_{\ell=0}^L \sqrt{\varepsilon_\ell^{2s} \cdot \ell^p} \end{cases} = \begin{cases} \varepsilon_L^{-2} \sqrt{\varepsilon_\ell^{2s+\gamma}} \sqrt{\varepsilon_0^{2s-\gamma}} \underbrace{\sum_{\ell=0}^L \sqrt{\eta^{(\gamma-2s)\ell}}}_{\leq C, \forall 2s > \gamma} \\ \varepsilon_L^{-2} \sqrt{\varepsilon_\ell^{2s} \cdot \ell^{-p}} \sqrt{\varepsilon_0^{2s}} \underbrace{\sum_{\ell=0}^L \sqrt{\eta^{-2s\ell} \cdot \ell^p}}_{\leq C} \end{cases} \tag{3.11}$$

The remaining summands can be bounded by a convergent geometric sequence. Thus, by the comparison test the infinite series ($L \rightarrow \infty$) converges to a constant, in the first case provided $2s > \gamma$.

Inserting (3.11) into the formula $W_L^{\text{tot}} = \sum_{\ell=0}^L M_\ell w_\ell$, we can write the work as a function of ε_L , which scales linearly with e_{tot}

$$W_L^{\text{tot}} \leq \begin{cases} C \varepsilon_L^{-2} \sqrt{\varepsilon_0^{2s-\gamma}} \sum_{\ell=0}^L \sqrt{\varepsilon_\ell^{2s-\gamma}} \stackrel{(3.11)}{\leq} C^2 \varepsilon_L^{-2} \varepsilon_0^{2s-\gamma}, & 2s > \gamma, \\ C \varepsilon_L^{-2} \sqrt{\varepsilon_0^{2s}} \sum_{\ell=0}^L \sqrt{\varepsilon_\ell^{2s} \cdot \ell^p} \stackrel{(3.11)}{\leq} C^2 \varepsilon_L^{-2} \varepsilon_0^{2s}, \end{cases} \\ = \mathcal{O}(\varepsilon_L^{-2}).$$

Thus, we obtain the optimal convergence rate

$$e_{\text{tot}} \propto (W_L^{\text{tot}})^{-\frac{1}{2}},$$

for the case $w_\ell = \mathcal{O}(\varepsilon_\ell^{-\gamma})$ provided $2s > \gamma$.

3.5 Measured values

The second approach to determining M_ℓ is based on estimating the work w_ℓ and the variance $\text{Var}[Y_\ell]$ on the different levels, rather than using their asymptotic convergence rates as in (3.11). In the absence of alternatives we still use the asymptotic convergence rate of the discretization error $\mathcal{O}(\varepsilon_L)$. As in [15] the required estimates are computed using the same samples already involved in the MLMC estimator.

A certain minimum number of samples (so-called warm-up samples) are needed on every level to provide an accurate estimate. Performance disadvantages arise if the number of required warm-up samples exceeds the optimal number of samples (3.10). This often happens for the finest level L . However, (3.11) proves that in this case, the optimal number of samples grows asymptotically, which implies that for small ε_L the required warm-up samples do not exceed the optimal number of samples. This technique complicates the implementation slightly, but does not increase the computational work for this application.

4 Numerical results

4.1 Model problems

In this section we formulate model problems on different domains to test our algorithm.

4.1.1 Hypercube

High dimensional problems suffer from the ‘‘curse of dimensionality’’ [3] when they are solved with standard deterministic methods, such as finite element or finite volume methods. The ‘‘curse of dimensionality’’ refers to the fact that the number of gridpoints in a d -dimensional hypercube when using a discretization with n points in each dimension grows like n^d , i.e. *exponentially* in the dimension.

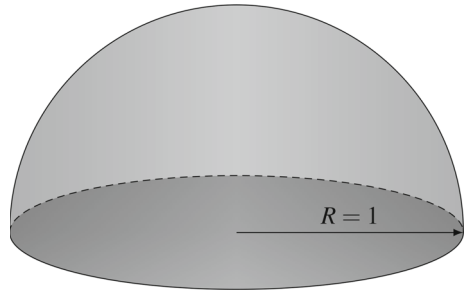
Here, we consider a ten-dimensional hypercube $D_\square := [0, 1]^{10}$ with the boundary conditions given by

$$f(\mathbf{x}) = \begin{cases} 4(x_1 - \frac{1}{2})^2, & \text{if } x_1 \leq \frac{1}{2} \\ 4(x_1 - \frac{3}{2})^2, & \text{if } x_1 \geq \frac{3}{2}, \\ 0, & \text{otherwise.} \end{cases} \quad (4.1)$$

4.1.2 Hemisphere

A three-dimensional problem to consider is the Laplace equation on a hemisphere $D_\Delta := \{\mathbf{x} \in \mathbb{R}^3 \mid |\mathbf{x}| \leq 1, x_3 \geq 0\}$, as illustrated in Fig. 6. The boundary conditions of the Laplace equation are chosen such that the analytical solution is given by $u(\mathbf{xx}) =$

Fig. 6 Hemisphere domain D_Δ in three dimensions



$[x_1^2 + x_2^2 + (x_3 + 1)^2]^{-\frac{1}{2}}$, as stated in [14]. This leads to the following formulation with Hölder exponent $\alpha = 1$:

$$\begin{aligned} \Delta u &= 0 \text{ in } D_\Delta \\ u(\mathbf{x}) &= [2(x_3 + 1)]^{-\frac{1}{2}} \text{ for } |\mathbf{x}| = 1 \\ u(\mathbf{x}) &= [x_1^2 + x_2^2 + 1]^{-\frac{1}{2}} \text{ for } x_3 = 0. \end{aligned}$$

The initial point was chosen to be $x_0 := (0.2, 0.3, 0.1)$.

4.2 Implementation

A generic C++ implementation for both the conventional and the multilevel version of the WoS algorithm was created. The M_ℓ are either determined analytically or based on measurements during the simulation. The code uses MPI to generate samples in parallel, such that larger problems can be solved. The code was tested with up to 1,000 cores. Further information on the implementation is found in [9].

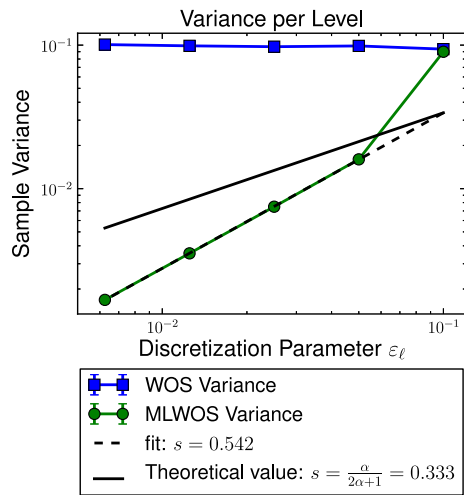
4.3 Measurement methodology

Monte Carlo methods are based on the approximation of an expectation by a sample mean. The resulting estimator is itself a random variable—thus, the total error is given by

$$\|\mathbb{E}[Y] - E[Y_{\varepsilon_L}]\|_{L^2(\Omega)} = \mathbb{E}\left[\|\mathbb{E}[Y] - E[Y_{\varepsilon_L}]\|^2\right]^{\frac{1}{2}}.$$

This outer expectation can again be approximated by a sample mean, corresponding to a repeated call of the corresponding algorithm. Thus, for a certain set of parameters one must call the implementation with different random number seeds and compute the sample mean over the resulting realizations. An estimate for a confidence interval can also be computed using the results of these repeated calls.

Fig. 7 D_{\square} Convergence of the variance as the discretization parameter ε_{ℓ} is reduced for the hypercube domain D_{\square} with $\alpha = 1$ using $\eta = 2$



4.4 Measured variance convergence rate

As stated in (3.1), the variance of the estimate on a level ℓ is assumed to converge like $\mathcal{O}(\varepsilon_{\ell}^s)$. In (3.4), it was shown that s is bounded by $\frac{\alpha}{2\alpha+1}$, α being the exponent of Hölder continuity of the boundary condition. The results presented here are empirical measurements showing the convergence of the variance of the estimator. The measurements were averaged over ten calls for the hypercube domain D_{\square} . In Fig. 7 the theoretical and the measured exponent s are shown.

Shown only for the hypercube domain, but observed for all described domains, the variance convergence rate is around 0.5, which is better than the value $\frac{1}{3}$ obtained from the analytical derivation assuming continuity of the boundary condition. This is not a contradiction since the derivation yields an upper bound. However, using the analytical upper bound will very likely result in suboptimal results, since obviously one can get away with fewer samples on the fine levels.

4.5 Measured convergence rate

In this section we compare the error versus work of the proposed multilevel $\mathbb{W}\circ\mathbb{S}$ algorithm with the plain $\mathbb{W}\circ\mathbb{S}$ algorithm. For the multilevel $\mathbb{W}\circ\mathbb{S}$ algorithm we distinguish between two strategies, one with analytically derived M_{ℓ} , w_{ℓ} and $\text{Var}[Y_{\ell}]$ (MLWOS) and the other with M_{ℓ} based on measured w_{ℓ} and $\text{Var}[Y_{\ell}]$ (MEAS) according to Sect. 3.5.

Measurements performed on the domain D_{\square} are shown in Fig. 8. We performed measurements with $\eta = 8$ and $\eta = 16$. The measured values for the average error and the average work are shown, together with the 1σ confidence interval, for different algorithms. Sampling was performed until the confidence interval was very small. In all cases the multilevel $\mathbb{W}\circ\mathbb{S}$ with analytically derived M_{ℓ} (MLWOS) performs poorly compared to the plain $\mathbb{W}\circ\mathbb{S}$ algorithm. The MLWOS performance improves for $\eta = 16$,

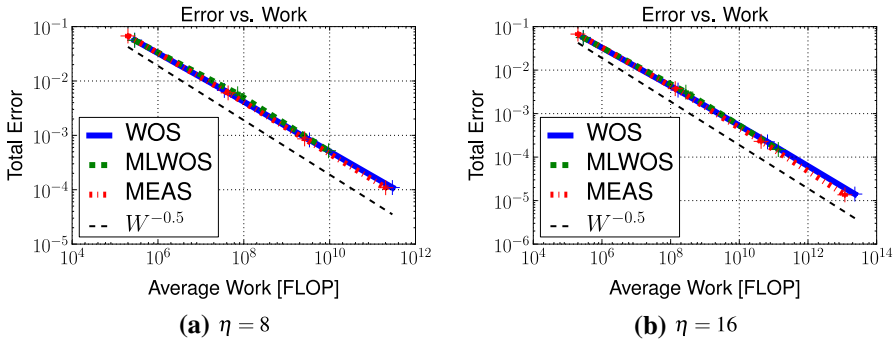
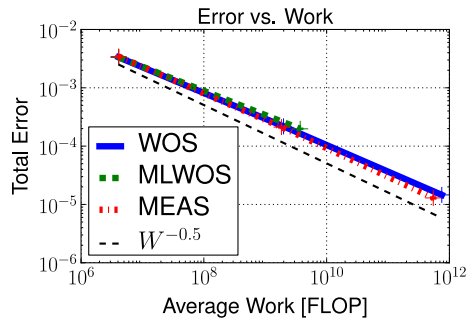


Fig. 8 Convergence of error versus work for the problem posed on D_{\square} , for various values of the refinement parameter η

Fig. 9 Convergence of the error versus work for the problem posed on D_{\triangle} , using $\eta = 16$



but an improvement over the plain $\mathbb{W}\circ\mathbb{S}$ algorithm is not measured. The multilevel $\mathbb{W}\circ\mathbb{S}$ with M_{ℓ} using measured values (MEAS) always performs better than the plain $\mathbb{W}\circ\mathbb{S}$ algorithm, especially for $\eta = 16$ where the computation is up to two times faster.

The analytically derived M_{ℓ} are based on the $\text{Var}[Y_{\ell}]$. As observed in Sect. 4.4, $\text{Var}[Y_{\ell}]$ converges faster than predicted in our theory. Hence it is not surprising that the multilevel $\mathbb{W}\circ\mathbb{S}$ algorithm with analytically derived M_{ℓ} performs suboptimally. The multilevel $\mathbb{W}\circ\mathbb{S}$ algorithm with M_{ℓ} based on measurements does not suffer from this problem, therefore we observe a better performance.

The parameter η does not influence the convergence rate of the multilevel scheme, but a clever choice may asymptotically reduce the work by a constant. Therefore it is expected that the performance of the multilevel $\mathbb{W}\circ\mathbb{S}$ algorithm depends on η . For certain problems (see e.g. [11]), this parameter is optimized analytically. Here, various values are tried on the domain D_{\square} in order to empirically find good values. We observed that $\eta = 16$ is a better choice than the other tested value $\eta = 8$.

Figure 9 shows measurements for the domain D_{\triangle} for $\eta = 16$. The results are similar to the ones seen for the domain D_{\square} . The multilevel $\mathbb{W}\circ\mathbb{S}$ with analytically derived M_{ℓ} (MLWOS) performs poorly, whereas the multilevel $\mathbb{W}\circ\mathbb{S}$ with M_{ℓ} using measured values (MEAS) is up to two times better than the plain $\mathbb{W}\circ\mathbb{S}$ algorithm.

5 Conclusions and outlook

In this work, a first application of the Multilevel Monte Carlo method to the Feynman–Kac formula was developed. A novel version of the Walk on Spheres process, the Multilevel Walk on Spheres was formulated and central quantities were derived. It was proven that the rate of convergence of the error versus the work of the Multilevel Walk on Spheres algorithm is optimal.

In order to measure the actual performance gain of the proposed method, an implementation of both the conventional and the multilevel version of the Walk on Spheres algorithm was created. By parallelizing the generation of samples, more thorough convergence results could be obtained. Additionally, a version of the multilevel method that chooses the number of samples based on measurements was implemented.

In order to test a variety of different situations, multiple domains and boundary conditions were implemented. The convergence rate of the error with respect to the work was measured, along with many other relevant quantities. Using the new Multilevel Walk on Spheres method, a reduction of the work by up to a factor 2 was achieved.

Using tighter bounds, it would perhaps be possible to find an upper bound for the variance convergence rate that better fits the measured behavior. This would hopefully allow to analytically determine the number of samples per level.

It would be beneficial to generalize the Multilevel Walk on Spheres to other elliptic equations, for example the Poisson equation with nonconstant right-hand side. This equation has applications in many fields, such as particle accelerator modeling or biochemical electrostatics in which a Multilevel Walk on Spheres formulation may be of use. This would require the formulation of a process that contains multiple discretization levels *along the path*, i.e., with level dependent maximal sphere radius. This would increase the expected work especially for fine discretizations and presumably increase the benefit of Multilevel Walk on Spheres over Walk on Spheres.

Acknowledgments We thank the anonymous reviewers for their valuable comments and suggestions that improved the quality of the paper.

References

1. Adelman, A., Arbenz, P., Ineichen, Y.: A fast parallel Poisson solver on irregular domains applied to beam dynamics simulations. *J. Comput. Phys.* **229**(12), 4554–4566 (2010)
2. Barth, A., Schwab, C., Zollinger, N.: Multi-level Monte Carlo finite element method for elliptic PDEs with stochastic coefficients. *Numer. Math.* **119**(1), 123–161 (2011)
3. Bellman, R.: *Dynamic Programming*. Princeton University Press, Princeton (1957)
4. Binder, I., Braverman, M.: The complexity of simulating Brownian motion. In: *Proceedings of the 20th annual ACM-SIAM symposium on discrete algorithms*, pp. 58–67 (2009)
5. Bruijn, N.: *Asymptotic Methods in Analysis*. Bibliotheca mathematica. Dover Publications, Mineola (1970)
6. Buchmann, F.M., Petersen, W.P.: Solving Dirichlet problems numerically using the Feynman–Kac representation. *BIT Numer. Math.* **43**, 519–540 (2003)
7. Chang, J.: *Stochastic Processes*. <http://www.stat.yale.edu/jtc5/251/stochastic-processes>. Lecture notes (2007)
8. Doob, J.L.: *Classical Potential Theory and Its Probabilistic Counterpart*. Classics in Mathematics Series. Springer, Berlin (2001)

9. Gantner, R. N.: Computing the Feynman–Kac formula efficiently with multilevel Monte Carlo. Master thesis, Computer Science Department, ETH Zurich (2013)
10. Garnett, J.B., Marshall, D.E.: Harmonic Measure. Cambridge University Press, Cambridge (2005)
11. Giles, M.B.: Multilevel Monte Carlo path simulation. *Oper. Res.* **56**(3), 607–617 (2008)
12. Hilber, N., Reichmann, O., Schwab, C., Winter, C.: Computational Methods for Quantitative Finance: Finite Element Methods for Derivative Pricing. Springer. Springer finance, Berlin (2013)
13. Kac, M.: On distributions of certain Wiener functionals. *Trans. Am. Math. Soc.* **65**(1), 1–13 (1949)
14. Mascagni, M., Hwang, C.O.: ε -Shell error analysis for “Walk on Spheres” algorithms. *Math. Comput. Simul.* **63**(2), 93–104 (2003)
15. Muller, M.E.: Some continuous Monte Carlo methods for the Dirichlet problem. *Ann. Math. Stat.* **27**(3), 569–589 (1955)
16. Müller, F., Jenny, P., Meyer, D.W.: Multilevel Monte Carlo for two phase flow and Buckley–Leverett transport in random heterogeneous porous media. *J. Comput. Phys.* **250**, 685–702 (2013)
17. Nevanlinna, R.H.: Analytic Functions. Princeton Mathematical Series. Princeton University Press, Princeton (1960)
18. Ransford, T.: Potential Theory in the Complex Plane. Cambridge University Press, Cambridge (1995)
19. Wendel, J.G.: Hitting spheres with Brownian motion. *Ann. Probab.* **8**(1), 164–169 (1980)