

COMPLEXITY ANALYSIS OF QUASI CONTINUOUS LEVEL MONTE CARLO *

CEDRIC AARON BESCHLE** AND ANDREA BARTH

Abstract. Continuous level Monte Carlo is an unbiased, continuous version of the celebrated multilevel Monte Carlo method. The approximation level is assumed to be continuous resulting in a stochastic process describing the quantity of interest. Continuous level Monte Carlo methods allow naturally for samplewise adaptive mesh refinements, which are indicated by (goal-oriented) error estimators. The samplewise refinement levels are drawn in the estimator from an exponentially-distributed random variable. Unfortunately in practical examples this results in higher costs due to high variance in the samples. In this paper we propose a variant of continuous level Monte Carlo, where a quasi Monte Carlo sequence is utilized to “sample” the exponential random variable. We provide a complexity theorem for this novel estimator and show that this theoretically and practically results in a variance reduction of the whole estimator.

Mathematics Subject Classification. 65C05, 65C10, 11K38, 65N30, 65N50.

Received May 25, 2023. Accepted May 20, 2024.

1. INTRODUCTION

During the last decade multilevel Monte Carlo methods and its variants as multiindex and multifidelity Monte Carlo have been successfully applied to reduce the costs of solving various uncertain problems (see *e.g.*, [1, 9, 15, 21, 23, 31]). The multilevel Monte Carlo estimator combines discretizations of a quantity of interest on a hierarchy of refinements in such a way that many samples on coarse refinement levels are combined with few samples on fine discretization levels to reduce the variance of the estimator when compared to a naive Monte Carlo approach. The multilevel Monte Carlo estimator is asymptotically optimal and reduces the costs to compute the quantity of interest considerably. This reduction comes at a cost: The quantity of interest has to be available on the hierarchy of discretizations and the variance of the details (difference of subsequent discretizations) has to decrease faster than the costs increase. Further, the weak and the strong error have to fulfill a certain ratio for the optimal cost reduction (see [15] for a detailed complexity theorem). However, in all cases the multilevel Monte Carlo estimator is biased (as is the singlelevel Monte Carlo estimator).

An unbiased variant was introduced in [10] with the continuous level Monte Carlo method. Here, the resolution levels are assumed to be continuous, resulting in a stochastic process describing the family of approximations of the quantity of interest. The refinement of each sample is determined by a (level) random variable. The samples are adaptively refined according to an *a-posteriori* error estimator. In practical terms the tail estimate of this

Keywords and phrases. Uncertainty quantification, Monte Carlo methods, continuous level Monte Carlo.

* *Dedicated to the memory of Assyr Abdulle.*

Institute of Applied Analysis and Numerical Simulation (IANS), University of Stuttgart, Allmandring 5b, 70569 Stuttgart, Germany.

**Corresponding author: cedric.beschle@ians.uni-stuttgart.de

level random variable is a crucial component in the performance of the estimator. In [7] a continuous level estimator was used to solve an elliptic problem with a discontinuous random coefficient and furthermore, the performance of the continuous level estimator was compared to its multilevel variant. The problem was chosen such that the continuous level estimator should have outperformed the multilevel method. Unfortunately the performance of the continuous level estimator is very sensitive to the tail estimate provided by the samples of the level random variable. To reduce this sensitivity the authors propose to use a quasi Monte Carlo sequence instead of i.i.d. samples of the level random variable to reduce the variance of the tail estimate and the whole estimator significantly. A similar idea was mentioned in a remark in [36] in the general framework of unbiased MLMC estimators. The so called quasi continuous level Monte Carlo estimator outperformed not only the continuous level but also the multilevel Monte Carlo estimator.

In this paper we investigate theoretically the quasi continuous level Monte Carlo method. We provide a complexity theorem which shows that the quasi continuous level Monte Carlo method has the same optimal complexity as the continuous level and multilevel estimator, but with a potentially lower variance and therefore an overall improved time to error performance as demonstrated by numerical experiments, where we apply both methods – continuous level Monte Carlo and quasi continuous level Monte Carlo – to an elliptic PDE problem with a log-Gauss random coefficient and compare their performance in different hyperparameter settings. As we only exchange the one-dimensional random variable with a quasi Monte Carlo sequence the cost increase is negligible. The quasi Monte Carlo sequence provides a more accurate tail estimate than pseudo-random numbers for the level random variable. We emphasise that we are not proposing a quasi (multilevel) Monte Carlo method to solve the uncertain PDE (as *e.g.* in [17, 18, 26, 27]). The performance of those methods depends on the dimensionality of the problem, the proposed quasi continuous level Monte Carlo method is robust in this aspect.

The paper is organized as follows: In Section 2 we briefly recap the continuous level Monte Carlo method and its complexity. We introduce quasi-random sequences and the notion of F -discrepancy in Section 3. We use the F -discrepancy in Section 4 to prove a complexity theorem for the quasi continuous level Monte Carlo method. In Section 5 we introduce the classical two-dimensional random elliptic PDE model we use as a test case. The coefficient is given by a log-Gauss random field and for the H^1 -norm as the quantity of interest we use a standard *a-posteriori* error estimator. For the performance comparison we estimate all parameters which are involved in the assumptions of the complexity theorem and compare the performances of the methods for different choices of the hyperparameters of the log-Gauss field based on respective theoretical upper bounds to the mean squared error of the methods stemming from the proofs of the complexity theorems and in a time to error performance over several simulation runs *via* a proposed algorithm.

2. CONTINUOUS LEVEL MONTE CARLO

We consider a stochastic model and denote by $\mathcal{Q}$ a real-valued quantity of interest of its solution. Denote by Q_L an approximation of $\mathcal{Q}$ by a discretization-based numerical scheme, for example a finite element method, to some resolution parameter $L \in \mathbb{N}_0$, *e.g.*, corresponding to the degrees of freedom (DOF) of a mesh. By assuming that $\mathbb{E}[Q_L] \rightarrow \mathbb{E}[\mathcal{Q}]$ for $L \rightarrow \infty$ $\mathbb{P}$ -almost surely, we are able to compute estimates $\hat{Q}_L^{\text{est}}$ for the mean value $\mathbb{E}[\mathcal{Q}]$ up to some desired accuracy with an average of independent approximation samples $(Q_L^{(k)}; k = 1, \dots, M)$ for $M \in \mathbb{N}$. The mean squared error (MSE) of an estimator $\hat{Q}_L^{\text{est}}$ for the mean value $\mathbb{E}[\mathcal{Q}]$ together with its decomposition into a variance and squared bias term is given by

$$\text{MSE} = \mathbb{E} \left[\left(\hat{Q}_L^{\text{est}} - \mathbb{E}[\mathcal{Q}] \right)^2 \right] = \mathbb{V} \left[\hat{Q}_L^{\text{est}} \right] + \left(\mathbb{E} \left[\hat{Q}_L^{\text{est}} - \mathcal{Q} \right] \right)^2. \quad (2.1)$$

In this work we investigate the recently developed quasi continuous level Monte Carlo method (QCLMC) [7], which is an improved version of the continuous level Monte Carlo method developed in [10]. The continuous level Monte Carlo (CLMC) method estimates the mean value of a quantity of interest with samplewise adaptive mesh hierarchies. This is realized by assuming a continuous resolution (level) $\ell \in \mathbb{R} > 0$ and a continuous family

of approximations $(Q(\ell); \ell \geq 0)$ of Q viewed as a stochastic process defined on a probability space $(\Omega, \mathcal{A}, \mathbb{P})$ with $\mathbb{E}[\|\frac{dQ}{d\ell}\|] \in L^1((0, \infty); \mathbb{R})$, such that $Q(\ell) \rightarrow Q$ as $\ell \rightarrow \infty$ $\mathbb{P}$ -almost surely. With these considerations, the CLMC estimator is defined by

$$\widehat{Q}_{0, L_{\max}}^{\text{CLMC}} := \frac{1}{M} \sum_{k=1}^M \int_0^{L_{\max}} \frac{1}{\mathbb{P}(L_r \geq \ell)} \left(\frac{dQ}{d\ell} \right)^{(k)}(\ell) \mathbb{1}_{[0, L_r^{(k)}]}(\ell) d\ell, \quad (2.2)$$

with deterministic maximal level $L_{\max} \in (0, +\infty]$, total sample number $M \in \mathbb{N}$ and a random variable L_r with finite expectation and $\mathbb{P}(L_r \geq \ell) > 0$ for all $\ell \in (0, L_{\max})$, that is independent of the stochastic process $(Q(\ell); \ell \geq 0)$. For each sample $k = 1, \dots, M$, the minimum of an i.i.d. copy $L_r^{(k)}$ of L_r and the predetermined $L_{\max}$ corresponds to the maximal computed resolution for this sample. Note, that the CLMC estimator (2.2) is defined as an estimator for the difference quantity $\mathbb{E}[Q - Q(0)]$, thus the indexing by $0, L_{\max}$, and in order to obtain an estimator for $\mathbb{E}[Q]$ it suffices to add an unbiased Monte Carlo estimator for $\mathbb{E}[Q(0)]$. The CLMC method is an unbiased estimator for $\mathbb{E}[Q(L_{\max}) - Q(0)]$ and in the case $L_{\max} = \infty$, it is an unbiased estimator for $\mathbb{E}[Q - Q(0)]$, i.e., $\mathbb{E}[\widehat{Q}_{0, \infty}^{\text{CLMC}}] = \mathbb{E}[Q - Q(0)]$ and thus, its MSE expansion (2.1) reduces to $\text{MSE}^{\text{CLMC}} = \mathbb{V}[\widehat{Q}_{\infty}^{\text{CLMC}}]$. As stated in [10], under the assumption that there exist positive constants $\alpha, \beta, \gamma, c_1, c_2, c_3$ such that for any $\ell > 0$ we have

$$\mathbb{E}\left[\frac{dQ(\ell)}{d\ell}\right] \leq c_1 e^{-\alpha\ell}, \quad \mathbb{V}\left[\frac{dQ(\ell)}{d\ell}\right] \leq c_2 e^{-\beta\ell}, \quad \frac{d\mathcal{C}[\ell]}{d\ell} \leq c_3 e^{\gamma\ell}, \quad (2.3)$$

where $\mathcal{C}[\ell]$ is the total accumulated cost to compute a sample of $Q(\ell)$ (cf. Rem. 2.2) and that $L_r \sim \text{Exp}(r)$ is exponentially distributed with $r \in [\min\{2\alpha, \beta, \gamma\}, \max\{\min\{\beta, 2\alpha\}, \gamma\}]$, then for any $\varepsilon \in (0, \frac{1}{e})$ there exist and $M \in \mathbb{N}$ and $\widetilde{C} > 0$ such that

$$\text{MSE}^{\text{CLMC}} \leq \varepsilon^2 \quad \text{and} \quad \mathcal{C}\left[\widehat{Q}_{0, L_{\max}}^{\text{CLMC}}\right] \leq \widetilde{C} \varepsilon^{-2 - \max\{0, \frac{\gamma - \min\{\beta, 2\alpha\}}{\alpha}\}} |\log(\varepsilon)|^{\delta_{r, \beta} + \delta_{r, 2\alpha} + \delta_{r, \gamma}} \quad (2.4)$$

where δ is the Kronecker delta. In the case $\min\{\beta, 2\alpha\} > \gamma$ and $r \in (\gamma, \min\{\beta, 2\alpha\})$ with $L_{\max} = \infty$ this reduces to

$$\text{MSE}^{\text{CLMC}} \leq \varepsilon^2 \quad \text{and} \quad \mathcal{C}\left[\widehat{Q}_{0, \infty}^{\text{CLMC}}\right] \leq \widetilde{C} \varepsilon^{-2}.$$

Next, we explain how CLMC generalizes MLMC. For any sample $1 \leq k \leq M$, suppose that $(Q_j^{(k)}; j \geq 1)$ denotes a countable sequence of approximations of $Q^{(k)}$ at levels $(\ell_j^{(k)}; j \geq 1)$. We choose a linear interpolation for the derivative samples

$$\left(\frac{dQ}{d\ell} \right)^{(k)}(\ell) := \frac{Q_j^{(k)} - Q_{j-1}^{(k)}}{\ell_j^{(k)} - \ell_{j-1}^{(k)}} \quad \text{for} \quad \ell \in \left(\ell_{j-1}^{(k)}, \ell_j^{(k)} \right], \quad (2.5)$$

which yields the estimator

$$\widehat{Q}_{0, L_{\max}}^{\text{CLMC}} = \frac{1}{M} \sum_{k=1}^M \sum_{j=1}^{J^{(k)}} \int_{\ell_{j-1}^{(k)}}^{\tilde{\ell}_j^{(k)}} \frac{1}{\mathbb{P}(L_r \geq \ell)} d\ell \frac{Q_j^{(k)} - Q_{j-1}^{(k)}}{\ell_j^{(k)} - \ell_{j-1}^{(k)}}, \quad (2.6)$$

with

$$J^{(k)} := \min\{j \geq 1 : \ell_j^{(k)} \geq L_r^{(k)} \wedge L_{\max}\}, \quad \tilde{\ell}_j^{(k)} := \ell_j^{(k)} \wedge L_r^{(k)} \wedge L_{\max}.$$

Note, that one is free to choose the level values $(\ell_j^{(k)}; k = 1, \dots, M, j \geq 1)$, in particular independent of the maximally drawn levels $(L_r^{(k)}; k = 1, \dots, M)$. A suitable choice in case of adaptive refinements is provided in Section 5.2. Setting $\ell_j^{(k)} = j$ for $j \in \mathbb{N}$ and all $k = 1, \dots, M$ to restrict the estimator to the integer level

framework and choosing $\mathbb{P}(L_r \geq j)$ to be a discrete distribution over the levels, that is constant over $(j-1, j)$, reduces the CLMC estimator (2.6) to, cf. [10],

$$\widehat{Q}_{0, L_{\max}}^{\text{CLMC}} = \frac{1}{M} \sum_{k=1}^M \sum_{j=1}^{J^{(k)}} \frac{1}{\mathbb{P}(L_r \geq j)} \left(Q_j^{(k)} - Q_{j-1}^{(k)} \right). \quad (2.7)$$

By setting $M_j := M\mathbb{P}(L \geq j)$ the estimator (2.7) is directly connected to the (unbiased) estimator

$$\widehat{Q}_{0, L_{\max}}^{\text{CLMC}} = \sum_{k=1}^M \sum_{j=1}^{J^{(k)}} \frac{1}{M_j} \left(Q_j^{(k)} - Q_{j-1}^{(k)} \right)$$

introduced by Rhee and Glynn [33]. Furthermore, with $0 < L_{\max} \in \mathbb{N}$, this may be expressed as, cf. [16],

$$\widehat{Q}_{0, L_{\max}}^{\text{MLMC}} := \sum_{j=1}^{L_{\max}} \frac{1}{M_j} \sum_{k=1}^{M_j} \left(Q_j^{(k)} - Q_{j-1}^{(k)} \right),$$

which is the formula of the standard multilevel Monte Carlo estimator, where $L_{\max}$ is the maximal level and the sample numbers M_j on each level $j = 1, \dots, L_{\max}$ are not probabilistic, but deterministic.

Remark 2.1. A benefit of the exponential distribution is that it naturally results in a decay in the number of samples with rising level value similar to MLMC. But, as demonstrated in the derivation of the MLMC estimator, the level random variable L_r is not restricted to the exponential distribution, but merely has to be independent of the stochastic process $(Q(\ell); \ell \geq 0)$ and have finite expectation with $\mathbb{P}(L_r \geq \ell) > 0$ for all $\ell \in (0, L_{\max})$.

Remark 2.2. The parameters c_1, α, c_2, β and c_3, γ from (2.3) depend not only on the approximations $Q_j^{(k)} = Q(\ell_j^{(k)})$ at refinement $j \in \mathbb{N}$ for $k = 1, \dots, M$ but also on the definition of the derivative $\frac{dQ}{d\ell}^{(k)}$ and the samplewise level $\ell_j^{(k)}$ in a practical setting, see, e.g., Section 5.2. The assumption on the bias and variance decay of the derivative quantity $\frac{dQ}{d\ell}$ and the assumption for the cost growth, which is an assumption for the rate of change of the total cost $d\mathcal{C}[\ell]$ to compute an approximation of $Q(\ell)$, scale with the change in the level $d\ell$ in a way, that changes the constants c_1, c_2 and c_3 correspondingly. In Theorem 2.3 of [10] the assumption on the cost growth is $\mathcal{C}[\ell] \leq c_3 e^{\gamma\ell}$, where $\mathcal{C}[\ell]$ is the cost to compute one sample of $Q(\ell)$. This assumption does not scale c_3 accordingly and allows to construct a practical CLMC estimator by a specific definition for the samplewise level $\ell_j^{(k)}$ at refinement $j \in \mathbb{N}$ for $k = 1, \dots, M$ such that the theoretical upper bound to the estimator's cost in equation (2.4) becomes arbitrarily small.

The key difference between the CLMC estimator and the QCLMC estimator is the choice of how to compute $L_r^{(k)}$. While in CLMC $L_r^{(k)}$ are i.i.d. copies of the random variable L_r for $k = 1, \dots, M$, in QCLMC we choose $L_r^{(k)}$ to be a deterministic quasi-random sequence for $k = 1, \dots, M$ yielding a better approximation of the underlying tail distribution *via* their improved F -discrepancy convergence. These concepts are introduced in the next section and the improvement in the approximation is demonstrated.

3. QUASI-RANDOM SEQUENCES AND F -DISCREPANCY

The discrepancy of a set of points P consisting of $x^{(1)}, \dots, x^{(M)} \in [0, 1]^s$ for $M \in \mathbb{N}$ and $s \in \mathbb{N}$ is given by

$$D_M(\mathcal{B}; P) = \sup_{B \in \mathcal{B}} \left| \frac{1}{M} \sum_{k=1}^M \mathbb{1}_B(x^{(k)}) - \lambda(B) \right|, \quad (3.1)$$

cf., e.g., [28], where λ is the Lebesgue measure and $\mathcal{B}$ a non-empty family of Lebesgue-measurable subsets of $[0, 1]^s$. For simplicity and since it fits our considerations we assume $s = 1$. Quasi-random sequences are numbers $x^{(1)}, \dots, x^{(M)} \in [0, 1]$ specifically designed such that the discrepancy converges to zero at a much faster rate than for pseudo-random number sequences, i.e.,

$$D_M(\mathcal{B}; P) \leq c_{\text{disc}} M^{\kappa-1} \text{ for some } 0 \leq \kappa < \frac{1}{2} \text{ and } c_{\text{disc}} > 0 \text{ independent of } M \in \mathbb{N}. \quad (3.2)$$

This is no probability convergence statement, because quasi-random numbers are essentially deterministic. The specific choice of $\mathcal{B}$ in the discrepancy definition (3.1) as the family of all subintervals $[0, x] \subset [0, 1]$, where $x \in (0, 1)$, leads to the star-discrepancy

$$D_M^*(P) = \sup_{[0, x] \subset [0, 1]} \left| \frac{1}{M} \sum_{k=1}^M \mathbb{1}_{[0, x]}(x^{(k)}) - x \right| = \sup_{x \in (0, 1)} \left| \frac{1}{M} \sum_{k=1}^M \mathbb{1}_{[0, x]}(x^{(k)}) - x \right|.$$

For a cumulative distribution function (CDF) $F : \mathbb{R} \rightarrow [0, 1]$, the empirical CDF of M samples $\tilde{x}^{(1)}, \dots, \tilde{x}^{(M)} \in \mathbb{R}$ is given by

$$F_M(x) := \frac{1}{M} \sum_{k=1}^M \mathbb{1}_{\{\tilde{x}^{(k)} \leq x\}} = \frac{1}{M} \sum_{k=1}^M \mathbb{1}_{[\tilde{x}^{(k)}, \infty)}(x) = \frac{1}{M} \sum_{k=1}^M \mathbb{1}_{(-\infty, x]}(\tilde{x}^{(k)}).$$

The F -discrepancy of P is defined by, cf. [13],

$$D_{F,P} = \sup_{x \in \mathbb{R}} |F_M(x) - F(x)|.$$

Considering the uniform distribution on $[0, 1]$ with CDF

$$F_U(x) = \begin{cases} 0 & \text{for } x < 0, \\ x & \text{for } 0 \leq x \leq 1, \\ 1 & \text{for } x > 1, \end{cases}$$

we observe that it holds

$$D_{F_U,P} = D_M^*(P).$$

Now, let $F_Y : \mathbb{R} \rightarrow [0, 1]$, be a continuous distribution function to a random variable Y , where the inverse F_Y^{-1} exists, is non-decreasing and continuous as well. Let $P_Y := \{y^{(k)}; k = 1, \dots, M\}$ be a sequence of points obtained through inverse sampling of quasi-random numbers $x^{(k)}$ via the inverse CDF F_Y^{-1} , i.e., $y^{(k)} = F_Y^{-1}(x^{(k)})$ for all $k = 1, \dots, M$. Further, assume that there exists $x \in [0, 1]$ such that $F_Y(y) = x$ and $F_Y^{-1}(x) = y$ for every $y \in \mathbb{R}$. With these assumptions we compute

$$\begin{aligned} D_{F_Y, P_Y} &= \sup_{y \in \mathbb{R}} |F_{M,Y}(y) - F_Y(y)| = \sup_{y \in \mathbb{R}} \left| \frac{1}{M} \sum_{k=1}^M \mathbb{1}_{\{y^{(k)} \leq y\}} - F_Y(y) \right| \\ &= \sup_{y \in \mathbb{R}} \left| \frac{1}{M} \sum_{k=1}^M \mathbb{1}_{\{F_Y^{-1}(x^{(k)}) \leq y\}} - F_Y(y) \right| = \sup_{x \in (0, 1)} \left| \frac{1}{M} \sum_{k=1}^M \mathbb{1}_{\{F_Y^{-1}(x^{(k)}) \leq F_Y^{-1}(x)\}} - x \right| \\ &= \sup_{x \in (0, 1)} \left| \frac{1}{M} \sum_{k=1}^M \mathbb{1}_{\{x^{(k)} \leq x\}} - x \right| = D_{F_U,P} = D_M^*(P). \end{aligned}$$

Thus, the F -discrepancy for a continuous random variable with continuous inverse is equal to the star discrepancy, cf. [13]. The same holds true when considering the F -discrepancy of the tail distribution function

$T_Y(y) = 1 - F_Y(y)$, because for the empirical tail distributions we have

$$\frac{1}{M} \sum_{k=1}^M \mathbb{1}_{[-\infty, y^{(k)}]}(y) = T_{M,Y}(y) = 1 - F_{M,Y}(y).$$

This leads to the following result, which is essential for the upcoming complexity analysis of the QCLMC method.

Lemma 3.1. *Let $(\Omega, \mathcal{A}, \mathbb{P})$ be a complete probability space and $Y : \Omega \rightarrow \mathbb{R}$ a real-valued random variable with continuous distribution function F_Y and a continuous inverse distribution function F_Y^{-1} . For the distribution function it holds $F_Y(y) = \mathbb{P}(Y \leq y)$ and for the tail distribution function $T_Y(y) = 1 - F_Y(y) = \mathbb{P}(Y \geq y)$. Let $y^{(k)}$ be a sequence generated via the inverse transformation*

$$y^{(k)} := F_Y^{-1}(x^{(k)}),$$

from a sequence $(x^{(k)}; k = 1, \dots, M)$ distributed in $[0, 1)$ that satisfies the star-discrepancy convergence property $D_M^((x^{(k)}; k = 1, \dots, M)) \leq c_{\text{disc}} M^{\kappa-1}$ for some $\kappa \geq 0$ and $c_{\text{disc}} > 0$ independent of $M \in \mathbb{N}$. Then, the following convergence result for estimating the tail distribution T_Y of the random variable Y via the empirical tail distribution $T_{M,Y}$ holds:*

$$\sup_{y \in \mathbb{R}} \left| \frac{1}{M} \sum_{k=1}^M \mathbb{1}_{[-\infty, y^{(k)}]}(y) - \mathbb{P}(Y \geq y) \right| \leq c_{\text{disc}} M^{\kappa-1}, \quad (3.3)$$

for $\kappa \geq 0$ and $c_{\text{disc}} > 0$ from the star-discrepancy convergence property of $(x^{(k)}; k = 1, \dots, M)$.

Remark 3.2. The assumptions in Lemma 3.1 hold for the specific case of an exponentially distributed random variable $Y \equiv L_r \sim \text{Exp}(r)$ for some $r > 0$. The distribution function is given by $\mathbb{P}(L_r \leq \ell) = 1 - e^{-r\ell}$ and the tail distribution function is $\mathbb{P}(L_r \geq \ell) = e^{-r\ell}$. The samples $(L_r^{(k)}; k \in \mathbb{N})$ are generated via the inverse transformation

$$L_r^{(k)} := -\frac{\ln(1 - x^{(k)})}{r} \quad (3.4)$$

from a sequence $(x^{(k)}; k = 1, \dots, M)$ distributed in $[0, 1)$. Thus, Lemma 3.1 bounds the error of the tail estimate of an exponentially-distributed random variable, approximated by an empirical tail estimate. An illustration of this for an empirical tail estimate obtained by a quasi-random sequence compared to a pseudo-random sequence is given in Figure 1.

Remark 3.3. For the F -discrepancy result of Lemma 3.1 the full rate of convergence with $\kappa = 0$ is achieved in one dimension by Sobol sequences or Hammersley points, among others, cf. [11] for an overview. The sequence of numbers $(x^{(k)}; k = 1, \dots, M)$ is not restricted to quasi-random numbers in general. Any sequence from the interval $[0, 1)$ may be used as long as the F -discrepancy convergence property with $\kappa = 0$ is satisfied. A sequence of i.i.d. copies of a $[0, 1]$ -uniformly distributed random variable for $x^{(k)}$ only yields $\kappa = \frac{1}{2}$, cf. Figures 1. The grid points $x^{(k)} = \frac{2k-1}{2M}$ for $k = 1, \dots, M$ have F -discrepancy $\frac{1}{2}M^{-1}$, cf. Corollary 1.2 of [25] and Lemma 1 of [13], which is the best achievable discrepancy in one dimension. However, this grid sequence is not nested and can not be reused for growing values of M . A great benefit of using quasi-random numbers is the possibility to sample one point after another in case the number of samples M is not known *a-priori*, e.g., in an on-the-fly-type algorithm.

Another great benefit of using quasi-random numbers for the low-discrepancy sequence is the possibility of creating independent low-discrepancy sequences *via* randomization, e.g., for Sobol numbers *via* Owen scrambling

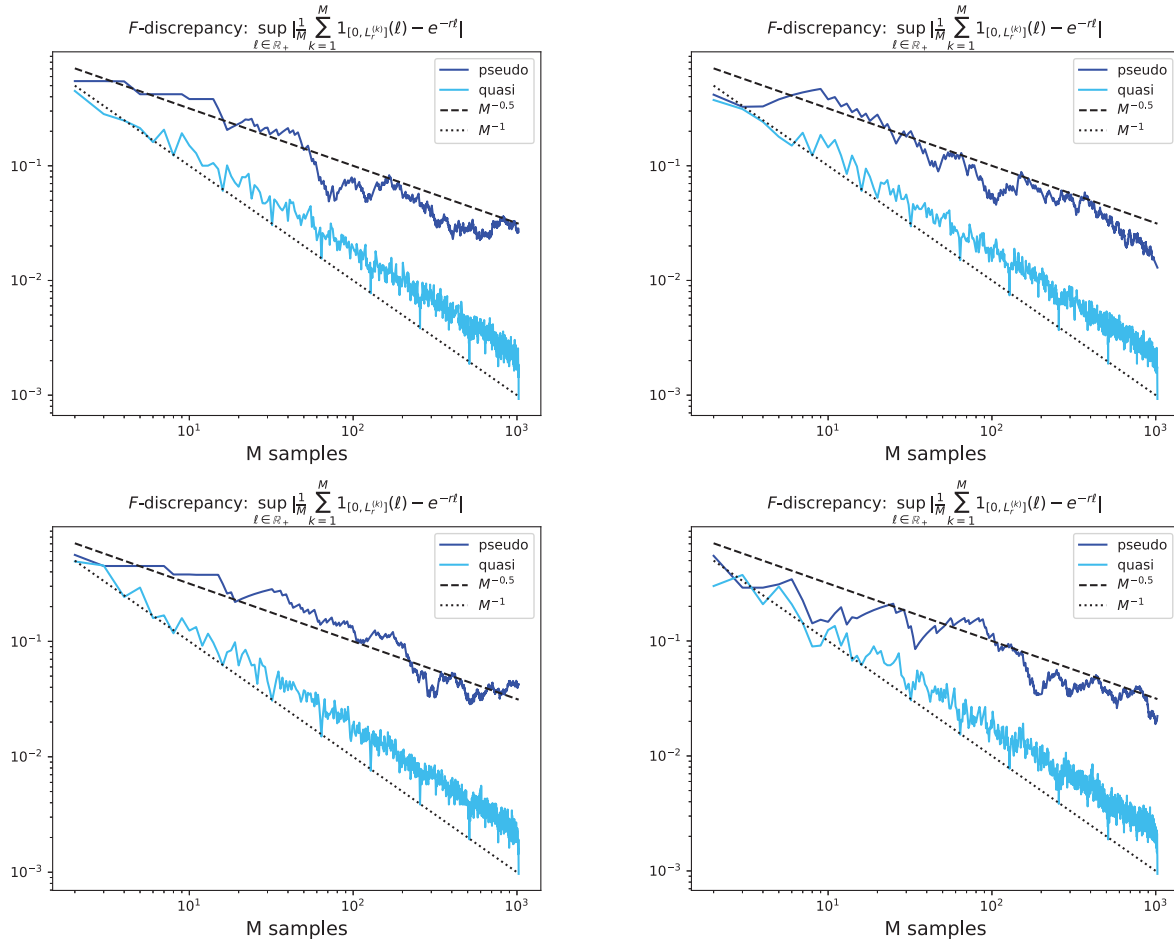


FIGURE 1. Demonstration of the convergence result of Lemma 3.1 for $r = 1.3$ and four independent runs (different seeds) of quasi-random Sobol numbers [34] with Owen scrambling, cf. [29, 30] generated through the scipy library [37], and pseudo-random numbers generated with the numpy library [22]. The quasi-random Sobol numbers have their optimal discrepancy for powers of two (location of downward spikes in the light blue lines), but we observe that for values in between powers of two, the discrepancy still converges with rate one, *i.e.*, $\kappa = 0$.

[29, 30], to obtain independent QCLMC estimators to estimate the MSE over several QCLMC runs as done in the numerical experiments in Section 5.4.

Furthermore, the continuous level framework may be extended to multiindex Monte Carlo (MIMC), cf. [21] and Section 5 from [10] for details. In this case, the level variable follows a multivariate probability distribution and quasi-random sequences with an optimal discrepancy property in higher dimensions are beneficial for such an extension of QCLMC. They obtain merely an additional logarithmic dependence on the dimension in the F -discrepancy convergence result.

4. QUASI CONTINUOUS LEVEL MONTE CARLO METHOD

As for the CLMC method in Section 2 we assume that for the level parameter $\ell > 0$ we have approximations to the quantity of interest $(Q(\ell); \ell > 0)$, the deterministic maximal level is given by $L_{\max} \in (0, \infty]$, $M \in \mathbb{N}$ is

the total number of samples and L_r is a random variable with finite expectation, $\mathbb{P}(L_r \geq \ell) > 0$ and that is independent of the stochastic process $(Q(\ell))_{\ell \geq 0}$. Then, the QCLMC estimator is defined by

$$\widehat{Q}_{0, L_{\max}}^{\text{QCLMC}} = \frac{1}{M} \sum_{k=1}^M \int_0^{L_{\max} \wedge \bar{L}} \frac{1}{\mathbb{P}(L_r \geq \ell)} \left(\frac{dQ}{d\ell} \right)^{(k)} \mathbb{1}_{[0, L_r^{(k)}]}(\ell) d\ell, \quad (4.1)$$

where furthermore and most importantly $(L_r^{(k)}; k = 1, \dots, M)$ is a deterministic sequence obtained *via* inverse transformation (see Lem. 3.1) with $\kappa = 0$ and $\bar{L} := \max\{L_r^{(k)}; k = 1, \dots, M\}$. Note, that the sequence $(L_r^{(k)}; k = 1, \dots, M)$ in this estimator does not consist of i.i.d. copies of the random variable L_r anymore.

For showing the unbiasedness result of the QCLMC estimator and the complexity theorem we choose $L_r \sim \text{Exp}(r)$ with parameter $r > 0$ as for the CLMC estimator and note that this is also not the only choice in QCLMC. The distribution generally has to satisfy the assumptions in Remark 2.1 and in Lemma 3.1.

Since we deal with a deterministic sequence for $(L_r^{(k)}; k = 1, \dots, M)$ the use of any distributional properties of random sequences is not reasonable here anymore. Thus, additionally to the expectation $\mathbb{E}[\widehat{Q}_{0, L_{\max}}^{\text{QCLMC}}]$ of the estimator we consider the limit $M \rightarrow \infty$, see, e.g., Chapter 3 of [14], when investigating the unbiasedness of the QCLMC estimator. Note, that the samples $(\frac{dQ}{d\ell}^{(k)}, k = 1, \dots, M)$ are still i.i.d. copies of the random variable $\frac{dQ}{d\ell}$ in QCLMC.

Proposition 4.1. *Assume that $\bar{L} = \ln(\tilde{c}^{-\frac{1}{r}} M^{\frac{1-\kappa}{r}})$ with $0 \leq \kappa < 1$, $r > 0$ and a constant $\tilde{c} > 0$ independent of $M \in \mathbb{N}$ and suppose there exist positive constants α and c_1 such that for any $\ell > 0$*

$$\left| \mathbb{E} \left[\frac{dQ(\ell)}{d\ell} \right] \right| \leq c_1 e^{-\alpha \ell}. \quad (4.2)$$

Then, in the limit $M \rightarrow \infty$, the QCLMC estimator (4.1) is an unbiased estimator for $\mathbb{E}[Q(L_{\max}) - Q(0)]$, i.e.,

$$\lim_{M \rightarrow \infty} \mathbb{E}[\widehat{Q}_{0, L_{\max}}^{\text{QCLMC}}] = \mathbb{E}[Q(L_{\max}) - Q(0)].$$

If $L_{\max} = \infty$ it holds $Q(L_{\max}) = Q(\infty) = \mathcal{Q}$ and the QCLMC estimator is an unbiased estimator to the real quantity of interest $\mathbb{E}[\mathcal{Q} - Q(0)]$.

Proof. First, we compute an equality for the expectation of the estimator (4.1) by adding zero in a suitable way and using that $\mathbb{P}(L_r \geq \ell) = e^{-r\ell}$

$$\begin{aligned} \mathbb{E}[\widehat{Q}_{0, L_{\max}}^{\text{QCLMC}}] &= \mathbb{E} \left[\frac{1}{M} \sum_{k=1}^M \int_0^{L_{\max} \wedge \bar{L}} e^{r\ell} \left(\frac{dQ}{d\ell} \right)^{(k)}(\ell) \mathbb{1}_{[0, L_r^{(k)}]}(\ell) d\ell \right] \\ &= \int_0^{L_{\max} \wedge \bar{L}} e^{r\ell} \mathbb{E} \left[\frac{dQ(\ell)}{d\ell} \right] \frac{1}{M} \sum_{k=1}^M \mathbb{1}_{[0, L_r^{(k)}]}(\ell) d\ell \\ &= \int_0^{L_{\max} \wedge \bar{L}} \left(e^{r\ell} \frac{1}{M} \sum_{k=1}^M \mathbb{1}_{[0, L_r^{(k)}]}(\ell) - \frac{e^{-r\ell}}{e^{-r\ell}} + 1 \right) \mathbb{E} \left[\frac{dQ(\ell)}{d\ell} \right] d\ell \\ &= \int_0^{L_{\max} \wedge \bar{L}} e^{r\ell} \left(\frac{1}{M} \sum_{k=1}^M \mathbb{1}_{[0, L_r^{(k)}]}(\ell) - e^{-r\ell} \right) \mathbb{E} \left[\frac{dQ(\ell)}{d\ell} \right] d\ell + \int_0^{L_{\max} \wedge \bar{L}} \mathbb{E} \left[\frac{dQ(\ell)}{d\ell} \right] d\ell =: \text{I} + \text{II}. \end{aligned}$$

We bound the integrand of I from below and above. By the F -discrepancy property (3.3) and Assumption (4.2) it holds for every $\ell > 0$

$$-c_1 c_{\text{disc}} M^{\kappa-1} e^{(r-\alpha)\ell} \leq e^{r\ell} \left(\frac{1}{M} \sum_{k=1}^M \mathbb{1}_{[0, L_r^{(k)}]}(\ell) - e^{-r\ell} \right) \mathbb{E} \left[\frac{dQ(\ell)}{d\ell} \right] \leq c_1 c_{\text{disc}} M^{\kappa-1} e^{(r-\alpha)\ell}.$$

We integrate the lower and upper bound from zero to $L_{\max} \wedge \bar{L}$ to obtain

$$\pm c_1 c_{\text{disc}} M^{\kappa-1} \int_0^{L_{\max} \wedge \bar{L}} e^{(r-\alpha)\ell} d\ell = \pm c_1 c_{\text{disc}} M^{\kappa-1} \begin{cases} \frac{1}{r-\alpha} \left(e^{(r-\alpha)(L_{\max} \wedge \bar{L})} - 1 \right) & \text{for } r \neq \alpha, \\ L_{\max} \wedge \bar{L} & \text{for } r = \alpha. \end{cases}$$

Note that $\lim_{M \rightarrow \infty} \bar{L} = \lim_{M \rightarrow \infty} \ln(\tilde{c}^{-\frac{1}{r}} M^{\frac{1-\kappa}{r}}) = \infty$ and thus, in the limit $M \rightarrow \infty$, for finite $L_{\max} < \infty$ the minimum $L_{\max} \wedge \bar{L}$ is attained for $L_{\max}$ and we trivially obtain $\lim_{M \rightarrow \infty} \text{I} = 0$ by the squeeze theorem and $\lim_{M \rightarrow \infty} \text{II} = \mathbb{E}[Q(L_{\max}) - Q(0)]$. If $L_{\max} = \infty$ we compute in the case $r = \alpha$

$$\lim_{M \rightarrow \infty} \pm c_1 c_{\text{disc}} M^{\kappa-1} (L_{\max} \wedge \bar{L}) = \lim_{M \rightarrow \infty} \pm c_1 c_{\text{disc}} M^{\kappa-1} \ln\left(\tilde{c}^{-\frac{1}{r}} M^{\frac{1-\kappa}{r}}\right) = 0,$$

for all $0 \leq \kappa < 1$. For $r \neq \alpha$ we obtain

$$\begin{aligned} \lim_{M \rightarrow \infty} \pm c_1 c_{\text{disc}} M^{\kappa-1} \frac{1}{r-\alpha} \left(e^{(r-\alpha)L_{\max} \wedge \bar{L}} - 1 \right) &= \lim_{M \rightarrow \infty} \pm c_1 c_{\text{disc}} M^{\kappa-1} \frac{1}{r-\alpha} \left(e^{(r-\alpha) \ln\left(\tilde{c}^{-\frac{1}{r}} M^{\frac{1-\kappa}{r}}\right)} - 1 \right) \\ &= \lim_{M \rightarrow \infty} \pm \left(\frac{c_1 c_{\text{disc}} \tilde{c}^{-\frac{r-\alpha}{r}}}{r-\alpha} M^{\kappa-1} M^{\frac{(r-\alpha)(1-\kappa)}{r}} - \frac{c_1 c_{\text{disc}}}{r-\alpha} M^{\kappa-1} \right) \\ &= \lim_{M \rightarrow \infty} \pm \left(\frac{c_1 c_{\text{disc}} \tilde{c}^{-\frac{r-\alpha}{r}}}{r-\alpha} M^{(\kappa-1)\frac{\alpha}{r}} - \frac{c_1 c_{\text{disc}}}{r-\alpha} M^{\kappa-1} \right) = 0, \end{aligned}$$

for all $0 \leq \kappa < 1$ and $\alpha > 0$. The squeeze theorem again yields $\lim_{M \rightarrow \infty} \text{I} = 0$ and $\lim_{M \rightarrow \infty} \text{II} = \mathbb{E}[Q(\infty) - Q(0)] = \mathbb{E}[Q - Q(0)]$. Thus, in any case we obtain in the limit $M \rightarrow \infty$ the final result

$$\lim_{M \rightarrow \infty} \mathbb{E}[\widehat{Q}_{0, L_{\max}}^{\text{QCLMC}}] = \mathbb{E}[Q(L_{\max}) - Q(0)],$$

for all $L_{\max} \in (0, \infty]$. □

Remark 4.2. For any sequence $x^{(k)}$ in $[0, 1)$ satisfying the discrepancy convergence property (3.2) for the star-discrepancy, it holds $\max\{x^{(k)}; k = 1, \dots, M\} = 1 - \tilde{c}M^{\kappa-1}$ for a constant $0 < \tilde{c} \leq c_{\text{disc}}$ independent of $M \in \mathbb{N}$ and we obtain for the sequence $L_r^{(k)}$ from Remark 3.2 that

$$\begin{aligned} \bar{L} := \max\{L_r^{(k)}; k = 1, \dots, M\} &= -\frac{\ln(1 - \max\{x^{(k)}; k = 1, \dots, M\})}{r} \\ &= -\frac{\ln(1 - (1 - \tilde{c}M^{\kappa-1}))}{r} = \ln\left(\tilde{c}^{-\frac{1}{r}} M^{\frac{1-\kappa}{r}}\right), \end{aligned} \tag{4.3}$$

showing that the assumption from Proposition 4.1 is satisfied.

Next, we prove a complexity theorem for the new QCLMC estimator with explicit treatment of the $(L_r^{(k)}; k = 1, \dots, M)$ as a deterministic quasi-random sequence with $\kappa = 0$.

Theorem 4.3 (QCLMC – complexity theorem). *Denote by $(Q(\ell); \ell \geq 0)$ a stochastic process defined on a probability space $(\Omega, \mathcal{A}, \mathbb{P})$ with $\mathbb{E}\left[\left|\frac{dQ}{d\ell}\right|\right] \in L^1((0, \infty); \mathbb{R})$, corresponding to a family of numerical approximations of $\mathcal{Q}$ such that $Q(\ell) \rightarrow \mathcal{Q}$ as $\ell \rightarrow \infty$ $\mathbb{P}$ -almost surely. Suppose there exist positive constants $\alpha, \beta, \gamma, c_1, c_2, c_3$ such that for any $\ell > 0$:*

$$\left| \mathbb{E}\left[\frac{dQ(\ell)}{d\ell}\right] \right| \leq c_1 e^{-\alpha\ell}, \tag{4.4a}$$

$$\mathbb{V}\left[\frac{dQ(\ell)}{d\ell}\right] \leq c_2 e^{-\beta\ell}, \quad (4.4b)$$

$$\frac{dC[\ell]}{d\ell} \leq c_3 e^{\gamma\ell}, \quad (4.4c)$$

where $C[\ell]$ is the total accumulated cost to compute a sample of $Q(\ell)$. Further, let $(L_r^{(k)}; k = 1, \dots, M)$ be a deterministic quasi-random sequence obtained by inverse transformation (see Lem. 3.1 with $\kappa = 0$), and let $r \in [\min\{\beta, 2\alpha, \gamma\}, \max\{\min\{\beta, 2\alpha\}, \gamma\}]$. Then, there exist $L_{\max} \in (0, \infty]$ and $M \in \mathbb{N}$, such that for any $\varepsilon \in (0, \frac{1}{e})$ it holds

$$\text{MSE}^{\text{QCLMC}} \leq \varepsilon^2 \quad \text{and} \quad \mathcal{C}\left[\widehat{Q}_{0, L_{\max}}^{\text{QCLMC}}\right] \leq C\varepsilon^{-2 - \max\{0, \frac{\gamma - \min\{\beta, 2\alpha\}}{\alpha}\}} |\ln(\varepsilon)|^{\delta_{r, \beta + \delta_{r, \gamma}}},$$

where δ is the Kronecker delta and $C > 0$ is independent of $L_{\max}$, M and ε .

Before stating the proof let us note, that with an appropriate choice of r , the complexity of the QCLMC estimator given in Theorem 4.3 is the same as for the CLMC estimator given in equation (2.4) and the MLMC estimator, cf. [15, 16], but with a potentially lower constant in the upper bound to the cost. This potential is investigated for the CLMC and QCLMC method in the numerical experiments in Section 5.

Proof. To deliberately use the F -discrepancy property of the quasi-random sequence, the proof is based on the standard decomposition of the MSE

$$\text{MSE}_{0, L_{\max}}^{\text{QCLMC}} = \mathbb{V}\left[\widehat{Q}_{0, L_{\max}}^{\text{QCLMC}}\right] + \mathbb{E}\left[\widehat{Q}_{0, L_{\max}}^{\text{QCLMC}} - (Q - Q(0))\right]^2,$$

into a variance and squared bias term, which is to be bounded by ε^2 for a given $0 < \varepsilon < e^{-1}$. The proof is split into three parts: In the first part we compute the squared bias of the QCLMC estimator and divide it into a term depending only on $L_{\max} \wedge \bar{L}$ and terms depending on $L_{\max} \wedge \bar{L}$ and M , but each of these terms again depends differently on the total sample size M . We choose $L_{\max} \wedge \bar{L}$ to bound the first term by $\frac{\varepsilon^2}{2}$ and M to bound the remaining terms by $\frac{\varepsilon^2}{4}$. Then, in the second part we bound the variance of the QCLMC estimator in terms of $\frac{\varepsilon^2}{4}$ by an appropriate choice of M . Let us note here, that in each of these bounds appear two types of terms again, each depending differently on the total sample size M . Finally, in the third part we bound the total cost of the QCLMC estimator with the aggregated choices of M from the first two parts, finishing the proof. Optimized MSE splits based on the problem parameters $\alpha, \beta, \gamma, c_1, c_2, c_3$ may as well be obtained, but are omitted for simplicity of notation. A key concept in the proof is the insertion of a zero in the integral quantities (as done in the proof of Prop. 4.1) in order to use the F -discrepancy property of the quasi-random sequence. Utilizing that $L_r \sim \text{Exp}(r)$, we the estimate

$$\begin{aligned} \left| \mathbb{E}\left[\widehat{Q}_{0, L_{\max}}^{\text{QCLMC}} - (Q - Q(0))\right] \right| &\leq \left| \mathbb{E}[Q - Q(L_{\max} \wedge \bar{L})] \right| + \left| \mathbb{E}\left[\widehat{Q}_{0, L_{\max}}^{\text{QCLMC}}\right] - \mathbb{E}[Q(L_{\max} \wedge \bar{L}) - Q(0)] \right| \\ &\leq \int_{L_{\max} \wedge \bar{L}}^{\infty} \left| \mathbb{E}\left[\frac{dQ(\ell)}{d\ell}\right] \right| d\ell + \frac{c_{\text{disc}}}{M} \int_0^{L_{\max} \wedge \bar{L}} e^{r\ell} \left| \mathbb{E}\left[\frac{dQ(\ell)}{d\ell}\right] \right| d\ell \\ &\leq c_1 \int_{L_{\max} \wedge \bar{L}}^{\infty} e^{-\alpha\ell} d\ell + \frac{c_{\text{disc}}}{M} c_1 \int_0^{L_{\max} \wedge \bar{L}} e^{(r-\alpha)\ell} d\ell \\ &= \frac{c_1}{\alpha} e^{-\alpha(L_{\max} \wedge \bar{L})} + \frac{c_{\text{disc}} c_1}{M} \cdot \begin{cases} \frac{1}{r-\alpha} (e^{(r-\alpha)(L_{\max} \wedge \bar{L})} - 1) & \text{for } r \neq \alpha, \\ L_{\max} \wedge \bar{L} & \text{for } r = \alpha, \end{cases} \end{aligned} \quad (4.5)$$

with explicit dependence on $c_{\text{disc}}, c_1, \alpha, r, M, \bar{L}$ and $L_{\max}$. We restrict to the case $r = \alpha$ and compute for the bound of the squared bias

$$\left| \mathbb{E}\left[\widehat{Q}_{0, L_{\max}}^{\text{QCLMC}} - (Q - Q(0))\right] \right|^2 \leq \frac{c_1^2}{\alpha^2} e^{-2\alpha(L_{\max} \wedge \bar{L})} + \frac{c_{\text{disc}}}{M} \frac{2c_1^2}{\alpha} (L_{\max} \wedge \bar{L}) e^{-\alpha(L_{\max} \wedge \bar{L})} + \frac{c_{\text{disc}}^2}{M^2} c_1^2 (L_{\max} \wedge \bar{L})^2.$$

Next, we consider the other case $r \neq \alpha$, which yields

$$\begin{aligned} \left| \mathbb{E} \left[\widehat{Q}_{0, L_{\max}}^{\text{QCLMC}} - (\mathcal{Q} - Q(0)) \right] \right|^2 &\leq \frac{c_1^2}{\alpha^2} e^{-2\alpha(L_{\max} \wedge \bar{L})} + \frac{c_{\text{disc}}}{M} \frac{2c_1^2}{\alpha} \frac{1}{r - \alpha} \left(e^{(r-2\alpha)(L_{\max} \wedge \bar{L})} - e^{-\alpha(L_{\max} \wedge \bar{L})} \right) \\ &\quad + \frac{c_{\text{disc}}^2}{M^2} c_1^2 \frac{1}{(r - \alpha)^2} \left(e^{2(r-\alpha)(L_{\max} \wedge \bar{L})} + 1 \right), \end{aligned}$$

where we bounded strictly negative terms by zero from above. As the upper bounds to the squared bias we define the functions

$$B_1(L_{\max} \wedge \bar{L}) := \frac{c_1^2}{\alpha^2} e^{-2\alpha(L_{\max} \wedge \bar{L})},$$

which depends only on $(L_{\max} \wedge \bar{L})$ and

$$B_2(L_{\max} \wedge \bar{L}, M) := \begin{cases} \frac{c_{\text{disc}}^2}{M^2} c_1^2 (L_{\max} \wedge \bar{L})^2 + \frac{c_{\text{disc}}}{M} \frac{2c_1^2}{\alpha} (L_{\max} \wedge \bar{L}) e^{-\alpha(L_{\max} \wedge \bar{L})} & \text{for } r = \alpha, \\ \frac{c_{\text{disc}}^2}{M^2} c_1^2 \frac{1}{(r-\alpha)^2} \left(e^{2(r-\alpha)(L_{\max} \wedge \bar{L})} + 1 \right) \\ \quad + \frac{c_{\text{disc}}}{M} \frac{2c_1^2}{\alpha} \frac{1}{r-\alpha} \left(e^{(r-2\alpha)(L_{\max} \wedge \bar{L})} - e^{-\alpha(L_{\max} \wedge \bar{L})} \right) & \text{for } r \neq \alpha, \end{cases}$$

which depends on $(L_{\max} \wedge \bar{L})$ and M . In order for B_1 to equal $\frac{\varepsilon^2}{2}$, we choose $(L_{\max} \wedge \bar{L})$ such that

$$B_1(L_{\max} \wedge \bar{L}) = \frac{c_1^2}{\alpha^2} e^{-2\alpha(L_{\max} \wedge \bar{L})} = \frac{\varepsilon^2}{2} \quad \text{leading to} \quad L_{\max} \wedge \bar{L} = \frac{1}{\alpha} \ln \left(\frac{c_1 \sqrt{2}}{\alpha \varepsilon} \right). \quad (4.6)$$

To avoid complicated case distinctions we set $L_{\max} = \frac{1}{\alpha} \ln \left(\frac{c_1 \sqrt{2}}{\alpha \varepsilon} \right)$ and choose $\bar{L} = \ln(\tilde{c}^{-\frac{1}{r}} M^{\frac{1}{r}}) \geq \frac{1}{\alpha} \ln \left(\frac{c_1 \sqrt{2}}{\alpha \varepsilon} \right)$ such that $L_{\max} \wedge \bar{L} = L_{\max}$ leading to

$$M \geq \tilde{c} \left(\frac{c_1 \sqrt{2}}{\alpha} \right)^{\frac{r}{\alpha}} \varepsilon^{-\frac{r}{\alpha}} = C_0 \varepsilon^{-\frac{r}{\alpha}}, \quad (4.7)$$

for a constant $C_0 := \tilde{c} \left(\frac{c_1 \sqrt{2}}{\alpha} \right)^{\frac{r}{\alpha}}$ independent of M , $L_{\max}$ and ε . Later, we see that this choice of M subsumes in with the other choices of M in the rest of this proof and is not restrictive with respect to its dependence on ε .

Using the definition for $L_{\max}$, we can estimate terms depending on $L_{\max}$ in terms of ε . For $\varepsilon < e^{-1}$ it holds $|\ln(\varepsilon)| > 1$ and thus we compute

$$\begin{aligned} L_{\max} &= \frac{1}{\alpha} \ln \left(\frac{c_1 \sqrt{2}}{\alpha \varepsilon} \right) = \frac{1}{\alpha} \ln \left(\frac{c_1 \sqrt{2}}{\alpha} \varepsilon^{-1} \right) \\ &= \frac{1}{\alpha} \ln \left(\frac{c_1 \sqrt{2}}{\alpha} \right) |\ln(\varepsilon)| + \frac{1}{\alpha} |\ln(\varepsilon)| \leq \frac{1}{\alpha} \max \left\{ \ln \left(\frac{c_1 \sqrt{2}}{\alpha} \right), 1 \right\} |\ln(\varepsilon)| = C_1 |\ln(\varepsilon)|, \end{aligned} \quad (4.8)$$

with $C_1 := \frac{1}{\alpha} \max \{ \ln \left(\frac{c_1 \sqrt{2}}{\alpha} \right), 1 \} > 0$ independent of M , $L_{\max}$ and ε . Further, we get for a scalar $\rho \in \mathbb{R}$

$$e^{\rho L_{\max}} = e^{\frac{\rho}{\alpha} \ln \left(\frac{c_1 \sqrt{2}}{\alpha \varepsilon} \right)} = \left(\frac{c_1 \sqrt{2}}{\alpha \varepsilon} \right)^{\frac{\rho}{\alpha}} = C_2 \varepsilon^{-\frac{\rho}{\alpha}}, \quad (4.9)$$

with a constant $C_2 := \left(\frac{c_1 \sqrt{2}}{\alpha} \right)^{\frac{\rho}{\alpha}} > 0$ independent of M , $L_{\max}$, and ε . Simply combining both relations we obtain

$$L_{\max} e^{\rho L_{\max}} \leq C_1 C_2 \varepsilon^{-\frac{\rho}{\alpha}} |\ln(\varepsilon)|. \quad (4.10)$$

Furthermore, exponential terms with negative exponent are bounded by one and negative terms are bounded by zero from above.

With these upper bounds at hand we distinct between five different relations between r and α . Recall, that we wish to bound $B_2(L_{\max}, M)$ in each case against $\frac{\varepsilon^2}{4}$ by using $L_{\max}$ from equation (4.6) and choosing M accordingly.

Case 1: ($r < \alpha$)

$$B_2(L_{\max}, M) = \frac{c_{\text{disc}}^2}{M^2} c_1^2 \frac{1}{(r - \alpha)^2} \left(e^{2(r - \alpha)L_{\max}} + 1 \right) + \frac{c_{\text{disc}}}{M} \frac{2c_1^2}{\alpha} \frac{1}{r - \alpha} \left(e^{(r - 2\alpha)L_{\max}} - e^{-\alpha L_{\max}} \right) \leq C_3 \frac{1}{M}.$$

Case 2: ($r = \alpha$)

$$B_2(L_{\max}, M) = \frac{c_{\text{disc}}^2}{M^2} c_1^2 L_{\max}^2 + \frac{c_{\text{disc}}}{M} \frac{2c_1^2}{\alpha} L_{\max} e^{-\alpha L_{\max}} \leq C_4 \left(\frac{\ln(\varepsilon)^2}{M} + 1 \right) \frac{1}{M}.$$

Case 3: ($\alpha < r < 2\alpha$)

$$\begin{aligned} B_2(L_{\max}, M) &= \frac{c_{\text{disc}}^2}{M^2} c_1^2 \frac{1}{(r - \alpha)^2} \left(e^{2(r - \alpha)L_{\max}} + 1 \right) + \frac{c_{\text{disc}}}{M} \frac{2c_1^2}{\alpha} \frac{1}{r - \alpha} \left(e^{(r - 2\alpha)L_{\max}} - e^{-\alpha L_{\max}} \right) \\ &\leq C_5 \left(\frac{1}{M} \varepsilon^{-2} + 1 \right) \frac{1}{M}. \end{aligned}$$

Case 4: ($r = 2\alpha$)

$$B_2(L_{\max}, M) = \frac{c_{\text{disc}}^2}{M^2} c_1^2 \frac{1}{\alpha^2} \left(e^{2\alpha L_{\max}} + 1 \right) \frac{c_{\text{disc}}}{M} \frac{2c_1^2}{\alpha^2} (1 - e^{-\alpha L_{\max}}) \leq C_6 \left(\frac{1}{M} \varepsilon^{-2} + 1 \right) \frac{1}{M}.$$

We note that no logarithmic ε contribution appears in the choice for the sample size M for $r = 2\alpha$, as opposed to the complexity theorem of the standard CLMC estimator.

Case 5: ($r > 2\alpha$)

$$\begin{aligned} B_2(L_{\max}, M) &= \frac{c_{\text{disc}}^2}{M^2} c_1^2 \frac{1}{(r - \alpha)^2} \left(e^{2(r - \alpha)L_{\max}} + 1 \right) + \frac{c_{\text{disc}}}{M} \frac{2c_1^2}{\alpha} \frac{1}{r - \alpha} \left(e^{(r - 2\alpha)L_{\max}} - e^{-\alpha L_{\max}} \right) \\ &\leq C_7 \left(\frac{\varepsilon^{-\frac{r}{\alpha}}}{M} + 1 \right) \frac{\varepsilon^{-\frac{r - 2\alpha}{\alpha}}}{M}. \end{aligned}$$

All constants $C_3, C_4, C_5, C_6, C_7 > 0$ are independent of M , $L_{\max}$ and ε . Overall, we obtain

$$B_2(L_{\max}, M) \leq C_{B_2} \frac{1}{M} \begin{cases} 1 & \text{for } r < \alpha, \\ \frac{\ln(\varepsilon)^2}{M} + 1 & \text{for } r = \alpha, \\ \frac{\varepsilon^{-2}}{M} + 1 & \text{for } \alpha < r \leq 2\alpha, \\ \left(\frac{\varepsilon^{-\frac{r}{\alpha}}}{M} + 1 \right) \varepsilon^{-\frac{r - 2\alpha}{\alpha}} & \text{for } r > 2\alpha \end{cases} \quad (4.11)$$

with a constant $C_{B_2} := \max\{C_3, C_4, C_5, C_6, C_7\} > 0$ of M , $L_{\max}$ and ε . Further to bound B_2 by $\frac{\varepsilon^2}{4}$ by an appropriate choice of M , in all cases $r \geq \alpha$ the additional contributions from ε need to be compensated for by the additional factor of $\frac{1}{M}$. We demonstrate this for the case $r > 2\alpha$ and start by choosing

$$M \geq C_{tmp} \varepsilon^{-2 - \frac{r - 2\alpha}{\alpha}}$$

for a constant C_{tmp} to be chosen subsequently. Then, we compute

$$\begin{aligned} C_7 \left(\frac{\varepsilon^{-\frac{r}{\alpha}}}{M} + 1 \right) \frac{\varepsilon^{-\frac{r-2\alpha}{\alpha}}}{M} &\leq C_7 \left(\frac{\varepsilon^{2+\frac{r-2\alpha}{\alpha}} \varepsilon^{-\frac{r}{\alpha}}}{C_{tmp}} + 1 \right) \frac{\varepsilon^{2+\frac{r-2\alpha}{\alpha}} \varepsilon^{-\frac{r-2\alpha}{\alpha}}}{C_{tmp}} \\ &= C_7 \left(\frac{\varepsilon^{2+\frac{r-2\alpha-r}{\alpha}}}{C_{tmp}} + 1 \right) \frac{\varepsilon^2}{C_{tmp}} = C_7 \left(\frac{1}{C_{tmp}} + 1 \right) \frac{\varepsilon^2}{C_{tmp}}, \end{aligned}$$

and choose $C_{tmp} := 2C_7(1 + \sqrt{1 + \frac{1}{C_7}})$ as the solution to the quadratic equation $C_7(\frac{1}{C_{tmp}} + 1)\frac{\varepsilon^2}{C_{tmp}} = \frac{\varepsilon^2}{4}$. Overall, we obtain $B_2 \leq \frac{\varepsilon^2}{4}$, dealing with all different relations between r and α by choosing M to be

$$M \geq \tilde{C}_{B_2} \varepsilon^{-2-\max\{0, \frac{r-2\alpha}{\alpha}\}}, \quad (4.12)$$

with an appropriate constant $\tilde{C}_{B_2} > 0$ of M , $L_{\max}$ and ε . The initial lower bound for M in equation (4.7) to obtain $L_{\max} \wedge \bar{L} = L_{\max}$ is satisfied by adapting the constant $\tilde{C}_{B_2}$ to be larger than C_0 , since it holds $\varepsilon^{-2-\max\{0, \frac{r-2\alpha}{\alpha}\}} \geq \varepsilon^{-\frac{r}{\alpha}}$ independent of the relation between r and α .

Next, we continue with the bound for the variance term. Note, that the copies of the stochastic process of approximations $(Q(\ell)^{(k)}; \ell \geq 0)$ as well as $(\frac{dQ(\ell)}{d\ell}^{(k)}; \ell > 0)$ are i.i.d. for $1 \leq k \leq M$. Using the Fubini–Tonelli theorem to exchange the covariance with the integration over the level domain we obtain by linearity of integration and bilinearity of the covariance

$$\begin{aligned} \mathbb{V}[\hat{Q}_{0, L_{\max}}^{\text{QCLMC}}] &= \text{Cov}[\hat{Q}_{0, L_{\max}}^{\text{QCLMC}}, \hat{Q}_{0, L_{\max}}^{\text{QCLMC}}] \\ &= \text{Cov} \left[\frac{1}{M} \sum_{k=1}^M \int_0^{L_{\max}} e^{r\ell} \left(\frac{dQ}{d\ell} \right)^{(k)}(\ell) \mathbb{1}_{[0, L_r^{(k)}]}(\ell) d\ell, \right. \\ &\quad \left. \frac{1}{M} \sum_{m=1}^M \int_0^{L_{\max}} e^{r\ell'} \left(\frac{dQ}{d\ell'} \right)^{(m)}(\ell') \mathbb{1}_{[0, L_r^{(m)}]}(\ell') d\ell' \right] \\ &= \frac{1}{M} \int_0^{L_{\max}} \int_0^{L_{\max}} \frac{1}{M} \sum_{k=1}^M \mathbb{1}_{[0, L_r^{(k)}]}(\ell) \mathbb{1}_{[0, L_r^{(k)}]}(\ell') e^{r\ell} e^{r\ell'} \text{Cov} \left[\left(\frac{dQ(\ell)}{d\ell} \right), \left(\frac{dQ(\ell')}{d\ell'} \right) \right] d\ell d\ell' \\ &\leq \frac{1}{M} \int_0^{L_{\max}} \int_0^{L_{\max}} \frac{1}{M} \sum_{k=1}^M \mathbb{1}_{[0, L_r^{(k)}]}(\ell) \mathbb{1}_{[0, L_r^{(k)}]}(\ell') e^{r\ell} e^{r\ell'} \mathbb{V} \left(\frac{dQ(\ell)}{d\ell} \right)^{\frac{1}{2}} \mathbb{V} \left(\frac{dQ(\ell')}{d\ell'} \right)^{\frac{1}{2}} d\ell d\ell' \\ &\leq \frac{c_2}{M} \int_0^{L_{\max}} \int_0^{L_{\max}} \frac{1}{M} \sum_{k=1}^M \mathbb{1}_{[0, L_r^{(k)}]}(\ell) \mathbb{1}_{[0, L_r^{(k)}]}(\ell') e^{(r-\frac{\beta}{2})\ell} e^{(r-\frac{\beta}{2})\ell'} d\ell d\ell', \end{aligned} \quad (4.13)$$

where we used the Cauchy–Schwarz inequality on the covariance and the convergence assumption on the variance of $\frac{dQ(\ell)}{d\ell}$ in (4.4b) from the complexity theorem. For fixed $L_r^{(k)}$ it holds $\mathbb{1}_{[0, L_r^{(k)}]}(\ell) \mathbb{1}_{[0, L_r^{(k)}]}(\ell') = \mathbb{1}_{[0, L_r^{(k)}]}(\max(\ell, \ell'))$ for $\ell, \ell' \in \mathbb{R}_{\geq 0}$ and we insert a zero by subtracting and adding $e^{-r \max\{\ell, \ell'\}}$ to obtain

$$\begin{aligned} &\frac{1}{M} \sum_{k=1}^M \mathbb{1}_{[0, L_r^{(k)}]}(\ell) \mathbb{1}_{[0, L_r^{(k)}]}(\ell') e^{(r-\frac{\beta}{2})\ell} e^{(r-\frac{\beta}{2})\ell'} \\ &= \left(\frac{1}{M} \sum_{k=1}^M \mathbb{1}_{[0, L_r^{(k)}]}(\max(\ell, \ell')) - e^{-r \max\{\ell, \ell'\}} \right) e^{(r-\frac{\beta}{2})\ell} e^{(r-\frac{\beta}{2})\ell'} + e^{-r \max\{\ell, \ell'\}} e^{(r-\frac{\beta}{2})\ell} e^{(r-\frac{\beta}{2})\ell'}. \end{aligned}$$

Inserting this back into equation (4.13), we obtain the following two integral terms,

$$\text{I} := \frac{c_2}{M} \int_0^{L_{\max}} \int_0^{L_{\max}} \left(\frac{1}{M} \sum_{k=1}^M \mathbb{1}_{[0, L_r^{(k)}]}(\max(\ell, \ell')) - e^{-r \max\{\ell, \ell'\}} \right) e^{(r-\frac{\beta}{2})\ell} e^{(r-\frac{\beta}{2})\ell'} d\ell d\ell' \quad (4.14)$$

and

$$\text{II} := \frac{c_2}{M} \int_0^{L_{\max}} \int_0^{L_{\max}} e^{-r \max\{\ell, \ell'\}} e^{(r-\frac{\beta}{2})\ell} e^{(r-\frac{\beta}{2})\ell'} d\ell d\ell', \quad (4.15)$$

that are estimated in Appendix A and lead to the following upper bound of the variance of the QCLMC estimator

$$\begin{aligned} \mathbb{V}[\hat{Q}_{0, L_{\max}}^{\text{QCLMC}}] &\leq \frac{c_{\text{disc}} c_2}{M^2} \begin{cases} \frac{4}{(2r-\beta)^2} (e^{(2r-\beta)L_{\max}} - 2e^{(r-\frac{\beta}{2})L_{\max}} + 1) & \text{for } r \neq \frac{\beta}{2}, \\ (L_{\max})^2 & \text{for } r = \frac{\beta}{2}, \end{cases} \\ &+ \frac{c_2}{M} \begin{cases} \frac{2}{(r-\beta)(r-\frac{\beta}{2})} e^{(r-\beta)L_{\max}} + \frac{4}{\beta(r-\frac{\beta}{2})} e^{-\frac{\beta}{2}L_{\max}} - \frac{4}{(r-\beta)\beta} & \text{for } r \neq \frac{\beta}{2}, \beta, \\ -\frac{8}{\beta^2} e^{-\frac{\beta}{2}L_{\max}} - \frac{4}{\beta} L_{\max} e^{-\frac{\beta}{2}L_{\max}} + \frac{8}{\beta^2} & \text{for } r = \frac{\beta}{2}, \\ \frac{4}{\beta} L_{\max} + \frac{8}{\beta^2} e^{-\frac{\beta}{2}L_{\max}} - \frac{8}{\beta^2} & \text{for } r = \beta, \end{cases} \end{aligned}$$

with explicit dependence on $c_{\text{disc}}, c_2, \beta, r, M, \bar{L}$ and $L_{\max}$. In order to bound the variance in terms of ε and M , we again use the choice of $L_{\max}$ from equation (4.6) and the assumption $\varepsilon < e^{-1}$, that yields $|\ln(\varepsilon)| > 1$. First, we trivially bound the variance further by removing all negative terms and by bounding the exponential terms with negative exponent by one to arrive at

$$\begin{aligned} \mathbb{V}[\hat{Q}_{0, L_{\max}}^{\text{QCLMC}}] &\leq \frac{c_{\text{disc}} c_2}{M^2} \begin{cases} \frac{4}{(2r-\beta)^2} (e^{(2r-\beta)L_{\max}} + 1) & \text{for } r \neq \frac{\beta}{2}, \\ (L_{\max})^2 & \text{for } r = \frac{\beta}{2}, \end{cases} \\ &+ \frac{c_2}{M} \begin{cases} \frac{2}{(r-\beta)(r-\frac{\beta}{2})} e^{(r-\beta)L_{\max}} + \frac{4}{\beta(r-\frac{\beta}{2})} e^{-\frac{\beta}{2}L_{\max}} - \frac{4}{(r-\beta)\beta} & \text{for } r \neq \frac{\beta}{2}, \beta, \\ \frac{8}{\beta^2} & \text{for } r = \frac{\beta}{2}, \\ \frac{4}{\beta} L_{\max} + \frac{8}{\beta^2} & \text{for } r = \beta, \end{cases} \end{aligned}$$

We bound the variance it in each distinct case by using the bounds from equations (4.8) to (4.10).

Case 1: ($r < \frac{\beta}{2}$)

$$\begin{aligned} \mathbb{V}[\hat{Q}_{0, L_{\max}}^{\text{QCLMC}}] &\leq \frac{c_{\text{disc}} c_2}{M^2} \frac{4}{(2r-\beta)^2} (e^{(2r-\beta)L_{\max}} + 1) \\ &+ \frac{c_2}{M} \left(\frac{2}{(r-\beta)(r-\frac{\beta}{2})} e^{(r-\beta)L_{\max}} + \frac{4}{\beta(r-\frac{\beta}{2})} e^{-\frac{\beta}{2}L_{\max}} - \frac{4}{(r-\beta)\beta} \right) \leq C_8 \frac{1}{M}. \end{aligned}$$

Case 2: ($r = \frac{\beta}{2}$)

$$\mathbb{V}[\hat{Q}_{0, L_{\max}}^{\text{QCLMC}}] \leq \frac{c_{\text{disc}} c_2}{M^2} (L_{\max})^2 + \frac{c_2}{M} \frac{8}{\beta^2} \leq C_9 \left(\frac{\ln(\varepsilon)^2}{M} + 1 \right) \frac{1}{M}.$$

Case 3: ($\frac{\beta}{2} < r < \beta$)

$$\begin{aligned} \mathbb{V}[\hat{Q}_{0,L_{\max}}^{\text{QCLMC}}] &\leq \frac{c_{\text{disc}} c_2}{M^2} \frac{4}{(2r-\beta)^2} (e^{(2r-\beta)L_{\max}} + 1) \\ &\quad + \frac{c_2}{M} \left(\frac{2}{(r-\beta)(r-\frac{\beta}{2})} e^{(r-\beta)L_{\max}} + \frac{4}{\beta(r-\frac{\beta}{2})} e^{-\frac{\beta}{2}L_{\max}} - \frac{4}{(r-\beta)\beta} \right) \\ &\leq C_{10} \left(\frac{\varepsilon^{-\frac{\beta}{\alpha}}}{M} + 1 \right) \frac{1}{M}. \end{aligned}$$

Case 4: ($r = \beta$) We compute with $|\ln(\varepsilon)| > 1$

$$\mathbb{V}[\hat{Q}_{0,L_{\max}}^{\text{QCLMC}}] \leq \frac{c_{\text{disc}} c_2}{M^2} \frac{4}{\beta^2} (e^{\beta L_{\max}} + 1) + \frac{c_2}{M} \left(\frac{4}{\beta} L_{\max} + \frac{8}{\beta^2} \right) \leq C_{11} \left(\frac{\varepsilon^{-\frac{\beta}{\alpha}}}{M} + 1 \right) \frac{|\ln(\varepsilon)|}{M}.$$

Case 5: ($r > \beta$)

$$\begin{aligned} \mathbb{V}[\hat{Q}_{0,L_{\max}}^{\text{QCLMC}}] &\leq \frac{c_{\text{disc}} c_2}{M^2} \frac{4}{(2r-\beta)^2} (e^{(2r-\beta)L_{\max}} + 1) \\ &\quad + \frac{c_2}{M} \left(\frac{2}{(r-\beta)(r-\frac{\beta}{2})} e^{(r-\beta)L_{\max}} + \frac{4}{\beta(r-\frac{\beta}{2})} e^{-\frac{\beta}{2}L_{\max}} - \frac{4}{(r-\beta)\beta} \right) \\ &\leq C_{12} \left(\frac{\varepsilon^{-\frac{r}{\alpha}}}{M} + 1 \right) \frac{\varepsilon^{-\frac{r-\beta}{\alpha}}}{M}. \end{aligned}$$

All constants $C_8, C_9, C_{10}, C_{11}, C_{12} > 0$ are independent of M , $L_{\max}$ and ε . Overall, we obtain

$$\mathbb{V}[\hat{Q}_{0,L_{\max}}^{\text{QCLMC}}] \leq C_V \frac{1}{M} \begin{cases} 1 & \text{for } r < \frac{\beta}{2}, \\ \frac{\ln(\varepsilon)^2}{M} + 1 & \text{for } r = \frac{\beta}{2}, \\ \frac{\varepsilon^{-\frac{2r-\beta}{\alpha}}}{M} + 1 & \text{for } \frac{\beta}{2} < r < \beta, \\ \left(\frac{\varepsilon^{-\frac{\beta}{\alpha}}}{M} + 1 \right) |\ln(\varepsilon)| & \text{for } r = \beta, \\ \left(\frac{\varepsilon^{-\frac{r}{\alpha}}}{M} + 1 \right) \varepsilon^{-\frac{r-\beta}{\alpha}} & \text{for } r > \beta, \end{cases} \quad (4.16)$$

with a constant $C_V := \max\{C_8, C_9, C_{10}, C_{11}, C_{12}\} > 0$, independent of M , $L_{\max}$ and ε . Further, to bound the variance by $\frac{\varepsilon^2}{4}$ by an appropriate choice of M , in all cases $r \geq \frac{\beta}{2}$ the additional contributions from ε need to be compensated for by the additional factor of $\frac{1}{M}$. This is realized, as shown in the bias part right before equation (4.12), by choosing

$$M \geq \tilde{C}_V \varepsilon^{-2-\max\{0, \frac{r-\min\{\beta, 2\alpha\}}{\alpha}\}} |\ln(\varepsilon)|^{\delta_{r,\beta}},$$

with an appropriate constant $\tilde{C}_V > 0$ independent of M , $L_{\max}$ and ε . The lower bound on M determined by the term B_2 (see (4.12)) is no more constraining than this newly established lower bound for M , with respect to its dependence on ε , since

$$\varepsilon^{-2-\max\{0, \frac{r-\min\{\beta, 2\alpha\}}{\alpha}\}} \geq \varepsilon^{-2-\max\{0, \frac{r-2\alpha}{\alpha}\}}.$$

For the upper bound to the MSE we therefore obtain

$$\text{MSE}_{0,L_{\max}}^{\text{QCLMC}} \leq \mathbb{V}[\hat{Q}_{0,L_{\max}}^{\text{QCLMC}}] + B_2(L_{\max} \wedge \bar{L}, M) + B_1(L_{\max} \wedge \bar{L}) \leq \frac{\varepsilon^2}{4} + \frac{\varepsilon^2}{4} + \frac{\varepsilon^2}{2} = \varepsilon^2,$$

by choosing

$$M \geq \left\lceil C_{MSE} \varepsilon^{-2-\max\{0, \frac{r-\min\{\beta, 2\alpha\}}{\alpha}\}} |\ln(\varepsilon)|^{\delta_{r,\beta}} \right\rceil, \quad (4.17)$$

where $\lceil \cdot \rceil$ denotes the Gauss bracket, for a constant $C_{MSE} := \max\{\tilde{C}_{B_2}, \tilde{C}_V\} > 0$ independent of M , $L_{\max}$ and ε . Finally, we compute an upper bound for the cost of the estimator by

$$\begin{aligned} \mathcal{C}(\hat{Q}_{0,L_{\max}}^{\text{QCLMC}}) &= \sum_{k=1}^M \int_0^{L_{\max}} \mathbb{1}_{[0,L^k]}(\ell) \frac{d\mathcal{C}[\ell]}{d\ell} d\ell \leq c_3 \sum_{k=1}^M \int_0^{L_{\max}} \mathbb{1}_{[0,L^k]}(\ell) e^{\gamma\ell} d\ell \\ &= c_3 M \int_0^{L_{\max}} \frac{1}{M} \sum_{k=1}^M \mathbb{1}_{[0,L^k]}(\ell) e^{\gamma\ell} d\ell = c_3 M \int_0^{L_{\max}} \left(\frac{1}{M} \sum_{k=1}^M \mathbb{1}_{[0,L^k]}(\ell) - e^{-r\ell} + e^{-r\ell} \right) e^{\gamma\ell} d\ell \\ &\leq c_3 M \sup_{\ell > 0} \left| \frac{1}{M} \sum_{k=1}^M \mathbb{1}_{[0,L^k]}(\ell) - e^{-r\ell} \right| \int_0^{L_{\max}} e^{\gamma\ell} d\ell + c_3 M \int_0^{L_{\max}} e^{-r\ell} e^{\gamma\ell} d\ell \\ &\leq c_3 c_{\text{disc}} \frac{1}{\gamma} (e^{\gamma L_{\max}} - 1) + c_3 M \begin{cases} \frac{1}{\gamma-r} (e^{(\gamma-r)L_{\max}} - 1) & \text{for } r \neq \gamma, \\ L_{\max} & \text{for } r = \gamma, \end{cases} \\ &\leq C_2 \varepsilon^{-\frac{\gamma}{\alpha}} + C_{13} M \varepsilon^{-\max\{0, \frac{\gamma-r}{\alpha}\}} |\ln(\varepsilon)|^{\delta_{r,\gamma}} \end{aligned}$$

for a constant $C_{13} > 0$ independent of M , $L_{\max}$ and ε . Bounding the Gauss bracket in equation (4.17) by adding one to it and inserting this for M we obtain for the cost

$$\begin{aligned} \mathcal{C}(\hat{Q}_{0,L_{\max}}^{\text{QCLMC}}) &\leq C_2 \varepsilon^{-\frac{\gamma}{\alpha}} + C_{13} \varepsilon^{-\max\{0, \frac{\gamma-r}{\alpha}\}} |\ln(\varepsilon)|^{\delta_{r,\gamma}} \\ &\quad + C_{13} \varepsilon^{-\max\{0, \frac{\gamma-r}{\alpha}\}} |\ln(\varepsilon)|^{\delta_{r,\gamma}} C_{MSE} \varepsilon^{-2-\max\{0, \frac{r-\min\{\beta, 2\alpha\}}{\alpha}\}} |\ln(\varepsilon)|^{\delta_{r,\beta}} \\ &\leq C \varepsilon^{-2-\max\{0, \frac{\gamma-\min\{\beta, 2\alpha\}}{\alpha}\}} |\ln(\varepsilon)|^{\delta_{r,\beta} + \delta_{r,\gamma}} \end{aligned}$$

for some constant $C > 0$, independent of M , $L_{\max}$ and ε , finishing the proof. In the last inequality we used the assumption $r \in [\min\{\beta, 2\alpha, \gamma\}, \max\{\min\{\beta, 2\alpha\}, \gamma\}]$ that yields $\varepsilon^{-\max\{0, \frac{\gamma-r}{\alpha}\}} \varepsilon^{-2-\max\{0, \frac{r-\min\{\beta, 2\alpha\}}{\alpha}\}} = \varepsilon^{-2-\max\{0, \frac{\gamma-\min\{\beta, 2\alpha\}}{\alpha}\}}$ and furthermore the trivial bounds for $0 < \varepsilon < e^{-1}$

$$\varepsilon^{-\frac{\gamma}{\alpha}} \leq \varepsilon^{-2-\max\{0, \frac{\gamma-\min\{\beta, 2\alpha\}}{\alpha}\}}, \quad \varepsilon^{-\max\{0, \frac{\gamma-r}{\alpha}\}} \leq \varepsilon^{-\max\{0, \frac{\gamma-\min\{\beta, 2\alpha\}}{\alpha}\}} \quad \text{and} \quad |\ln(\varepsilon)|^{\delta_{r,\gamma}} \leq \varepsilon^{-2}.$$

□

Remark 4.4. For simplicity of notation the proof is given for the choice $L_{\max} < \infty$ such that the bias term is bounded by $\frac{\varepsilon}{\sqrt{2}}$ and an appropriate value of M (see (4.7)) to ensure $L_{\max} \wedge \bar{L} = L_{\max}$ is considered. Note, however, that the complexity theorem still holds in the case $L_{\max} = \infty$, where $L_{\max} \wedge \bar{L} = \bar{L}$, with the same choice of M . Most importantly this means, that for $L_{\max} = \infty$ all error contributions in the MSE, namely the squared bias and the variance, decrease automatically with growing M while the cost does not blow up, because the maximal level $\bar{L} = \ln(\tilde{c}^{-\frac{1}{r}} M^{\frac{1}{r}})$ grows with M at just the right speed. Further note, that taking the limit $M \rightarrow \infty$ in the complexity theorem as done to proof the unbiasedness properties of the QCLMC estimator in Proposition 4.1 is not sensible, since $\lim_{M \rightarrow \infty} \text{MSE}^{\text{QCLMC}} = 0$ with $\lim_{M \rightarrow \infty} \mathcal{C}(\hat{Q}_{0,L_{\max}}^{\text{QCLMC}}) = \infty$.

Remark 4.5. It is important to note, that in order to prove the Complexity Theorem 4.3 using the F -discrepancy property as done in this work, the sequence $(L_r^{(k)}; k = 1, \dots, M)$ must satisfy $\kappa = 0$. This is not the case for an i.i.d. sequence with $\kappa = \frac{1}{2}$. The additional dependence on ε stemming from the use of the F -discrepancy property in the upper bounds to B_2 (see (4.11)) and upper bound to the variance (see (4.16)), cannot be compensated by $M^{\kappa-1}$ with $\kappa = \frac{1}{2}$ instead of $\kappa = 0$.

5. NUMERICAL EXPERIMENTS

The derived upper bounds to the MSE from the proofs of the complexity theorems for QCLMC and CLMC allow us to compare the methods to one another in terms of their computational time to error performance on the basis of the underlying stochastic model parameters. In order to do so we introduce a random PDE model as our stochastic model problem and approximate it by a spatial discretization *via* h -adaptive finite elements. Further, in Algorithm 1 we formulate a practical (Q)CLMC algorithm, state how to obtain the sample adaptive meshes *via a-posteriori* error estimation and we demonstrate how to numerically estimate the model parameters, that are the basis of the performance comparison.

As a followup experiment we compare the performance of CLMC and QCLMC by estimating the real achieved MSE over a series of runs of the (Q)CLMC algorithm for a growing sequence of sample sizes.

The numerical experiments are implemented in Python, where all finite element computations are implemented in FEniCS [5]. The linear systems are solved with its integrated optimized direct LU -decomposition. The computations are done on an Intel(R) Core(TM) i7-4770 CPU running at 3.4 GHz with 4 cores and 2 threads per core.

5.1. Random PDE model and its discretization

For the comparison of performances between the CLMC and QCLMC method we consider the quantity of interest $\mathcal{Q} = \|\cdot\|_{H^1(\mathcal{D})}$ to be the H^1 -norm of the solution of a random PDE, that we introduce next. Let $(\Omega, \mathcal{A}, \mathbb{P})$ be a complete probability space and $\mathcal{D} \subset \mathbb{R}^d$, $d = 1, 2, 3$ be a bounded and connected Lipschitz domain. The linear, random elliptic PDE with solution $u : \Omega \times \mathcal{D} \rightarrow \mathbb{R}$ is given by

$$-\nabla \cdot (a(\omega, x) \nabla u(\omega, x)) = f(x) \quad \text{in } \Omega \times \mathcal{D}, \quad (5.1)$$

where $f : \mathcal{D} \rightarrow \mathbb{R}$ is the source term and $a : \Omega \times \mathcal{D} \rightarrow \mathbb{R}$ is the random coefficient. The boundary $\partial\mathcal{D}$ is assumed to be Lipschitz continuous and equipped with homogeneous Dirichlet boundary conditions

$$u(\omega, x) = 0 \quad \text{on } \Omega \times \partial\mathcal{D}.$$

This is a simple mathematical model for subsurface flow through porous media and has been a common model problem in various works on uncertainty quantification (see, *e.g.*, [6, 9, 35]). For simplicity we choose $\mathcal{D} := [0, 1]^2$ and set $f \equiv 1$ in equation (5.1). We consider a log-Gauss random field as the random coefficient a with a covariance function of the Matérn class, *i.e.*,

$$\text{Cov}(x, y) := v \frac{2^{1-\nu}}{\Gamma(\nu)} \left(\frac{\sqrt{2\nu} \|x - y\|_2}{\lambda} \right)^\nu K_\nu \left(\frac{\sqrt{2\nu} \|x - y\|_2}{\lambda} \right) \quad \text{for } x, y \in \mathbb{R}^d, \quad (5.2)$$

where $\|\cdot\|_2$ denotes the Euclidean norm on $\mathcal{D}$ and where $v > 0$ is the variance, $\lambda > 0$ the correlation length and $\nu > 0$ a parameter steering the roughness of the field. The functions Γ and K_ν are the Gamma function and modified Bessel function, cf. [2], respectively. As an approximation to a we consider a truncated Karhunen–Loève expansion (cf. [4]), for $R \in \mathbb{N}$

$$a_R(x, \omega) = \exp \left(\sum_{m=1}^R \sqrt{\mu_m} \phi_m(x) \xi_m \right), \quad (5.3)$$

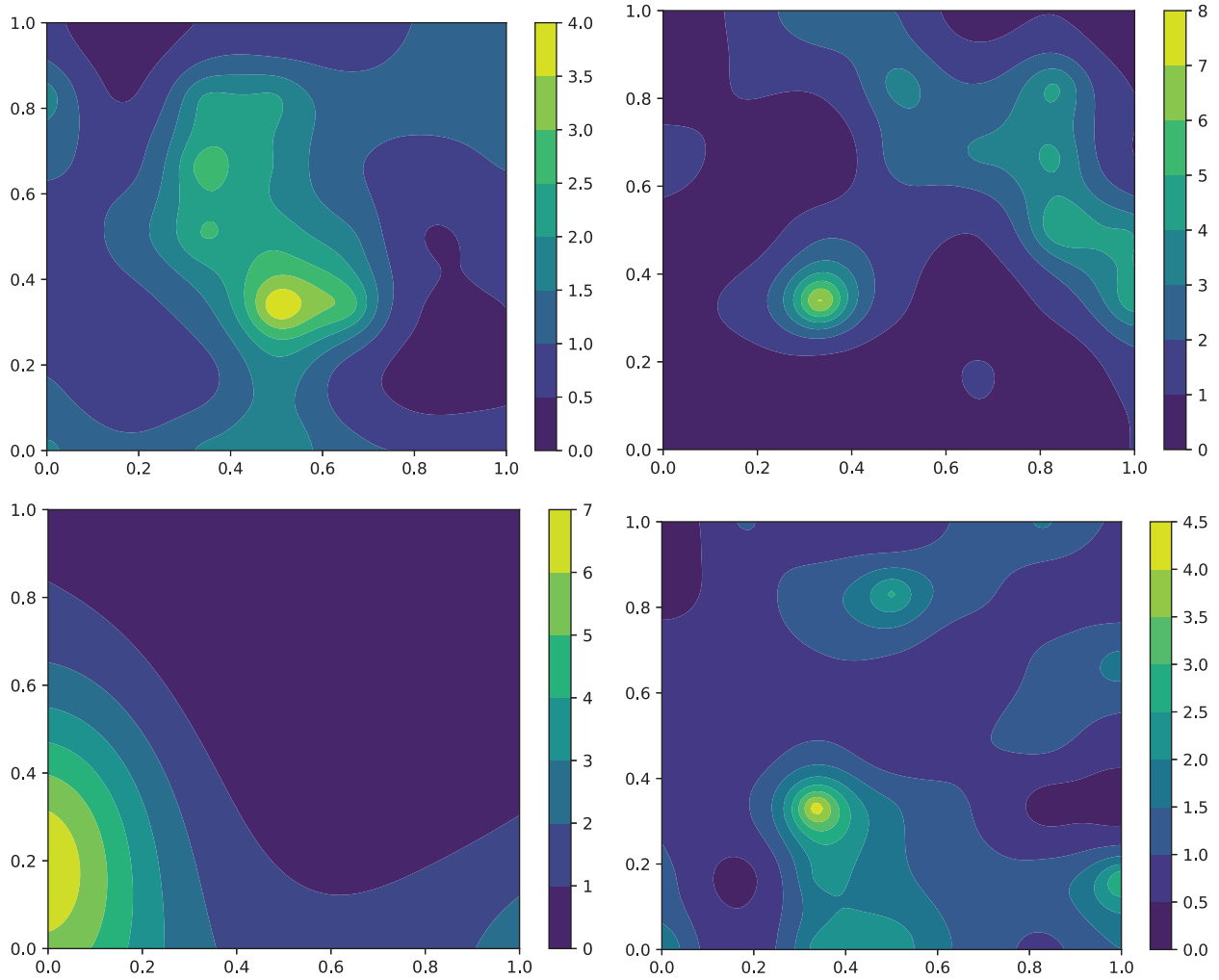


FIGURE 2. Visualization of samples of the log-Gauss random coefficient for $\nu = 1.5$, $\lambda = 0.1$, $v = 0.5$ (*upper left*), $\nu = 1.5$, $\lambda = 0.1$, $v = 1$ (*upper right*), $\nu = 10$, $\lambda = 0.5$, $v = 2$ (*lower left*), $\nu = 1$, $\lambda = 0.1$, $v = 0.5$ (*lower right*). The KL-expansion (5.3) was truncated after $R = 36$ terms in each case.

where $\xi_m \stackrel{d}{=} \mathcal{N}(0, 1)$ are standard normal-distributed random variables and the eigenvalues μ_m and eigenfunctions ϕ_m of the covariance kernel (5.2) are approximated *via* the Nyström method (cf. [32]), for $1 \leq m \leq R$. For the numerical discretization of the PDE (5.1) we consider the Finite Element method (FE), see, *e.g.*, [8, 20, 24], with sample-dependent adaptive meshes. Exemplary visualizations of the log-Gauss random coefficient are given in Figure 2 and a numerical approximation to a pathwise solution of the PDE (5.1) on an adaptive mesh is given in Figure 3.

5.2. Practical estimator, *a-posteriori* error and parameter estimates

The continuous stochastic process $\frac{dQ}{d\ell}^{(k)}$ for each sample $k \in \mathbb{N}$ has to be approximated in order to be computable numerically. As described in [10] and as done in the beginning when deriving the MLMC estimator

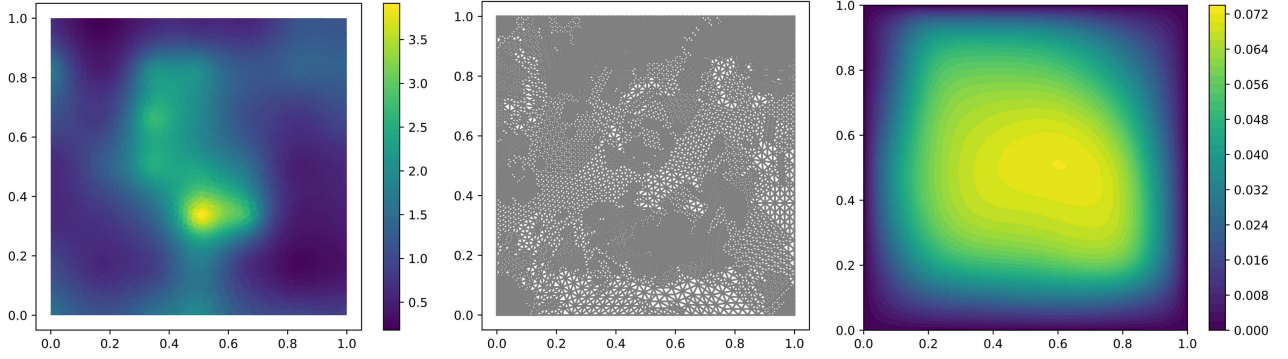


FIGURE 3. Single sample of the log-Gaussian random coefficient for $\nu = 1.5$, $\lambda = 0.1$, $v = 0.5$ (*left*) on adaptive mesh (*middle*) generated by 5 iterative refinement steps, see Remark 5.1, and corresponding PDE solution (*right*). The KL-expansion was truncated after $R = 36$ terms.

from the CLMC estimator, a straightforward approximation is given *via* linear interpolation

$$\frac{dQ^{(k)}}{d\ell}(\ell) := \frac{Q_j^{(k)} - Q_{j-1}^{(k)}}{\ell_j^{(k)} - \ell_{j-1}^{(k)}} \quad \text{for } \ell \in \left(\ell_{j-1}^{(k)}, \ell_j^{(k)}\right],$$

with samples $Q_j^{(k)}$ as approximations to the quantity of interest at levels $\ell_j^{(k)}$ for $j \geq 1$, $k \in \mathbb{N}$. This is only one possible choice of many, *e.g.*, a particular regression function or polynomial interpolant may be used to match the global trend of the process $Q(\ell)$, cf. Section 3.3 of [10] for details. Inserting the linear interpolation into the QCLMC (respectively CLMC) estimator we obtain

$$\hat{Q}_{0, L_{\max}}^{\text{QCLMC}} = \frac{1}{M} \sum_{k=1}^M \sum_{j=1}^{J^{(k)}} \int_{\ell_{j-1}^{(k)}}^{\tilde{\ell}_j^{(k)}} \frac{1}{\mathbb{P}(L_r \geq \ell)} d\ell \frac{Q_j^{(k)} - Q_{j-1}^{(k)}}{\ell_j^{(k)} - \ell_{j-1}^{(k)}}, \quad (5.4)$$

with $M \in \mathbb{N}$ and

$$J^{(k)} := \min \left\{ j \geq 1 : \ell_j^{(k)} \geq L_r^{(k)} \wedge L_{\max} \right\}, \quad \tilde{\ell}_j^{(k)} := \ell_j^{(k)} \wedge L_r^{(k)} \wedge L_{\max}.$$

Different to the derivation of the MLMC estimator in equation (2.6), the level random variable L_r is exponentially-distributed to some parameter $r > 0$ and the integral in the estimator (5.4) computes to

$$\int_{\ell_{j-1}^{(k)}}^{\tilde{\ell}_j^{(k)}} \frac{1}{\mathbb{P}(L_r \geq \ell)} d\ell = \int_{\ell_{j-1}^{(k)}}^{\tilde{\ell}_j^{(k)}} e^{r\ell} d\ell = \frac{\exp(r\tilde{\ell}_j^{(k)}) - \exp(r\ell_{j-1}^{(k)})}{r},$$

and the practical (Q)CLMC estimator is given by

$$\hat{Q}_{0, L_{\max}}^{\text{QCLMC}} = \frac{1}{M} \sum_{k=1}^M \sum_{j=1}^{J^{(k)}} \frac{\exp(r\tilde{\ell}_j^{(k)}) - \exp(r\ell_{j-1}^{(k)})}{r(\ell_j^{(k)} - \ell_{j-1}^{(k)})} (Q_j^{(k)} - Q_{j-1}^{(k)}).$$

The samplewise continuous level of refinement for each sample k is defined *via* a samplewise *a-posteriori* error estimator defined below in equation (5.6) by

$$\ell_j^{(k)} := -\ln(e_j^{(k)}/e_0^{(k)}) \quad \text{for } j = 0, \dots, J, \text{ for } J \in \mathbb{N} \quad \text{and} \quad k = 1, \dots, M, \quad (5.5)$$

naturally providing values $\ell_0^{(k)} = 0$ for all $k = 1, \dots, M$. In order to use sample adaptive meshes in the (Q)CLMC method, we use a standard energy norm error estimator for each sample $k \in \mathbb{N}$ (we refer to [19] for an overview of different *a-posteriori* error estimation techniques). The values $(e_j^{(k)}; j = 0, \dots, J)$ are computable *a-posteriori* error estimators that satisfy

$$\left| \mathcal{Q}^{(k)} - Q_j^{(k)} \right| = \left| \|u^{(k)}\|_{H^1(\mathcal{D})} - \|u_j^{(k)}\|_{H^1(\mathcal{D})} \right| \leq C_{\text{est}}^{(k)} \left(\sum_{K \in \mathcal{K}_j^{(k)}} \left(\varphi_K^{(k)} \right)^2 \right)^{\frac{1}{2}} =: e_j^{(k)}, \quad (5.6)$$

for each sample $k \in \mathbb{N}$ with a constant $C_{\text{est}}^{(k)} > 0$ independent of $u^{(k)}$ and the FE approximation $u_j^{(k)}$. The elementwise error indicator $\varphi_K^{(k)}$ is given by the formula

$$\left(\varphi_K^{(k)} \right)^2 = h_K^2 \|f_j + \nabla \cdot (a_j \nabla u_j)\|_{L^2(K)}^2 + \frac{1}{2} \sum_{\gamma \in \mathcal{E}_K} h_\gamma \left\| [\vec{n}_\gamma|_K \cdot (a_j \nabla u_j)]_\gamma \right\|_{L^2(\gamma)}^2,$$

where we omitted the dependence of the right hand side terms on k for a better readability. The quantities f_j and a_j are approximations to f and a , h_K and h_γ are the element diameter and edge length of element K and edge γ and $\vec{n}_\gamma$ is the outward pointing unit normal vector to edge γ . Further, $\mathcal{E}_K$ is the collection of all edges of element $K \in \mathcal{K}_j^{(k)}$ and $[\cdot]_\gamma$ denotes the jump of a quantity over the edge γ . Details on the derivation of the estimator are found in, e.g., [3, 19].

As in [7], we estimate the underlying model parameters and constants for CLMC and QCLMC, $\alpha, \beta, \gamma, c_1, c_2, c_3$ from Theorem 4.3 numerically, since for the considered model problem and various real-world applications they are not available theoretically. We apply the natural logarithm to equations (4.4a)–(4.4c) to obtain the linear relationships $\ln(\mathbb{E}[\frac{dQ}{d\ell}]) \leq \tilde{c}_1 - \alpha\ell$, $\ln(\mathbb{V}[\frac{dQ}{d\ell}]) \leq \tilde{c}_2 - \beta\ell$ and $\ln(\frac{dC}{d\ell}) \leq \tilde{c}_3 + \gamma\ell$, where $\tilde{c}_i = \ln(c_i)$ for $i = 1, 2, 3$. Using the definition for $\frac{dQ}{d\ell}$ from above and a similar definition for $\frac{dC}{d\ell}$, the mean, variance and cost quantities are estimated by sample averages at refinement steps $j = 0, \dots, J$ with corresponding approximations to the levels $\ell_j \approx \frac{1}{M} \sum_{k=1}^M \ell_j^{(k)}$. Finally, the parameters and constants from the linear relationships are obtained by linear fitting.

Remark 5.1. The adaptive refinement procedure (throughout this work) of $J \in \mathbb{N}$ refinement steps is the classical Dörfler marking strategy from [12]. Starting on an initial unstructured uniform mesh, all elements that exceed 50% of the total *a-posteriori* error bound according to equation (5.6) are refined in each step, i.e., for $j = 0, \dots, J - 1$.

5.3. Comparison of upper bounds to the MSE

We compare the theoretical performance of QCLMC and unbiased CLMC (cf. [7, 10]) based on the derived upper bounds to the MSE from the proofs of their respective complexity theorems. The parameter estimates given in Table 1 provide the convergence regime. All experiments share the properties $\gamma \approx 2$ as opposed to $\gamma \approx 1$, which is usually expected by an optimal direct solver to solve a 2-dimensional PDE problem. But, as indicated in Remark 2.2 the rate γ in equation (4.4c) scales with the average growth of the computed levels $\ell_j^{(k)}$ (see Eq. (5.5)) for each sample $k \in \mathbb{N}$ over the refinements $j \in \mathbb{N}$. This average growth of the level is halved since the *a-posteriori* error estimator from equation (5.6) is actually an upper bound to the samplewise strong error $\|u^{(k)} - u_j^{(k)}\|_{H^1(\mathcal{D})}$ and, hence converges with half the rate as the samplewise weak error $|\|u^{(k)}\|_{H^1(\mathcal{D})} - \|u_j^{(k)}\|_{H^1(\mathcal{D})}|$. Overall, it still holds $\gamma < \min\{\beta, 2\alpha\}$, since α and β are scaled the same way. Further, for all upcoming numerical experiments the truncation index for the KL expansion (5.3) is $R = 36$ and we choose $r = (\gamma + \min\{\beta, 2\alpha\})/2$ to satisfy the assumption on r for QCLMC and CLMC from their complexity theorems. Next, we state the bias and variance bounds from the proof of the complexity theorem for QCLMC

TABLE 1. Estimates for the parameters from equations (4.4a)–(4.4c) for different values of the hyperparameters ν, λ and v of the log-Gauss random coefficient (5.2) for $M = 500$ independent PDE samples, generated by a pseudo-random number generator. The error estimator for the adaptive refinement procedure is defined *via* equation (5.6) and $J = 11$ adaptive refinement steps are used, see Remark 5.1.

Matérn parameters	c_1	α	c_1^2	c_2	β	γ	r
$\nu = 1, \lambda = 0.1, v = 0.5$	5.21e-02	1.85	2.72e-03	4.13e-04	3.69	1.83	2.76
$\nu = 10, \lambda = 0.5, v = 2$	1.47e-01	1.87	2.18e-02	6.32e-02	3.55	1.84	2.7

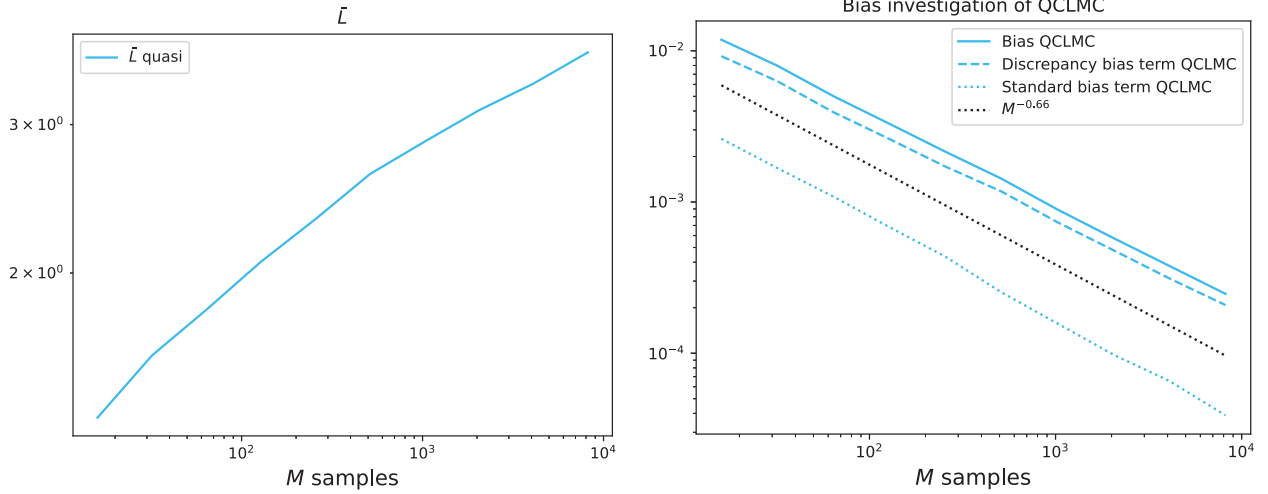


FIGURE 4. Mean of maximal levels $\bar{L}$ generated by quasi-random Sobol sequence for M samples (*left*) and corresponding mean of the upper bound to the bias of QCLMC (*right*) over 100 independent runs realized *via* Owen Scrambling. Hyperparameters for log-Gauss field (5.2): $\nu = 1, \lambda = 0.1, v = 0.5$.

in dependence of the problem parameters $\alpha, \beta, c_1, c_2, r$. The bias is given by

$$\left| \mathbb{E} \left[\hat{Q}_{0, L_{\max}}^{\text{QCLMC}} - (Q - Q(0)) \right] \right| \leq \frac{c_{\text{disc}} c_1}{M} \frac{1}{r - \alpha} \left(e^{(r - \alpha)(L_{\max} \wedge \bar{L})} - 1 \right) + \frac{c_1}{\alpha} e^{-\alpha(L_{\max} \wedge \bar{L})} \quad (5.7)$$

= discrepancy bias term + standard bias term,

where the first term stands for the additional bias introduced by the F -discrepancy of the quasi-random sequence (see Lem. 3.1). The variance is bounded by

$$\begin{aligned} \mathbb{V} \left[\hat{Q}_{0, L_{\max}}^{\text{QCLMC}} \right] &\leq \frac{c_{\text{disc}} c_2}{M^2} \left(\frac{4}{(2r - \beta)^2} \left(e^{(2r - \beta)(L_{\max} \wedge \bar{L})} - 2e^{(r - \frac{\beta}{2})(L_{\max} \wedge \bar{L})} + 1 \right) \right) \\ &\quad + \frac{c_2}{M} \left(\frac{2}{(r - \beta)(r - \frac{\beta}{2})} e^{(r - \beta)(L_{\max} \wedge \bar{L})} + \frac{4}{\beta(r - \frac{\beta}{2})} e^{-\frac{\beta}{2}(L_{\max} \wedge \bar{L})} + \frac{4}{(\beta - r)\beta} \right) \quad (5.8) \\ &= \text{discrepancy variance term} + \text{variance convergence term}, \end{aligned}$$

where we see the split in an additional term introduced by the F -discrepancy of the quasi-random sequence and the term stemming from the assumption on the convergence of the variance decay in the complexity theorem (see (4.4b)). The MSE, see equation (2.1), of QCLMC consists of the variance and the squared bias of the estimator

$$\text{MSE}^{\text{QCLMC}} = \mathbb{V}[\hat{Q}_{0,L_{\max}}^{\text{QCLMC}}] + \left| \mathbb{E}[\hat{Q}_{0,L_{\max}}^{\text{QCLMC}} - (\mathcal{Q} - Q(0))] \right|^2, \quad (5.9)$$

and it is bounded by the respective variance upper bound (see (5.8)) and bias upper bound (see (5.7)). For CLMC the bias is bounded by, cf. [7, 10],

$$\begin{aligned} \left| \mathbb{E}[\hat{Q}_{0,L_{\max}}^{\text{CLMC}} - (\mathcal{Q} - Q(0))] \right| &= |\mathbb{E}[\mathcal{Q} - Q(L_{\max})]| \leq \frac{c_1}{\alpha} e^{-\alpha L_{\max}} \\ &= \text{standard bias term}, \end{aligned} \quad (5.10)$$

which vanishes in the case $L_{\max} = \infty$. For the variance we have

$$\begin{aligned} \mathbb{V}[\hat{Q}_{0,L_{\max}}^{\text{CLMC}}] &\leq \frac{c_2}{M} \left[\frac{1}{\left(r - \frac{\beta}{2}\right)^2} \left(\frac{2r - \beta}{r - \beta} e^{(r-\beta)L_{\max}} + \frac{4r - 2\beta}{\beta} e^{-\frac{\beta}{2}L_{\max}} \right) + \frac{4}{(\beta - r)\beta} \right] \\ &\quad + \frac{c_1^2}{M} \left[\frac{1}{(r - \alpha)^2} \left(\frac{2r - 2\alpha}{r - 2\alpha} e^{(r-2\alpha)L_{\max}} - \frac{(r - \alpha)^2}{\alpha^2} e^{-2\alpha L_{\max}} + \frac{2r^2 - 2r\alpha}{\alpha^2} e^{-\alpha L_{\max}} \right) + \frac{r}{(2\alpha - r)\alpha^2} \right], \end{aligned} \quad (5.11)$$

which in the case $L_{\max} = \infty$ and $r < \min\{\beta, 2\alpha\}$ boils down to

$$\begin{aligned} \mathbb{V}[\hat{Q}_{0,\infty}^{\text{CLMC}}] &\leq \frac{1}{M} \frac{4c_2}{(\beta - r)\beta} + \frac{1}{M} \frac{c_1^2 r}{(2\alpha - r)\alpha^2} \\ &= \text{variance convergence term} + \text{bias convergence term}. \end{aligned} \quad (5.12)$$

Here, the first term corresponds to the assumption on the variance decay and the second to the assumption on the bias decay, see (2.3). The MSE of CLMC is given by

$$\text{MSE}^{\text{CLMC}} = \begin{cases} \mathbb{V}[\hat{Q}_{0,L_{\max}}^{\text{CLMC}}] + |\mathbb{E}[\mathcal{Q} - Q(L_{\max})]|^2 & \text{for } L_{\max} \neq \infty, \\ \mathbb{V}[\hat{Q}_{0,\infty}^{\text{CLMC}}] & \text{for } L_{\max} = \infty, \end{cases} \quad (5.13)$$

which is bounded in the respective case by the variance upper bound (see (5.11) and (5.12)) and bias upper bound (see (5.10)). For our first performance comparison of both methods, we compare the upper bounds to the MSE based on the parameter estimates for the different hyperparameter settings for the log-Gauss PDE coefficient as listed in Table 1 for a range of sample sizes $M = 16 \cdot 2^i$ for $i = 0, 1, \dots, 9$. For CLMC we set $L_{\max} = \infty$ resulting in the unbiased version. For QCLMC we set $L_{\max} = \infty$ as well and thus the upper bounds for the bias and variance are independent of $L_{\max}$, but depend on $\bar{L}$. For QCLMC we compute for each hyperparameter setting in the PDE coefficient an average of the upper bounds over 100 independent runs of a quasi-random sequence $L_r^{(k)}$ yielding different values for $\bar{L}$ in each run. The independence of the quasi-random sequence in QCLMC was realized by Owen scrambling (see [29, 30]) of a Sobol sequence. For QCLMC the maximal generated level $\bar{L}$ in each run, see Figure 4 (left), dictates the bias, illustrated in Figure 4 (right). We see that the additional bias introduced by the quasi-random sequence is significant, but decays faster than $M^{-\frac{1}{2}}$ resulting in a diminishing contribution to the upper bound to the MSE by the squared bias, emphasizing the automatic compensation of the bias error by $\bar{L}$ as described in Remark 4.4. The upper bounds to the variances of both methods are given in Figure 5 (left), where we observe that QCLMC achieves a much lower upper bound to the variance than CLMC. The upper bound to the variance of CLMC is dominated by the bias convergence term. The discrepancy variance

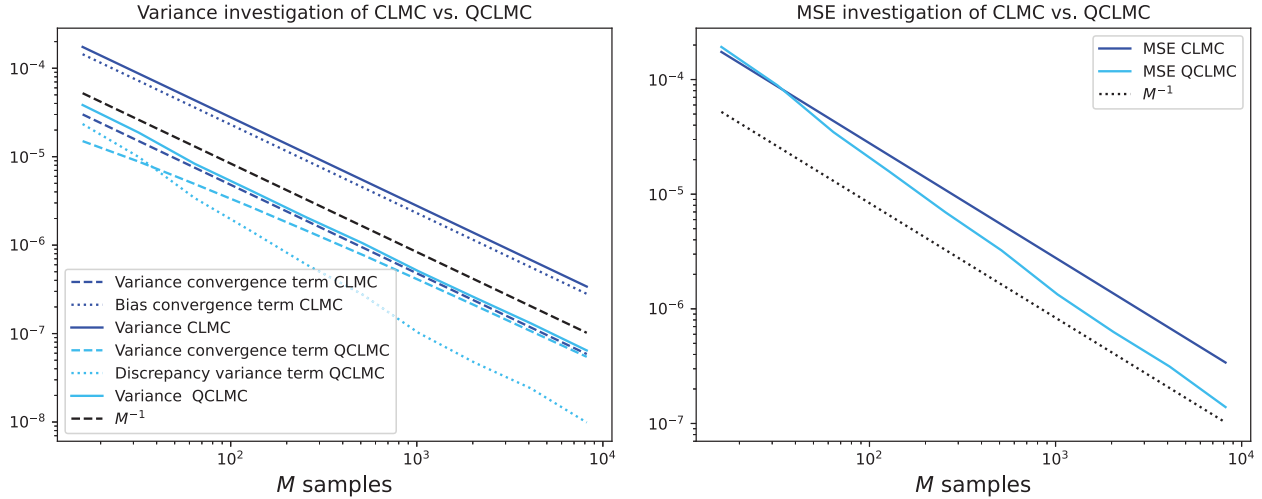


FIGURE 5. Upper bounds to the variance (*left*) and MSE (*right*) for CLMC and QCLMC estimated over 100 independent runs. Hyperparameters for log-Gauss field (5.2): $\nu = 1$, $\lambda = 0.1$, $v = 0.5$.

term decays at a faster rate and is therefore not dominant in the QCLMC estimate. This means that the upper bound to the variance of the QCLMC estimator is essentially only the variance convergence term, whereas the CLMC estimator is dominated by the bias convergence term. The resulting upper bound to the MSE for both methods is shown in Figure 5 (right). We observe a smaller upper bound to the MSE for QCLMC in comparison to CLMC, as a direct consequence of the variance reduction and the natural bias compensation of QCLMC. In Figure 6 (right) we again observe for QCLMC that the additional bias introduced by the quasi-random sequence decays faster than $M^{-\frac{1}{2}}$ resulting in a diminishing contribution to the upper bound to the MSE by the squared bias. In Figure 6 (middle) we observe the same magnitude in the upper bounds to the variance by both CLMC and QCLMC. Here, the bias convergence term in CLMC is insignificant compared to its variance convergence term. For both CLMC and QCLMC the variance convergence term contributes the most to the variance upper bound. Overall, this results in the same upper bound to the MSE for both methods in Figure 6 (right).

The constant in the upper bound to the variance of QCLMC is influenced by the constant c_2 from the variance decay assumption (4.4b), where the upper bound to the variance of CLMC is influenced by both c_2 and c_1^2 from the bias decay assumption (4.4a) and both converge in M with rate one. The constant c_1 enters in QCLMC only in the upper bound to the bias, and the squared bias converges like $M^{-0.66 \cdot 2} < M^{-1}$ for the given examples. This means that for a larger quotient of the constants c_1^2/c_2 given in Table 1 we see a better result for QCLMC compared to CLMC. This behaviour may be explained, since $c_2 > c_1^2$ resembles a high variance of the problem relative to the squared bias, leading to large sample sizes necessary to reduce the variance of the estimator and the effect of accurately sampling the level distribution L_r by fewer samples becomes less significant.

5.4. Performance comparison of CLMC and QCLMC

With the practical estimator defined in Section 5.2 at hand, an algorithm for CLMC and QCLMC to compare their real performance is formulated in Algorithm 1. The algorithm is defined for $L_{\max} = \infty$ and $\gamma < \min\{\beta, 2\alpha\}$. This algorithm is used to evaluate the performance of CLMC and QCLMC for a sequence of sample sizes $M_i = 16 \cdot 2^i$ for $i = 0, 1, \dots, 9$. We compute $K = 100$ independent runs for each of the sample sizes and each

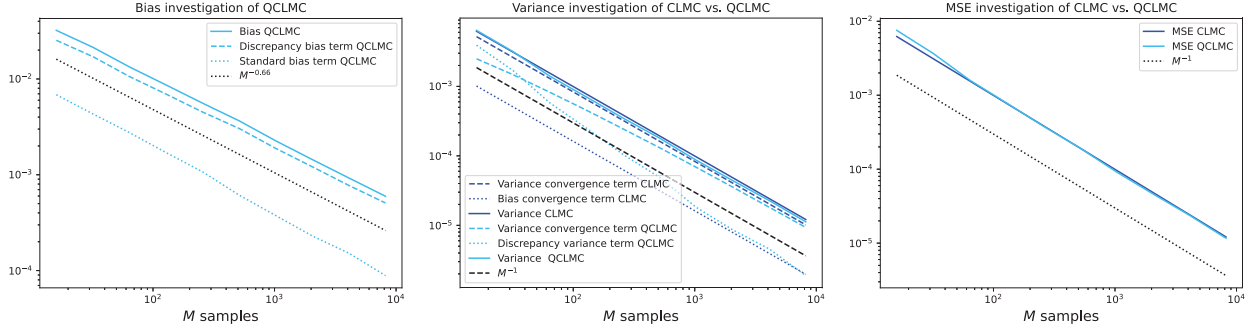


FIGURE 6. Upper bounds to the bias (*left*), variance (*middle*) and MSE (*right*) for CLMC and QCLMC estimated over 100 independent runs. Hyperparameters for log-Gauss field (5.2): $\nu = 10$, $\lambda = 0.5$, $v = 2$.

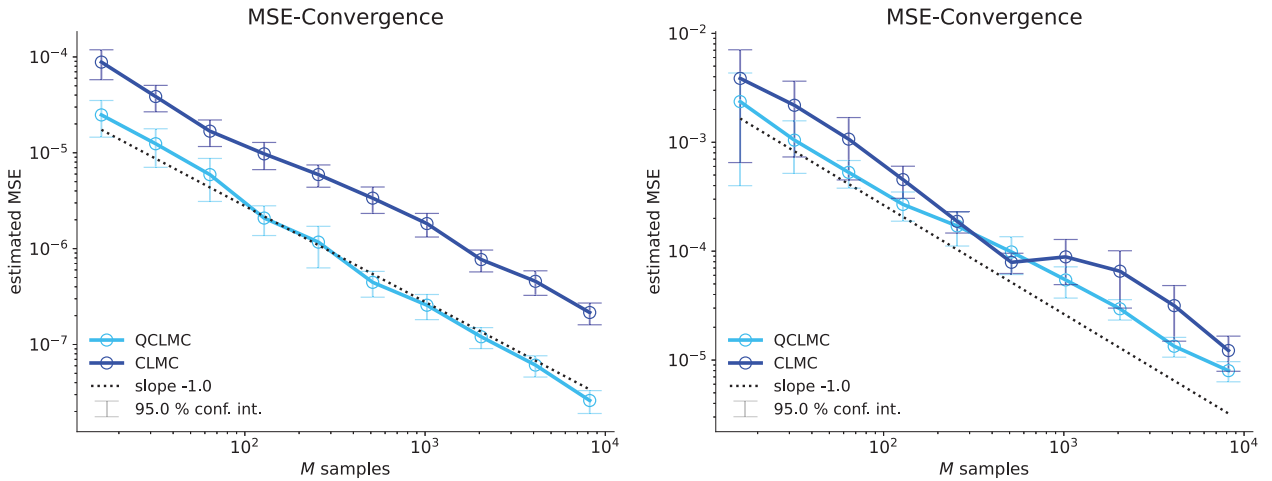


FIGURE 7. Estimated MSE (y -axis) for CLMC and QCLMC over 100 independent runs for sample sizes $M_i = 16 \cdot 2^i$ for $i = 0, \dots, 9$ (x -axis). Hyperparameters for log-Gauss field (5.2): $\nu = 1$, $\lambda = 0.1$, $v = 0.5$ (*left*) and $\nu = 10$, $\lambda = 0.5$, $v = 2$ (*right*).

method and estimate the respective achieved MSE by

$$\text{MSE}^{(\text{Q})\text{CLMC}} = \mathbb{E} \left[\left(\hat{Q}_{0,\infty}^{(\text{Q})\text{CLMC}} - \mathbb{E}[Q - Q(0)] \right)^2 \right] \approx \frac{1}{K} \sum_{k=1}^K \left(\left(\hat{Q}_{0,\infty}^{(\text{Q})\text{CLMC}} \right)^{(k)} - \hat{Q}_{\text{ref}} \right)^2,$$

where the reference solution $\hat{Q}_{\text{ref}} \approx \mathbb{E}[Q - Q(0)]$ is computed by an optimized MLMC algorithm to a very small tolerance, cf. [7, 15]. The convergence results are given in Figure 7, where the 95% confidence intervals are computed *via* the central limit theorem. We observe that both methods achieve their expected cost (measured in sample sizes) to MSE convergence rate of -1 . As already indicated by the previous experiment in Figure 5 (right), we observe in Figure 7 (left) a significant improvement of the MSE for the QCLMC method in comparison to the CLMC method for the same number of samples M . Comparing the MSE curves for CLMC and QCLMC in Figure 7 (left) to the MSE upper bounds given in Figure 5 (right) we observe that the improvement of the upper bounds occurs only for larger values of M , whereas the real estimated MSE for QCLMC is significantly

reduced compared to CLMC right from the start. This goes along with the observation that the upper bound to the MSE for CLMC is tighter than the one for QCLMC, which gets tight for larger values of M . On the other hand, we observe a similar performance of CLMC and QCLMC in Figure 7 (right), which is expected when comparing it to Figure 6 (right). In Tables 2 and 3 the corresponding achieved MSE values are given for each method and each sample size, together with the quotient of improvement by the QCLMC method. The average quotient of improvement for the values in Table 2 is about 5.6 and the quotient between the constants is $c_1^2/c_2 \approx 6.6 > 1$. We conducted several more experiment and observed in the case $c_1^2/c_2 > 1$, that the factor of improvement of QCLMC over CLMC may be related to the ratio c_1^2/c_2 . The average quotient of improvement for the values in Table 3 is about 1.7 and the quotient between the constants is $c_1^2/c_2 \approx 0.34 < 1$. In case $c_1^2/c_2 \leq 1$ both CLMC and QCLMC are expected to perform similarly and the exact ratio c_1^2/c_2 does not seem to be of importance.

Algorithm 1. (Q)CLMC.

Require: $r \in (\gamma, \min\{\beta, 2\alpha\})$ the exponential distribution parameter, $M \in \mathbb{N}$ the total number of samples, the estimator $est \in \{\text{CLMC}, \text{QCLMC}\}$

for $k = 1 : M$ **do**

if $est = \text{CLMC}$ **then**

 Draw sample $L_r^{(k)} \sim \text{Exp}(r)$

else if $est = \text{QCLMC}$ **then**

 Generate quasi-random number $x^{(k)} \in [0, 1)$

 Obtain sample $L_r^{(k)}$ from $x^{(k)}$ via inverse transformation with the exponential CDF $L_r^{(k)} = -\frac{\ln(1-x^{(k)})}{r}$

end if

$j \leftarrow -1$

$\ell_{tmp} \leftarrow 0$

while $\ell_{tmp} \leq L_r^{(k)}$ **do**

$j \leftarrow j + 1$

 Evaluate and save sample $Q_j^{(k)}$

 Evaluate *a-posteriori* error estimate $e_j^{(k)}$ of $|\mathcal{Q}^{(k)} - Q_j^{(k)}|$

 Compute and save level $\ell_j^{(k)} = -\ln(e_j^{(k)}/e_0^{(k)})$

$\ell_{tmp} \leftarrow \ell_j^{(k)}$

end while

 Save $J^{(k)} \leftarrow j$

 Save $\tilde{\ell}_j^{(k)} = \min\{L_r^{(k)}, \ell_j^{(k)}\}$

end for

$$\hat{Q}_{0,\infty}^{\text{QCLMC}} = \frac{1}{M} \sum_{k=1}^M \sum_{j=1}^{J^{(k)}} \frac{\exp(r\tilde{\ell}_j^{(k)}) - \exp(r\ell_{j-1}^{(k)})}{r(\ell_j^{(k)} - \ell_{j-1}^{(k)})} (Q_j^{(k)} - Q_{j-1}^{(k)})$$

6. CONCLUSION

We conclude that QCLMC significantly outperforms CLMC for the provided numerical examples in case $c_1^2/c_2 > 1$ and the factor by which the MSE is improved may be related to the ratio c_1^2/c_2 . Overall, by looking at the upper bounds to the bias and variance, see equations (5.7) and (5.8) for QCLMC and equation (5.12) for CLMC, and the conducted numerical experiments in this work, it is reasonable to expect a similar performance of both methods in case $c_1^2/c_2 \leq 1$, because the error contributions by terms including c_2 converge at a rate M^{-1} for both methods. More importantly, we expect that QCLMC outperforms CLMC in cases where $c_1^2/c_2 > 1$, because the error contributions by terms including c_1^2 converge faster than M^{-1} for QCLMC as a result of the superior convergence properties of the quasi-random sequence in the empirical tail estimate (F -discrepancy).

TABLE 2. Estimated MSE values for CLMC and QCLMC for different values of M together with their quotient, to be able to compare the performances. Hyperparameters for log-Gauss field (5.2): $\nu = 1$, $\lambda = 0.1$, $v = 0.5$.

Method/ M	16	32	64	128	256
MSE^{CLMC}	8.8e−05	3.7e−05	1.7e−05	9.7e−06	5.9e−06
$\text{MSE}^{\text{QCLMC}}$	2.5e−05	1.2e−05	5.9e−06	2.1e−06	1.2e−06
$\frac{\text{MSE}^{\text{CLMC}}}{\text{MSE}^{\text{QCLMC}}}$	3.6	3.1	2.8	4.6	5.1
Method/ M	512	1024	2048	4096	8192
MSE^{CLMC}	3.4e−06	1.8e−06	7.7e−07	4.5e−07	2.1e−07
$\text{MSE}^{\text{QCLMC}}$	4.5e−07	2.5e−07	1.2e−07	6.1e−08	2.6e−08
$\frac{\text{MSE}^{\text{CLMC}}}{\text{MSE}^{\text{QCLMC}}}$	7.5	7.1	6.5	7.5	8.2

TABLE 3. Estimated MSE values for CLMC and QCLMC for different values of M together with their quotient, to be able to compare the performances. Hyperparameters for log-Gauss field (5.2): $\nu = 10$, $\lambda = 0.5$, $v = 2$.

Method/ M	16	32	64	128	256
MSE^{CLMC}	3.9e−03	2.2e−03	1.1e−03	4.5e−04	1.9e−04
$\text{MSE}^{\text{QCLMC}}$	2.4e−03	1.1e−03	5.3e−04	2.7e−04	1.7e−04
$\frac{\text{MSE}^{\text{CLMC}}}{\text{MSE}^{\text{QCLMC}}}$	1.6	2.1	2.0	1.7	1.1
Method/ M	512	1024	2048	4096	8192
MSE^{CLMC}	7.9e−05	8.9e−05	6.5e−05	3.2e−05	1.2e−05
$\text{MSE}^{\text{QCLMC}}$	9.8e−05	5.4e−05	2.9e−05	1.3e−05	7.9e−06
$\frac{\text{MSE}^{\text{CLMC}}}{\text{MSE}^{\text{QCLMC}}}$	0.8	1.6	2.2	2.4	1.5

Due to the above mentioned advantages and essentially the same involved effort in the implementation, we generally recommend to use QCLMC over CLMC.

APPENDIX A. INTEGRAL COMPUTATIONS

Here, we show the bounds for the integral terms I (4.14) and II (4.15) from the proof of the QCLMC complexity theorem. For I we compute

$$\begin{aligned}
I &= \frac{c_2}{M} \int_0^{L_{\max}} \int_0^{L_{\max}} \left(\frac{1}{M} \sum_{k=1}^M \mathbb{1}_{[0, L_r^{(k)}]}(\max(\ell, \ell')) - e^{-r \max\{\ell, \ell'\}} \right) e^{(r-\frac{\beta}{2})\ell} e^{(r-\frac{\beta}{2})\ell'} d\ell d\ell' \\
&= \frac{c_2}{M} \int_0^{L_{\max}} \left(\frac{1}{M} \sum_{k=1}^M \mathbb{1}_{[0, L_r^{(k)}]}(\ell') - e^{-r\ell'} \right) \int_0^{\ell'} e^{(r-\frac{\beta}{2})\ell} e^{(r-\frac{\beta}{2})\ell'} d\ell d\ell' \\
&\quad + \frac{c_2}{M} \int_0^{L_{\max}} \int_{\ell'}^{L_{\max}} \left(\frac{1}{M} \sum_{k=1}^M \mathbb{1}_{[0, L_r^{(k)}]}(\ell) - e^{-r\ell} \right) e^{(r-\frac{\beta}{2})\ell} e^{(r-\frac{\beta}{2})\ell'} d\ell d\ell' \\
&\leq \frac{c_2}{M} \sup_{\ell' > 0} \left| \frac{1}{M} \sum_{k=1}^M \mathbb{1}_{[0, L_r^{(k)}]}(\ell') - e^{-r\ell'} \right| \int_0^{L_{\max}} \int_0^{\ell'} e^{(r-\frac{\beta}{2})\ell} e^{(r-\frac{\beta}{2})\ell'} d\ell d\ell' \\
&\quad + \frac{c_2}{M} \sup_{\ell > 0} \left| \frac{1}{M} \sum_{k=1}^M \mathbb{1}_{[0, L_r^{(k)}]}(\ell) - e^{-r\ell} \right| \int_0^{L_{\max}} \int_{\ell'}^{L_{\max}} e^{(r-\frac{\beta}{2})\ell} e^{(r-\frac{\beta}{2})\ell'} d\ell d\ell' \\
&\leq \frac{c_{\text{disc}} c_2}{M^2} \int_0^{L_{\max}} \int_0^{L_{\max}} e^{(r-\frac{\beta}{2})\ell} e^{(r-\frac{\beta}{2})\ell'} d\ell d\ell' = \frac{c_{\text{disc}} c_2}{M^2} \left(\int_0^{L_{\max}} e^{(r-\frac{\beta}{2})\ell} d\ell \right)^2 \\
&= \frac{c_{\text{disc}} c_2}{M^2} \begin{cases} \left(\frac{1}{r-\frac{\beta}{2}} e^{(r-\frac{\beta}{2})L_{\max}} - \frac{1}{r-\frac{\beta}{2}} \right)^2 & \text{for } \beta \neq \frac{r}{2} \\ (L_{\max})^2 & \text{for } \beta = \frac{r}{2}, \end{cases}
\end{aligned}$$

using the Fubini–Tonelli theorem to compute the double integral as the square of the respective single integral. Further, we compute the square in the case $r \neq \frac{\beta}{2}$ to obtain

$$\begin{aligned}
\left(\frac{1}{r-\frac{\beta}{2}} e^{(r-\frac{\beta}{2})L_{\max}} - \frac{1}{r-\frac{\beta}{2}} \right)^2 &= \frac{4}{(2r-\beta)^2} e^{(2r-\beta)L_{\max}} - \frac{8}{(2r-\beta)^2} e^{(r-\frac{\beta}{2})L_{\max}} + \frac{4}{(2r-\beta)^2} \\
&= \frac{4}{(2r-\beta)^2} \left(e^{(2r-\beta)L_{\max}} - 2e^{(r-\frac{\beta}{2})L_{\max}} + 1 \right).
\end{aligned}$$

For II , we compute the following double integral using the Fubini–Tonelli theorem:

$$II = \frac{c_2}{M} \int_0^{L_{\max}} \int_0^{L_{\max}} e^{-r \max\{\ell, \ell'\}} e^{(r-\frac{\beta}{2})\ell} e^{(r-\frac{\beta}{2})\ell'} d\ell d\ell' = c_2 \int_0^{L_{\max}} e^{(r-\frac{\beta}{2})\ell} \int_0^{L_{\max}} e^{-r \max(\ell, \ell')} e^{(r-\frac{\beta}{2})\ell'} d\ell' d\ell.$$

The inner integral computes in the case $r \neq \frac{\beta}{2}$ as

$$\begin{aligned}
\int_0^{L_{\max}} e^{-r \max(\ell, \ell')} e^{(r-\frac{\beta}{2})\ell'} d\ell' &= e^{-r\ell} \int_0^{\ell} e^{(r-\frac{\beta}{2})\ell'} d\ell' + \int_{\ell}^{L_{\max}} e^{-\frac{\beta}{2}\ell'} d\ell' \\
&= \frac{1}{r-\frac{\beta}{2}} e^{-r\ell} \left(e^{(r-\frac{\beta}{2})\ell} - 1 \right) + \frac{1}{-\frac{\beta}{2}} \left(e^{-\frac{\beta}{2}L_{\max}} - e^{-\frac{\beta}{2}\ell} \right) = \frac{1}{r-\frac{\beta}{2}} e^{-\frac{\beta}{2}\ell} - \frac{1}{r-\frac{\beta}{2}} e^{-r\ell} + \frac{2}{\beta} e^{-\frac{\beta}{2}\ell} - \frac{2}{\beta} e^{-\frac{\beta}{2}L_{\max}} \\
&= \frac{2r}{\left(r-\frac{\beta}{2}\right)\beta} e^{-\frac{\beta}{2}\ell} - \frac{1}{r-\frac{\beta}{2}} e^{-r\ell} - \frac{2}{\beta} e^{-\frac{\beta}{2}L_{\max}}.
\end{aligned}$$

Inserting this on top again leaves us to compute 3 more integrals and a further case distinction, where we start with $r \neq \beta$

$$\begin{aligned}
& \frac{2r}{\left(r - \frac{\beta}{2}\right)\beta} \int_0^{L_{\max}} e^{(r-\beta)\ell} d\ell - \frac{1}{r - \frac{\beta}{2}} \int_0^{L_{\max}} e^{-\frac{\beta}{2}\ell} d\ell - \frac{2}{\beta} e^{-\frac{\beta}{2}L_{\max}} \int_0^{L_{\max}} e^{(r-\frac{\beta}{2})\ell} d\ell \\
&= \frac{2r}{\left(r - \frac{\beta}{2}\right)\beta(r - \beta)} e^{(r-\beta)L_{\max}} - \frac{2r}{\left(r - \frac{\beta}{2}\right)\beta(r - \beta)} \\
&\quad + \frac{2}{\left(r - \frac{\beta}{2}\right)\beta} e^{-\frac{\beta}{2}L_{\max}} - \frac{2}{\left(r - \frac{\beta}{2}\right)\beta} - \frac{2}{\beta\left(r - \frac{\beta}{2}\right)} e^{(r-\beta)L_{\max}} + \frac{2}{\beta\left(r - \frac{\beta}{2}\right)} e^{-\frac{\beta}{2}L_{\max}} \\
&= \frac{2}{(r - \beta)\left(r - \frac{\beta}{2}\right)} e^{(r-\beta)L_{\max}} + \frac{4}{\beta\left(r - \frac{\beta}{2}\right)} e^{-\frac{\beta}{2}L_{\max}} - \frac{4}{(r - \beta)\beta},
\end{aligned}$$

and for the case $r = \beta$ we compute

$$\begin{aligned}
& \frac{2r}{\left(r - \frac{\beta}{2}\right)\beta} \int_0^{L_{\max}} e^{(r-\beta)\ell} d\ell - \frac{1}{r - \frac{\beta}{2}} \int_0^{L_{\max}} e^{-\frac{\beta}{2}\ell} d\ell - \frac{2}{\beta} e^{-\frac{\beta}{2}L_{\max}} \int_0^{L_{\max}} e^{(r-\frac{\beta}{2})\ell} d\ell \\
&= \frac{4}{\beta} L_{\max} + \frac{4}{\beta^2} e^{-\frac{\beta}{2}L_{\max}} - \frac{4}{\beta^2} - \frac{4}{\beta^2} + \frac{4}{\beta^2} e^{-\frac{\beta}{2}L_{\max}} = \frac{4}{\beta} L_{\max} + \frac{8}{\beta^2} e^{-\frac{\beta}{2}L_{\max}} - \frac{8}{\beta^2}.
\end{aligned}$$

For the special case $r = \frac{\beta}{2}$ we obtain

$$\begin{aligned}
& \int_0^{L_{\max}} e^{-\frac{\beta}{2} \max(\ell, \ell')} d\ell' = e^{-\frac{\beta}{2}\ell} \int_0^{\ell} d\ell' + \int_{\ell}^{L_{\max}} e^{-\frac{\beta}{2}\ell'} d\ell' \\
&= \ell e^{-\frac{\beta}{2}\ell} + \frac{1}{-\frac{\beta}{2}} \left(e^{-\frac{\beta}{2}L_{\max}} - e^{-\frac{\beta}{2}\ell} \right) = \ell e^{-\frac{\beta}{2}\ell} + \frac{2}{\beta} e^{-\frac{\beta}{2}\ell} - \frac{2}{\beta} e^{-\frac{\beta}{2}L_{\max}},
\end{aligned}$$

and inserting this on top leads to

$$\begin{aligned}
& \int_0^{L_{\max}} \ell e^{-\frac{\beta}{2}\ell} d\ell + \frac{2}{\beta} \int_0^{L_{\max}} e^{-\frac{\beta}{2}\ell} d\ell - \frac{2}{\beta} e^{-\frac{\beta}{2}L_{\max}} \int_0^{L_{\max}} d\ell \\
&= -\frac{4}{\beta^2} e^{-\frac{\beta}{2}L_{\max}} - \frac{2}{\beta} L_{\max} e^{-\frac{\beta}{2}L_{\max}} + \frac{4}{\beta^2} - \frac{4}{\beta^2} e^{-\frac{\beta}{2}L_{\max}} + \frac{4}{\beta^2} - \frac{2}{\beta} L_{\max} e^{-\frac{\beta}{2}L_{\max}} \\
&= -\frac{8}{\beta^2} e^{-\frac{\beta}{2}L_{\max}} - \frac{4}{\beta} L_{\max} e^{-\frac{\beta}{2}L_{\max}} + \frac{8}{\beta^2}.
\end{aligned}$$

Overall, we obtain the bounds

$$\text{I} \leq \frac{c_{\text{disc}} c_2}{M^2} \begin{cases} \left(\frac{4}{(2r-\beta)^2} \left(e^{(2r-\beta)L_{\max}} - 2e^{(r-\frac{\beta}{2})L_{\max}} + 1 \right) \right) & \text{for } r \neq \frac{\beta}{2}, \\ (L_{\max})^2 & \text{for } r = \frac{\beta}{2}, \end{cases}$$

and

$$\text{II} \leq \frac{c_2}{M} \begin{cases} \left(\frac{2}{(r-\beta)\left(r - \frac{\beta}{2}\right)} e^{(r-\beta)L_{\max}} + \frac{4}{\beta\left(r - \frac{\beta}{2}\right)} e^{-\frac{\beta}{2}L_{\max}} - \frac{4}{(r-\beta)\beta} \right) & \text{for } r \neq \frac{\beta}{2}, \beta, \\ -\frac{8}{\beta^2} e^{-\frac{\beta}{2}L_{\max}} - \frac{4}{\beta} L_{\max} e^{-\frac{\beta}{2}L_{\max}} + \frac{8}{\beta^2} & \text{for } r = \frac{\beta}{2}, \\ \frac{4}{\beta} L_{\max} + \frac{8}{\beta^2} e^{-\frac{\beta}{2}L_{\max}} - \frac{8}{\beta^2} & \text{for } r = \beta. \end{cases}$$

ACKNOWLEDGEMENTS

We thank the anonymous referees for constructive comments leading to a significant improvement of the manuscript and Fabio Musco for helpful discussions on quasi-random numbers.

FUNDING

The work of Cedric Aaron Beschle is funded by the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) – Project-ID 251654672 – SFBTRR 161., The work of Andrea Barth is funded by Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) under Germany’s Excellence Strategy – EXC 2075 – 390740016.

REFERENCES

- [1] A. Abdulle, A. Barth and C. Schwab, Multilevel Monte Carlo methods for stochastic elliptic multiscale PDEs. *Multiscale Model. Simul.* **11** (2013) 1033–1070.
- [2] M. Abramowitz and I.A. Stegun, Handbook of Mathematical Functions with Formulas, Graphs, and Mathematical Tables. *National Bureau of Standards Applied Mathematics Series*. For sale by the Superintendent of Documents. Vol. 55. US Government Printing Office, Washington, DC (1964).
- [3] M. Ainsworth and J.T. Oden, A posteriori error estimation in finite element analysis. *Comput. Methods Appl. Mech. Eng.* **142** (1997) 1–88.
- [4] A. Alexanderian, A brief note on the Karhunen–Loève expansion. Preprint [arXiv:1509.07526](https://arxiv.org/abs/1509.07526) (2015).
- [5] M.S. Alnaes, J. Blechta, J. Hake, A. Johansson, B. Kehlet, A. Logg, C. Richardson, J. Ring, M.E. Rognes and G.N. Wells, The FEniCS project version 1.5. *Arch. Numer. Softw.* **3** (2015) 9–23.
- [6] A. Barth and A. Stein, A study of elliptic partial differential equations with jump diffusion coefficients. *SIAM/ASA J. Uncertain. Quantif.* **6** (2018) 1707–1743.
- [7] C.A. Beschle and A. Barth, Quasi continuous level Monte Carlo for random elliptic PDEs, in Monte Carlo and Quasi-Monte Carlo Methods 2022, edited by A. Hinrichs, P. Kritzer and F. Pillichshammer. Springer Verlag (2023).
- [8] S.C. Brenner and L.R. Scott, The Mathematical Theory of Finite Element Methods, 3rd edition. Vol. 15 of *Texts in Applied Mathematics*. Springer, New York (2008).
- [9] K.A. Cliffe, M.B. Giles, R. Scheichl and A.L. Teckentrup, Multilevel Monte Carlo methods and applications to elliptic PDEs with random coefficients. *Comput. Vis. Sci.* **14** (2011) 3–15.
- [10] G. Detommaso, T. Dodwell and R. Scheichl, Continuous level Monte Carlo and sample-adaptive model hierarchies. *SIAM/ASA J. Uncertain. Quantif.* **7** (2019) 93–116.
- [11] J. Dick and F. Pillichshammer, Discrepancy theory and quasi-Monte Carlo integration, in A Panorama of Discrepancy Theory. Vol. 2107 of *Lecture Notes in Math*. Springer, Cham (2014) 539–619.
- [12] W. Dörfler, A convergent adaptive algorithm for Poisson’s equation. *SIAM J. Numer. Anal.* **33** (1996) 1106–1124.
- [13] K.T. Fang, Y. Wang and P.M. Bentler, Some applications of number-theoretic methods in statistics. *Statist. Sci.* **9** (1994) 416–428.
- [14] J.E. Gentle, Random number generation and Monte Carlo methods, in Statistics and Computing, 2nd edition. Springer, New York (2003).
- [15] M.B. Giles, Multilevel Monte Carlo path simulation. *Oper. Res.* **56** (2008) 607–617.
- [16] M.B. Giles, Multilevel Monte Carlo methods. *Acta Numer.* **24** (2015) 259–328.
- [17] 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) 3668–3694.
- [18] I.G. Graham, F.Y. Kuo, J.A. Nichols, R. Scheichl, C. Schwab and I.H. Sloan, Quasi-Monte Carlo finite element methods for elliptic PDEs with lognormal random coefficients. *Numer. Math.* **131** (2015) 329–368.
- [19] T. Grätsch and K.-J. Bathe, A posteriori error estimation techniques in practical finite element analysis. *Comput. Struct.* **83** (2005) 235–265.
- [20] W. Hackbusch, Elliptic Differential Equations: Theory and Numerical Treatment, 2nd edition. Vol. 18 of *Springer Series in Computational Mathematics*. Springer-Verlag, Berlin (2017).
- [21] A.-L. Haji-Ali, F. Nobile and R. Tempone, Multi-index Monte Carlo: when sparsity meets sampling. *Numer. Math.* **132** (2016) 767–806.
- [22] C.R. Harris, K.J. Millman, S.J. van der Walt, R. Gommers, P. Virtanen, D. Cournapeau, E. Wieser, J. Taylor, S. Berg, N.J. Smith, R. Kern, M. Picus, S. Hoyer, M.H. van Kerkwijk, M. Brett, A. Haldane, J.F. del Río, M. Wiebe, P. Peterson, P. Gérard-Marchant, K. Sheppard, T. Reddy, W. Weckesser, H. Abbasi, C. Gohlke and T.E. Oliphant, Array programming with NumPy. *Nature* **585** (Sept. 2020) 357–362.
- [23] S. Heinrich, Multilevel monte carlo methods, in Large-Scale Scientific Computing. Springer Berlin Heidelberg, Berlin, Heidelberg (2001) 58–67.

- [24] P. Knabner and L. Angermann, Numerical Methods for Elliptic and Parabolic Partial Differential Equations. Vol. 44 of *Texts in Applied Mathematics*. Springer-Verlag, New York (2003).
- [25] L. Kuipers and H. Niederreiter, Uniform Distribution of Sequences. *Pure and Applied Mathematics*. Wiley-Interscience [John Wiley & Sons], New York-London-Sydney (1974).
- [26] F.Y. Kuo, C. Schwab and I.H. Sloan, Quasi-Monte Carlo finite element methods for a class of elliptic partial differential equations with random coefficients. *SIAM J. Numer. Anal.* **50** (2012) 3351–3374.
- [27] F.Y. Kuo, C. Schwab and I.H. Sloan, Multi-level quasi-Monte Carlo finite element methods for a class of elliptic PDEs with random coefficients. *Found. Comput. Math.* **15** (2015) 411–449.
- [28] H. Niederreiter, Random Number Generation and Quasi-Monte Carlo Methods. Vol. 63 of *CBMS-NSF Regional Conference Series in Applied Mathematics*. Society for Industrial and Applied Mathematics (SIAM), Philadelphia, PA (1992).
- [29] A.B. Owen, Randomly permuted (t, m, s) -nets and (t, s) -sequences, in Monte Carlo and Quasi-Monte Carlo Methods in Scientific Computing (Las Vegas, NV, 1994). Vol. 106 of *Lect. Notes Stat.* Springer, New York (1995) 299–317.
- [30] A.B. Owen, Scrambling Sobol’ and Niederreiter-Xing points. *J. Complexity* **14** (1998) 466–489.
- [31] B. Peherstorfer, K. Willcox and M. Gunzburger, Optimal model management for multifidelity Monte Carlo estimation. *SIAM J. Sci. Comput.* **38** (2016) A3163–A3194.
- [32] C.E. Rasmussen and C.K.I. Williams, Gaussian Processes for Machine Learning. *Adaptive Computation and Machine Learning*. MIT Press, Cambridge, MA (2006).
- [33] C.-H. Rhee and P.W. Glynn, Unbiased estimation with square root convergence for SDE models. *Oper. Res.* **63** (2015) 1026–1043.
- [34] I.M. Sobol, Distribution of points in a cube and approximate evaluation of integrals. *Ž. Vyčisl. Mat i Mat. Fiz.* **7** (1967) 784–802.
- [35] A.L. Teckentrup, R. Scheichl, M.B. Giles and E. Ullmann, Further analysis of multilevel Monte Carlo methods for elliptic PDEs with random coefficients. *Numer. Math.* **125** (2013) 569–600.
- [36] M. Vihola, Unbiased estimators and multilevel Monte Carlo. *Oper. Res.* **66** (2018) 448–462.
- [37] P. Virtanen, R. Gommers, T.E. Oliphant, M. Haberland, T. Reddy, D. Cournapeau, E. Burovski, P. Peterson, W. Weckesser, J. Bright, S.J. van der Walt, M. Brett, J. Wilson, K.J. Millman, N. Mayorov, A.R.J. Nelson, E. Jones, R. Kern, E. Larson, C.J. Carey, Í. Polat, Y. Feng, E.W. Moore, J. VanderPlas, D. Laxalde, J. Perktold, R. Cimrman, I. Henriksen, E.A. Quintero, C.R. Harris, A.M. Archibald, A.H. Ribeiro, F. Pedregosa, P. van Mulbregt and SciPy 1.0 Contributors, SciPy 1.0: fundamental algorithms for scientific computing in Python. *Nat. Methods* **17** (2020) 261–272.

Please help to maintain this journal in open access!



This journal is currently published in open access under the Subscribe to Open model (S2O). We are thankful to our subscribers and supporters for making it possible to publish this journal in open access in the current year, free of charge for authors and readers.

Check with your library that it subscribes to the journal, or consider making a personal donation to the S2O programme by contacting subscribers@edpsciences.org.

More information, including a list of supporters and financial transparency reports, is available at <https://edpsciences.org/en/subscribe-to-open-s2o>.