

Article

Randomly Shifted Lattice Rules with Importance Sampling and Applications

Hejin Wang ^{1,2,*} and Zhan Zheng ^{2,†}¹ École Nationale de la Statistique et de L'administration Économique Paris, 91120 Paris, France² Department of Mathematical Sciences, Tsinghua University, Beijing 100084, China; zhengz15@tsinghua.org.cn

* Correspondence: hejin.wang@ensae.fr; Tel.: +33-07-49-67-05-67

† These authors contributed equally to this work.

Abstract: In financial and statistical computations, calculating expectations often requires evaluating integrals with respect to a Gaussian measure. Monte Carlo methods are widely used for this purpose due to their dimension-independent convergence rate. Quasi-Monte Carlo is the deterministic analogue of Monte Carlo and has the potential to substantially enhance the convergence rate. Importance sampling is a widely used variance reduction technique. However, research into the specific impact of importance sampling on the integrand, as well as the conditions for convergence, is relatively scarce. In this study, we combine the randomly shifted lattice rule with importance sampling. We prove that, for unbounded functions, randomly shifted lattice rules combined with a suitably chosen importance density can achieve convergence as quickly as $O(N^{-1+\epsilon})$, given N samples for arbitrary ϵ values under certain conditions. We also prove that the conditions of convergence for Laplace importance sampling are stricter than those for optimal drift importance sampling. Furthermore, using a generalized linear mixed model and Randleman–Bartter model, we provide the conditions under which functions utilizing Laplace importance sampling achieve convergence rates of nearly $O(N^{-1+\epsilon})$ for arbitrary ϵ values.

Keywords: importance sampling; quasi-Monte Carlo; lattice rules

MSC: 65C20; 65C05; 91G60



Citation: Wang, H.; Zheng, Z.

Randomly Shifted Lattice Rules with Importance Sampling and Applications. *Mathematics* **2024**, *12*, 630. <https://doi.org/10.3390/math12050630>

Academic Editors: David Carfi and Manuel Alberto M. Ferreira

Received: 25 December 2023

Revised: 27 January 2024

Accepted: 19 February 2024

Published: 21 February 2024



Copyright: © 2024 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

In finance and statistics, many computational challenges involve calculating expectations, frequently represented as integrals over Gaussian measures. Typical examples include the valuation of assets driven by Brownian motion in security pricing. Monte Carlo (MC) methods, characterized by an $O(N^{-1/2})$ convergence rate for square-integrable integrands of sample size N , are widely used in expectation estimation [1]. Despite their dimension-independent rate, they may fall short in practical applications. To enhance the order of convergence and reduce variance, we combine quasi-Monte Carlo (QMC) methods with importance sampling (IS). QMC methods are deterministic counterparts of MC that use low-discrepancy sequences [2]. In the last 30 years, the use of these methods in scientific computations has grown significantly, particularly in the fields of finance and statistics. Given certain conditions, the error bound of QMC quadrature can reach $O(N^{-1+\epsilon})$ for a sample size N and any positive ϵ . This ϵ value is associated with the dimension of the integrand [2]. In practical scenarios, randomized quasi-Monte Carlo (RQMC) methods are commonly used. They preserve the convergence characteristics of QMC while providing the added advantage of making computational efficiency analysis more accessible, including tasks like error estimation and confidence interval construction. RQMC quadratures come in various forms, including scrambled digital nets and randomly shifted lattice rules, as explored in [3,4]. For a comprehensive understanding of both QMC and RQMC, readers are referred to [5,6]. Recently, researchers [7] successfully combined scrambled digital

nets with IS, demonstrating that an approximate $O(N^{-1+\epsilon})$ root mean square error can be achieved. This level of accuracy depends on the specific boundary growth conditions of the integrand and the optimization of the IS density. This study focuses on the effects of combining randomly shifted lattice rules with IS.

The use of IS as an effective technique for reducing variance has been well-established in the MC literature, with its application in security pricing having been explored in earlier studies, such as [8]. Over the years, further research has highlighted the effectiveness of IS in dealing with rare events, as detailed in sources like [9]. This has revealed that the usefulness of IS surpasses traditional methods of variance reduction. The success of IS heavily depends on choosing the appropriate importance density. For a detailed discussion on IS, refer to [10]. When selecting an importance density, two types of importance densities stand out: optimal drift importance sampling (ODIS) and Laplace importance sampling (LapIS). ODIS utilizes a multivariate normal density that shares the covariance matrix with the prior density, while LapIS uses a general multivariate normal density, aligning its mean and covariance matrices with the mode and curvature of the integrand [11]. Combining IS into QMC quadrature makes deriving a theoretical convergence rate more complex than with MC methods [7]. The effectiveness of MC is generally unaffected by the dimension and smoothness of the problem, except for the requirement of square-integrability, and these factors significantly influence the performance of QMC. Kuo and Dunsmuir et al. [12] have shown that QMC outperforms MC in the context of log-likelihood integrals. Furthermore, Dick et al. [13] developed a weighted discrepancy bound specifically for QMC when combined with IS, providing clear error bounds for integrands that are sufficiently regular. The study in [7] highlighted potential boundary singularities in the unit hypercube due to the IS density. These works prompted us to investigate the error bounds associated with the combination of IS and randomly shifted lattice rules. It is important to note that the core theoretical basis of our study is based on the reproducing kernel Hilbert space (RKHS) framework.

This contributions of this study are as follows. Our research is focused on the combination of the shifted rank-1 lattice rule with IS. This combination has been shown to effectively reduce variance. Our study also analyzes the convergence error of the integral estimation of unbounded functions, using importance sampling combined with the shifted lattice rule. Moreover, we conduct a preliminary analysis to investigate how varying importance densities affect functions. This analysis is crucial for determining the effectiveness of different IS strategies in diverse scenarios. In the realms of finance and statistics, we identify specific conditions under which functions utilizing IS demonstrate enhanced convergence rates. Although our study does not investigate the optimization of weights, the numerical results that we have obtained indicate that the observed numerical effects are relatively robust toward different weights.

The structure of this paper is as follows. Section 2 offers an overview of IS and the randomly shifted lattice rule. Section 3 primarily focuses on the convergence error incurred when estimating the integrals of unbounded functions within RKHS by using importance sampling combined with shifted lattice rules. Section 4 presents case studies illustrating the effects of ODIS and LapIS, and Section 5 concludes the paper.

2. Preliminary

2.1. Monte Carlo and Quasi-Monte Carlo Methods

Consider the computation of a d -dimensional integral over a unit hypercube:

$$I(g) = \int_{(0,1)^d} g(z) dz. \quad (1)$$

For numerical approximation, we employ

$$\hat{I}_N(g) = \frac{1}{N} \sum_{i=1}^N g(z_i). \quad (2)$$

In the MC approach, points z_i are randomly selected within the unit hypercube. If the variance of $g(z)$ is finite, i.e.,

$$\sigma^2 := \int_{(0,1)^d} (g(z) - I(g))^2 dz < \infty, \quad (3)$$

then, by the central limit theorem, the root mean square error (RMSE) of the MC method is

$$\sqrt{\mathbb{E}[(\hat{I}_N(g) - I(g))^2]} = \frac{\sigma}{\sqrt{N}}. \quad (4)$$

This suggests an RMSE rate of $O(N^{-1/2})$ for square-integrable integrands.

To improve the convergence, we use QMC methods. The QMC estimator is similar to that in MC methods, but it utilizes a deterministic set of low-discrepancy points $\mathcal{P} = \{z_1, \dots, z_N\}$, in contrast to the independently and identically distributed points used in MC. We employ lattice rules as the low-discrepancy point sets.

The typical error bound for QMC methods follows the Koksma–Hlawka inequality:

$$|\hat{I}_N(g) - I(g)| \leq D^*(\mathcal{P}) V_{\text{HK}}(g), \quad (5)$$

where $D^*(\mathcal{P})$ denotes the star-discrepancy of the point set $\mathcal{P}$ and $V_{\text{HK}}(\cdot)$ represents the variation in the Hardy–Krause sense of a function over the unit hypercube. With low-discrepancy sequences achieving a star-discrepancy of $O(N^{-1}(\log N)^d)$ and assuming bounded variation, the QMC estimator guarantees a deterministic error bound of $O(N^{-1}(\log N)^d)$, which is superior to MC for a fixed dimension d .

2.2. Importance Sampling

In this study, we examine integrals with respect to

$$C = \int_{\mathbb{R}^d} G(z) p(z; \mu_0, \Sigma_0) dz, \quad (6)$$

where $p(z; \mu, \Sigma)$ represents the probability density function of a d -dimensional normal distribution with mean vector μ and covariance matrix Σ , that is,

$$\frac{1}{\sqrt{(2\pi)^k |\Sigma|}} \exp\left(-\frac{1}{2}(z - \mu)^\top \Sigma^{-1}(z - \mu)\right). \quad (7)$$

Employing IS involves using a proposal importance density q to reformulate the integral as

$$\begin{aligned} C &= \int_{\mathbb{R}^d} G(z) \frac{p(z; \mu_0, \Sigma_0)}{q(z)} q(z) dz \\ &= \int_{\mathbb{R}^d} G(z) W(z) q(z) dz \\ &= \int_{\mathbb{R}^d} G_{\text{IS}}(z) q(z) dz, \end{aligned} \quad (8)$$

where

$$W(z) = \frac{p(z; \mu_0, \Sigma_0)}{q(z)} \quad (9)$$

is defined as the likelihood ratio (LR) function and $G_{\text{IS}}(z) = G(z)W(z)$ denotes the transformed integrand.

Choosing an IS density among normal densities, i.e., $q(z) = p(z; \mu, \Sigma)$, leads to

$$W(z) = \sqrt{\frac{\det(\Sigma)}{\det(\Sigma_0)}} \exp\left(\frac{1}{2}(z - \mu)^\top \Sigma^{-1}(z - \mu) - \frac{1}{2}(z - \mu_0)^\top \Sigma_0^{-1}(z - \mu_0)\right). \quad (10)$$

Using the affine transformation property of the normal distribution, we can express the integral C as

$$\begin{aligned} C &= \int_{\mathbb{R}^d} G_{\text{IS}}(\mu + Lz) p(z; 0, I_d) dz \\ &= \int_{(0,1)^d} G_{\text{IS}}(\mu + L\Phi^{-1}(u)) du, \end{aligned} \quad (11)$$

where L signifies any square root decomposition of the covariance matrix Σ (satisfying $LL^T = \Sigma$). Here, $\Phi(\cdot)$ denotes the cumulative distribution function (CDF) of the standard normal distribution in d dimensions, and $\Phi^{-1}(\cdot)$ is its inverse, which is applied component-wise. The LR function is now represented as

$$W(\mu + Lz) = \sqrt{\frac{\det(\Sigma)}{\det(\Sigma_0)}} \exp\left(\frac{1}{2}z^T z - \frac{1}{2}(\mu - \mu_0 + Lz)^T \Sigma_0^{-1}(\mu - \mu_0 + Lz)\right). \quad (12)$$

It is noteworthy that it is sufficient to use $\mu = \mu_0$ and $\Sigma = \Sigma_0$ when IS is not employed.

The proper selection of parameters μ and Σ can reduce the error of this system. In the context of MC simulations, this is tantamount to reducing the variance σ^2 , as indicated in Equation (4). Assuming that G is nonnegative for all z , the optimal zero-variance IS density would be

$$q_{\text{opt}}(z) := \frac{1}{C} G(z) p(z; \mu_0, \Sigma_0), \quad (13)$$

where C is unknown. Although this is impractical due to the unknown value of C , this formulation inspires the selection of an IS density that approximates the behavior of q_{opt} .

We define $H(z) = \log(q_{\text{opt}}(z))$. Assuming that $H(z)$ is differentiable and unimodal, the mode of $H(z)$ by μ_* may be denoted via

$$\mu_* = \arg \max_{z \in \mathbb{R}^d} H(z) = \arg \max_{z \in \mathbb{R}^d} G(z) p(z; \mu_0, \Sigma_0). \quad (14)$$

A second-order Taylor expansion about μ_* yields

$$H(z) \approx H(\mu_*) - \frac{1}{2}(z - \mu_*)^\top \Sigma_*^{-1}(z - \mu_*), \quad (15)$$

where Σ_* is determined by

$$\Sigma_* = -\nabla^2 H(\mu_*)^{-1} = \left(\Sigma_0 - \nabla^2 \log(G(\mu_*))\right)^{-1}. \quad (16)$$

Consequently, we approximate $q_{\text{opt}}(z)$ as

$$q_{\text{opt}}(z) \approx \exp\left(H(\mu_*) - \frac{1}{2}(z - \mu_*)^\top \Sigma_*^{-1}(z - \mu_*)\right) \propto p(z; \mu_*, \Sigma_*). \quad (17)$$

Thus, $p(z; \mu_*, \Sigma_*)$ represents a density closely resembling the optimal density. The LapIS density adopts $\mu = \mu_*$ and $\Sigma = \Sigma_*$, $\Sigma_* = L_*^T L_*$, while the ODIS density selects $\mu = \mu_*$ but maintains $\Sigma = \Sigma_0$.

2.2.1. ODIS Estimator

The importance density for ODIS is $p(z; \mu_*, \Sigma_0)$. The function $f_O(z)$ is defined as

$$f_O(z) = (2\pi)^{-d/2} \exp(H(L_* z + \mu_*)) \prod_{j=1}^d \frac{1}{\rho(z_j)}, \quad (18)$$

and the integral becomes

$$C = \int_{\mathbb{R}^d} f_O(z) \prod_{j=1}^d \rho(z_j) dz = \int_{(0,1)^d} f_O(\Phi^{-1}(u)) du. \quad (19)$$

2.2.2. LapIS Estimator

The importance density for LapIS is $p(z; \mu_*, \Sigma_*)$. The function $f_L(z)$ for LapIS is defined as

$$f_L(z) = (2\pi)^{-d/2} \frac{\det(L_*)}{\sqrt{\det(\Sigma)}} \exp(H(L_* z + \mu_*)) \prod_{j=1}^d \frac{1}{\rho(z_j)}, \quad (20)$$

and the integral takes the form

$$\int_{\mathbb{R}^d} f_L(z) \prod_{j=1}^d \rho(z_j) dz = \int_{(0,1)^d} f_L(\Phi^{-1}(u)) du. \quad (21)$$

It is noteworthy that LapIS does not always outperform ODIS (for example, see [7,11]). Generally, in an MC framework, LapIS tends to be effective if $q_{\text{opt}}(z)$ closely resembles a normal distribution. From a QMC perspective, this situation becomes increasingly complex as the performance of QMC is further influenced by additional factors such as the boundary growth of the integrand when using scrambled digital nets. This aspect is further explored in the context of randomly shifted lattice rules in the subsequent sections.

2.3. Reproducing Kernel Hilbert Space

RKHS theory offers a framework for error analysis. Recent developments in this system can be found in [14–16].

Consider integrals with respect to a probability density, which is a product of univariate marginal densities

$$C(f) = \int_{\mathbb{R}^d} f(z) \prod_{i=1}^d \phi(z_i) dz. \quad (22)$$

Let Φ denote the cumulative distribution function (CDF) of ϕ :

$$\Phi(x) = \int_{-\infty}^x \phi(y) dy, \quad (23)$$

and Φ^{-1} denote its inverse. For instance, if $\phi(x) = p(x; 0, 1)$, Φ represents the CDF of the standard normal distribution. Assume that $\mathcal{F}$ is a Hilbert space on $\mathbb{R}^d$. Suppose that $K_{\mathcal{F}}$ is a reproducing kernel, and the function $K_{\mathcal{F}} : \mathbb{R}^d \times \mathbb{R}^d \rightarrow \mathbb{R}$ has the following properties:

- Functional property: $K_{\mathcal{F}}(\cdot, x) \in \mathcal{F}, \forall x \in \mathbb{R}^d$;
- Symmetry: $K_{\mathcal{F}}(x, y) = K_{\mathcal{F}}(y, x), \forall x, y \in \mathbb{R}^d$;
- Reproducing property: $\langle f, K_{\mathcal{F}}(\cdot, x) \rangle = f(x), \forall x \in \mathbb{R}^d$ and $f \in \mathcal{F}$.

Additionally, it is required that

$$\int_{\mathbb{R}^d} K_{\mathcal{F}}(x, x) \prod_{i=1}^d \phi(x_i) dx < \infty. \quad (24)$$

This leads to

$$C(f) = \int_{\mathbb{R}^d} \langle f, K_{\mathcal{F}}(\cdot, x) \rangle \prod_{i=1}^d \phi(x_i) dx = \langle f, h \rangle, \quad (25)$$

where

$$h(x) = \int_{\mathbb{R}^d} K_{\mathcal{F}}(x, y) \prod_{i=1}^d \phi(y_i) dy. \quad (26)$$

Equation (24) ensures that Equations (25) and (26) are finite. The initial error is

$$e_0(\mathcal{F}) = \sup_{\|f\|_{\mathcal{F}} \leq 1} |C(f)| = \sqrt{\langle h, h \rangle}, \quad (27)$$

yielding

$$|e_0(\mathcal{F})|^2 = \langle h, h \rangle = \int_{\mathbb{R}^d} \int_{\mathbb{R}^d} K_{\mathcal{F}}(x, y) \prod_{i=1}^d [\phi(x_i) \phi(y_i)] dx dy. \quad (28)$$

The transformation ϕ establishes a one-to-one correspondence between functions in the space $\mathcal{F}$ and those in a Hilbert space $\mathcal{G}$ defined on $(0, 1)^d$. This mapping induces an isometry of

$$f \in \mathcal{F} \longleftrightarrow g = f(\Phi^{-1}(\cdot)) \in \mathcal{G}.$$

Consequently, the following equality holds:

$$\|f\|_{\mathcal{F}} = \|g\|_{\mathcal{G}}.$$

Recalling the isometric space $\mathcal{G}$ of $\mathcal{F}$, the integral formulation converts to

$$C(f) = \int_{(0,1)^d} g(u) du := I(g). \quad (29)$$

Furthermore, $\mathcal{G}$, similar to $\mathcal{F}$, is an RKHS equipped with a reproducing kernel of

$$K_{\mathcal{G}}(u, v) = K_{\mathcal{F}}(\Phi^{-1}(u), \Phi^{-1}(v)). \quad (30)$$

This implies that the initial error in $\mathcal{G}$ is analogous to that in $\mathcal{F}$:

$$e_0(g) = \sup_{\|g\|_{\mathcal{G}} \leq 1} |I(g)| = \sup_{\|f\|_{\mathcal{F}} \leq 1} |C(f)| = e_0(f). \quad (31)$$

Consider applying a QMC rule to estimate the integral Equation (29) via

$$\hat{I}_N(g) = \frac{1}{N} \sum_{i=1}^N g(u_i), \quad (32)$$

where the worst-case error in $\mathcal{G}$ is defined as

$$e_w(\hat{I}_N, \mathcal{G}) = \sup_{\|g\|_{\mathcal{G}} \leq 1} |\hat{I}_N(g) - I(g)|. \quad (33)$$

Therefore, the error in approximating $C(f)$ using the transformed QMC rule is bounded by

$$|\hat{I}_N(f \circ \Phi^{-1}) - C(f)| = |\hat{I}_N(g) - I(g)| \leq e_w(\hat{I}_N, \mathcal{G}) \|g\|_{\mathcal{G}} = e_w(\hat{I}_N, \mathcal{G}) \|f\|_{\mathcal{F}}. \quad (34)$$

In the context of QMC methods, we utilize a randomly shifted rank-1 lattice rule defined by the sequence $u_i, i = 1, 2, \dots, N$, where each u_i value is given by

$$u_i = \left\{ \frac{iz}{N} + \zeta \right\}, \quad i = 1, 2, \dots, N. \quad (35)$$

Here, $\{\cdot\}$ denotes the component-wise fractional part, $\zeta \in (0, 1)^d$ is a uniformly distributed random variable, and $z \in \mathcal{Z}_N^d$ represents the generating vector.

Let us define

$$\mathcal{Z}_N = \{z \in \mathbb{Z} \mid 1 \leq z \leq N-1, \gcd(z, N) = 1\}, \quad (36)$$

which represents the set of positive integers not exceeding $N-1$ and coprime to N . We consider the shift-average worst-case error, depending solely on the generating vector, as described in [17]:

$$\begin{aligned} |e_N^{\text{sa}}(z)|^2 &= \int_{(0,1)^d} |e_w(\hat{I}_N, \mathcal{G})|^2 d\zeta \\ &= - \int_{(0,1)^d} \int_{(0,1)^d} K_{\mathcal{G}}(u, v) du dv + \frac{1}{N} \sum_{i=1}^N K_{\mathcal{G}}^{\text{sik}}\left(\left\{\frac{iz}{N}\right\}\right), \end{aligned} \quad (37)$$

where $K_{\mathcal{G}}^{\text{sik}}$ is the shift-invariant kernel, defined as

$$K_{\mathcal{G}}^{\text{sik}}(u, v) = \int_{(0,1)^d} K_{\mathcal{G}}(\{u + \zeta\}, \{v + \zeta\}) d\zeta. \quad (38)$$

In our analysis, we simplify the notation of $K_{\mathcal{G}}^{\text{sik}}$ based on the difference between u and v as

$$K_{\mathcal{G}}^{\text{sik}}(\{u - v\}) := K_{\mathcal{G}}^{\text{sik}}(\{u - v\}, 0) = K_{\mathcal{G}}^{\text{sik}}(u, v). \quad (39)$$

For a linear functional space $\mathcal{F}$ over $\mathbb{R}^d$ with a weight function ψ and positive weight coefficients $\gamma_k, k = 1, 2, \dots, d$, the inner product is defined as [18]

$$\langle f, g \rangle_{\mathcal{F}} = \sum_{u \subseteq \{1:d\}} \left(\prod_{k \in u} \frac{1}{\gamma_k} \int_{\mathbb{R}^{|u|}} \frac{\partial^{|u|} f}{\partial x_u}(x_u, 0) \frac{\partial^{|u|} g}{\partial x_u}(x_u, 0) \prod_{k \in u} \psi^2(x_k) dx_u \right), \quad (40)$$

where $(x_u, 0)$ denotes a d -dimensional vector, with x_k for $k \in u$ and otherwise zero. The reproducing kernel components satisfy

$$\int_{\mathbb{R}} K_{\mathcal{F},j}(y, y) \psi(y) dy < \infty, \quad (41)$$

where $K_{\mathcal{F},j}$ depends on γ_j ,

$$K_{\mathcal{F},j}(x, y) = 1 + \gamma_j \eta(x, y), \quad (42)$$

with η defined as

$$\eta(x, y) = \begin{cases} \int_0^{\max(x,y)} \frac{1}{\psi^2(t)} dt, & x, y > 0, \\ \int_{\max(x,y)}^0 \frac{1}{\psi^2(t)} dt, & x, y < 0, \\ 0, & \text{otherwise.} \end{cases} \quad (43)$$

The shift-invariant kernel, as detailed in [14], is expressed as

$$K_{\mathcal{G}}^{\text{sik}}(u, v) = \sum_{u \in \{1:d\}} \gamma_u \prod_{i \in u} \theta(\{u_i - v_i\}), \quad u, v \in (0, 1)^d, \quad (44)$$

where

$$\theta(x) = \int_{\Phi^{-1}(x)}^0 \frac{\Phi(t) - x}{\psi^2(t)} dt + \int_{\Phi^{-1}(1-x)}^0 \frac{\Phi(t) - 1 + x}{\psi^2(t)} dt, \quad x \in (0, 1). \quad (45)$$

Expanding $\theta(x)$ into a Fourier series, we obtain

$$\theta(x) = \sum_{h \in \mathbb{Z}} \hat{\theta}(h) \exp(2\pi i h x), \quad (46)$$

with the Fourier coefficients given by

$$\hat{\theta}(h) = \int_0^1 \theta(x) \exp(-2\pi i h x) dx. \quad (47)$$

The explicit form of $\hat{\theta}(h)$ in the Gaussian setting is shown in Appendix A. In this space setting, to satisfy Equation (41), we require that, for any finite c ,

$$\int_{-\infty}^c \frac{\Phi(t)}{\psi^2(t)} dt + \int_c^{\infty} \frac{1 - \Phi(t)}{\psi^2(t)} dt < \infty \quad (48)$$

as shown in [15].

3. Main Results

This section primarily addresses the theoretical convergence rate of the randomly shifted lattice rule combined with IS methods for functions defined in $\mathbb{R}^d$. We establish the conditions under which IS-based estimators achieve a convergence rate of $O(N^{-1+\epsilon})$, with $\epsilon > 0$ being arbitrarily small.

Error Bounds for Importance Sampling

We first review a Theorem proposed by Nichols and Kou [15] that established a connection between Fourier analysis and the shift-average worst-case error.

Assumption 1. Assume there exist $1/2 < r < 1$ and $M > 0$ such that

$$\hat{\theta}(h) \leq M|h|^{-2r} \text{ for all } h \neq 0. \quad (49)$$

Theorem 1. Let $f \in \mathcal{F}$, with density function ϕ , weight parameters γ_u , and weight function ψ . If Assumption 1 is satisfied with a parameter $1/2 < r < 1$, then a randomly shifted lattice rule with N points in d dimensions can be constructed by a CBC algorithm such that, for all $\lambda \in [\frac{1}{2}, r)$, the RMSE is

$$\sqrt{E^\Delta |\hat{I}_N(f \circ \Phi^{-1}) - C(f)|^2} \leq \left(\frac{1}{\varphi(N)} \sum_{\emptyset \neq u \subset 1:s} \gamma_u^{\frac{1}{2\lambda}} \left(2M^{\frac{1}{2\lambda}} \zeta\left(\frac{r}{\lambda}\right) \right)^{|u|} \right)^\lambda \|f\|_{\mathcal{F}}, \quad (50)$$

where $\zeta(x) := \sum_{k=1}^{\infty} k^{-x}$ denotes the Riemann zeta function and $\varphi(N) := |\{1 \leq i \leq N-1 : \gcd(i, N) = 1\}|$ denotes the Euler function. If

$$\sum_{|u| < \infty} \gamma_u^{\frac{1}{2\lambda}} \left(2M^{\frac{1}{2\lambda}} \zeta\left(\frac{r}{\lambda}\right) \right)^{|u|} < \infty \quad (51)$$

further holds, the RMSE has an upper bound independently of d .

Proof. The proof is the combination of Theorems 7 and 8 in [15]. $\square$

Note that the Euler function $\varphi(N) = N - 1$ when N is prime. The parameter r of commonly encountered combinations of distribution and weight functions have been computed in [15]. Here, we select some combinations of the probability density ϕ and weight function ψ that satisfy (48) from [15] in Table 1, which are used later.

Table 1. The choice of the parameter r in Assumption 1 for some common combinations of the probability density ϕ and weight function ψ .

$\phi(x)$	$\frac{\exp(-x^2/2\nu)}{\sqrt{2\pi\nu}}$	$\frac{1}{\sqrt{\nu\pi}} \frac{\Gamma(\frac{\nu+1}{2})}{\Gamma(\frac{\nu}{2})} (1 + \frac{x^2}{\nu})^{-\frac{\nu+1}{2}}$
$\psi(x) = \exp(-x^2/(2\alpha^2))$	$r = 1 - \frac{\nu}{\alpha^2}, \alpha > 2\nu$	-
$\psi(x) = (1 + x)^{-\alpha}$	$r = 1 - \delta, \delta \in (0, \min(\frac{1}{2}, \frac{9}{8}\alpha\nu))$	$r = 1 - \frac{2\alpha+1}{2\nu}, 2\alpha+1 < \nu$

Since the Hardy–Krause variation of unbounded functions is infinite, if we want to determine whether the QMC method is suitable for these functions, RKHS theory provides us with a great tool. To ascertain the achievable convergence order using the lattice rule, it suffices to examine whether the function adheres to the conditions stipulated in RKHS.

According to [15], when we select product weights of the form $\gamma_u = \prod_{j \in u} \gamma_j$, which satisfy the condition

$$\sum_{j=1}^{\infty} \gamma_j^{\frac{1}{2r_2}} < \infty, \quad (52)$$

then (51) holds and it is possible to achieve a convergence error (Equation (50)) for a fixed dimension d .

The remaining hidden condition indicates that G_{IS} needs to be in the RKHS. We next list several conditions for the log integrand $\hat{g} = \log(G_{IS}(z))$.

Assumption 2 (Upper Bound). $\hat{g}(z) = O(\|z\|^\beta)$ with $\beta < 2$.

Assumption 3 (Minimax Eigenvalue). $-\min_{i \in \{1, \dots, d\}} h_i \cdot \max_{j \in \{1, \dots, d\}} \lambda_j < \frac{1}{\alpha}$, where h_i and λ_j are the eigenvalues of $\nabla^2 \hat{g}(\mu_*)$ and Σ_* , respectively.

Assumption 4 (Semi-Positive Definiteness). $\nabla^2 \hat{g}(\mu_*)$ is semi-definite.

The fulfillment of these assumptions is vital in ensuring that the function G_{IS} belongs to the RKHS, thereby facilitating the effective application of the established results.

We can now state the theoretical convergence rate of the randomly shifted lattice rule combined with ODIS or LapIS.

Theorem 2. Let the function G_{IS} be the integrand in Equation (11). Assume that the function G_{IS} belongs to an RKHS equipped with the weight function $\psi(z) \equiv \exp(-\frac{z^2}{2\alpha^2})$, with $\alpha^2 > 2$. Then, using the lattice rule obtained by the CBC algorithm yields an error rate of the estimator as

$$\sqrt{\mathbb{E}|(\hat{I}_N(G_{IS} \circ \Phi^{-1}) - C)|^2} = O(N^{-1+\frac{1}{\alpha^2}+\delta}), \quad \text{where } \delta > 0. \quad (53)$$

For ODIS, Assumption 2 offers a sufficient condition for G_{IS} to qualify as a member of the RKHS. For LapIS, both Assumptions 2 and 3 are required.

Proof. Under the condition $\beta < 2$, Assumption 2 ensures that the expression $\hat{g}(L_*z + \mu_*)$ is dominated by $\frac{z^T z}{2\alpha^2}$. This implies that $\exp(\hat{g}(L_*z + \mu_*) - \frac{z^T z}{2\alpha^2})$ approaches zero as x tends towards infinity. For ODIS, Equation (18) confirms that Assumption 2 suffices for f_O to belong to RKHS. In contrast, for LapIS, Equation (20) requires additional conditions, which may be addressed by examining the dominant term $\exp[D(z)]$, where

$$D(z) = \hat{g}(L_*z + \mu_*) - \frac{1}{2}(L_*z)^T \nabla^2 \hat{g}(\mu_*) (L_*z). \quad (54)$$

Using Assumption 2, we can govern the first term. Incorporating Assumption 3, the singular value decomposition of $\nabla^2 g(\mu_*)$, represented as $Q^T T Q$, shows that Q is an orthogonal matrix and T is a diagonal matrix containing the eigenvalues. The term

$$-\frac{1}{2}(L_* z)^T \nabla^2 \hat{g}(\mu_*)(L_* z) = -\frac{1}{2}(Q L_* z)^T T (Q L_* z), \quad (55)$$

is then dominated by

$$-\frac{h}{2}\|Q L_* z\|^2 \leq -\frac{h}{2}\|Q\|_2^2 \|L_*\|_2^2 \|z\|^2 \leq -\frac{h}{2} \max_{j \in 1:d} \lambda_j \|z\|^2. \quad (56)$$

The second inequality holds because the spectral norm of an orthogonal matrix is 1, and the spectral norm of L_* is the nonnegative square root of the maximum eigenvalue of $L_*^T L_*$, which equals Σ_* . Assumption 3 implies that $\frac{1}{2\alpha} > -\frac{h}{2} \max_{j \in 1:d} \lambda_j$, ensuring that the integrand remains within the RKHS. Then, we can obtain the results by using Table 1. $\square$

Assumption 3 can also be substituted with Assumption 4, which provides an alternative solution.

Remark 1. Considering β in Assumption 2, if $\beta = 2$, then

$$\begin{aligned} \hat{g}(L_* z + \mu_*) &= O(\|L_* z\|^2) \\ &\leq O(\|L_*\|^2 \|z\|^2) \\ &= O((\max_{j \in 1:d} \lambda_j) \|z\|^2). \end{aligned}$$

As the RKHS contains $\exp(\gamma \|z\|^2)$ for $\gamma < \frac{1}{2\alpha}$, it follows that $\max_{j \in 1:d} \lambda_j < \frac{1}{2\alpha}$. Therefore, a more precise format of Assumption 2 asserts that

$$\hat{g}(z) = O(\|z\|^\beta), \beta < 2, \quad (57)$$

or

$$\hat{g}(z) = O(\|z\|^2), \quad \max_{j \in 1:d} \lambda_j < \frac{1}{2\alpha}. \quad (58)$$

And we also find this result if the function has a second-order exponential growth. Then, using ODIS, we can obtain a convergence rate of nearly $O(N^{-1})$ by letting α grow to infinity while using LapIS; we cannot let α be larger, to make sure that the integrand belongs to RKHS.

Remark 2. Regarding the weight functions, they should be carefully chosen for LapIS. For example, $\psi(x) = \exp(-\frac{|x|}{\alpha})$ cannot be selected as the weight function. LapIS introduces the term

$$\exp(-\frac{1}{2}(L_* z)^T \nabla^2 \hat{g}(\mu_*)(L_* z))$$

to the function, which brings a second-order growth to the function and may not belong to the corresponding RKHS, thereby potentially conflicting with the desired conclusion. This restriction, however, does not apply to ODIS, which does not introduce a second-order growth term. As a result, ODIS offers more flexibility in the choice of weight functions than LapIS, indicating that the convergence conditions for LapIS are more stringent.

4. Case Studies

The requirements of growth for the integrand in LapIS are less stringent compared to those in ODIS. Therefore, in the subsequent sections on the generalized linear mixed model (GLMM) and the Randleman–Bartter model, it is sufficient to verify that the integrand for LapIS lies within the RKHS, providing conditions under which the function resides in the

RKHS. When calculating the expected values for Asian options, although the conditions for the inclusion of the function in the RKHS cannot be provided, favorable results are still observed by applying IS combined with the lattice rule. The theoretical justification of these observations is a subject of our future research. The code behind these examples can be found in <https://github.com/wanghj20/Importance-sampling-with-lattice-rule.git>. (accessed on 26 January 2024).

4.1. Statistical Models

In this section, we consider two commonly encountered models in statistical problems.

4.1.1. Generalized Linear Mixed Model

Next, we discuss a class of highly structured statistical models known as GLMMs (see [12]). For simplicity, the likelihood is represented as

$$L(\beta, \kappa, \sigma) = \int_{\mathbb{R}^d} \prod_{j=1}^d \frac{\exp(y_j(\omega_j + \beta) - \exp(\omega_j + \beta))}{y_j} \frac{\exp(-\frac{1}{2}\omega^T \Sigma^{-1} \omega)}{\sqrt{(2\pi)^d \det(\Sigma)}} d\omega, \quad (59)$$

where $y = (y_1, y_2, \dots, y_d)$ denotes the data of nonnegative integers and Σ is the covariance matrix with entries $\Sigma_{i,j} = \frac{\sigma^2 \kappa^{|i-j|}}{1-\kappa^2}$ for $i, j = 1, 2, \dots, d$. The objective is to maximize the log-likelihood $\log L(\beta, \kappa, \sigma)$ given y with respect to (β, κ, σ) , subject to $\kappa \in (0, 1)$ and $\sigma > 0$.

The integral in Equation (59) can be rewritten as $\int_{\mathbb{R}^d} \exp(F(\omega)) d\omega$, neglecting a normalizing constant, where

$$F(\omega) = \sum_{j=1}^d (y_j(\omega_j + \beta) - \exp(\omega_j + \beta)) - \frac{1}{2} \omega^T \Sigma^{-1} \omega. \quad (60)$$

As F is a unimodal function, the optimum value of this system is given by

$$\omega_{\star} = \arg \max_{\omega \in \mathbb{R}^d} F(\omega), \quad (61)$$

which is determined by solving

$$\nabla F(\omega) = y - e^{\beta} \exp(\omega) - \Sigma^{-1} \omega = 0. \quad (62)$$

The Hessian matrix at the optimum value is expressed as

$$\nabla^2 F(\omega_{\star}) = -e^{\beta} \text{diag}(e^{\omega_{\star 1}}, e^{\omega_{\star 2}}, \dots, e^{\omega_{\star d}}) - \Sigma^{-1}. \quad (63)$$

In the framework of LapIS, the covariance matrix $\Sigma_{\star}$ is defined by

$$\Sigma_{\star} = (-\nabla^2 F(\omega_{\star}))^{-1}, \quad (64)$$

and the integral $\int_{\mathbb{R}^d} \exp(F(\omega)) d\omega$ is transformed as

$$\begin{aligned} \int_{\mathbb{R}^d} \exp(F(\omega)) d\omega &= \det(L_{\star}) \int_{\mathbb{R}^d} \exp(F(L_{\star} v) + \omega_{\star}) dv \\ &= \det(L_{\star}) \int_{(0,1)^d} \exp(F(L_{\star} \Phi^{-1}(u) + \omega_{\star})) \prod_{j=1}^d \frac{1}{\rho(\Phi^{-1}(u_j))} du, \end{aligned} \quad (65)$$

where ρ and Φ represent the PDF and CDF of the standard normal distribution, respectively.

Following the insights in [18] and previous analyses, the weight function $\psi(x) = \exp\left(-\frac{x^2}{2\alpha^2}\right)$ is adopted to construct the RKHS, requiring that $\alpha^2 > 2$.

Theorem 3. *If the condition*

$$\max_{j \in 1:d} \lambda_j < \frac{\exp(-\beta)}{\alpha^2 \exp(\max_{j \in 1:d} \omega_{*j})} \quad (66)$$

is satisfied, then the integrand obtained using LapIS in the GLMM belongs to an RKHS, as characterized by the weight function $\psi(x) = \exp\left(-\frac{x^2}{2\alpha^2}\right)$. Under these conditions, there exists a randomly shifted lattice rule that achieves an error bound of $O(N^{-1+1/\alpha^2+\delta})$ for this model.

Proof. The function $g(\omega)$ is defined as

$$\begin{aligned} g(\omega) &= \log \left(\prod_{j=1}^d \frac{\exp(y_j(\omega_j + \beta)) - \exp(\omega_j + \beta)}{y_j} \right) \\ &= \sum_{j=1}^d [y_j(\omega_j + \beta) - \exp(\omega_j + \beta) - \log y_j]. \end{aligned} \quad (67)$$

Because g is upper bounded, it satisfies Assumption 2. Differentiation of g with respect to each component yields

$$\frac{\partial g(\omega)}{\partial \omega_k} = y_k - \exp(\omega_k + \beta), \quad (68)$$

and

$$\frac{\partial^2 g(\omega)}{\partial \omega_j \partial \omega_k} = -\exp(\omega_k + \beta) \delta_{jk}. \quad (69)$$

Therefore, $\nabla^2 g$ is a negative definite diagonal matrix. According to Theorem 2, it is sufficient to verify whether Assumption 3 is valid. Clearly, this is equivalent to Equation (66). $\square$

Remark 3. *First, it is necessary to have $\alpha^2 > 2$ to ensure that the RKHS is well-defined. Additionally, according to Equation (66), the following is a necessary condition:*

$$2 < \alpha^2 < \frac{\exp(-\beta)}{\max_{j \in 1:d} \lambda_j \exp(\max_{j \in 1:d} \omega_{*j})}, \quad (70)$$

or, equivalently,

$$2e^\beta \max_{j \in 1:d} \lambda_j \exp(\max_{j \in 1:d} \omega_{*j}) < 1. \quad (71)$$

This condition may not always be valid, especially when d is large. This is because LapIS can result in an unbounded boundary, potentially diverging faster than $\exp(x^T x / 2\alpha^2)$, which is unfavorable for using QMC points to estimate.

4.1.2. Binary Regression Model

We define these functions such that their integral corresponds to the marginal likelihood integral $\int \pi_0(\beta) L(y|\beta) d\beta$ for a Bayesian statistical model. Here, $\beta \in \mathbb{R}^s$, $\pi_0(\beta)$ denotes a Gaussian prior density with mean zero and covariance $5\mathbf{I}_s$. The likelihood function $L(y|\beta)$ is given by the logistic regression model $L(y|\beta) = \prod_{i=1}^n F(y_i \beta^T x_i)$, where $F(z) = \frac{1}{1+e^{-z}}$, and the data $(x_i, y_i)_{i=1}^n$ comprise predictors $x_i \in \mathbb{R}^s$ and labels $y_i \in \{-1, 1\}$. Here, we adapt the importance sampling approach with MC and QMC to approximate the quantities above as in [19]. We consider the Pima dataset, which has 10 predictors. For $s = 8$, we take the first eight predictors (and same for $s = 4$). Here, we also fix the number of datapoints as 100. And, in this experiment, we use the product weight, which is generated by the CBC algorithm. Then, we calculate the relative RMSE of the four estimators: classic MC (NONE-MC), LapIS-QMC, ODIS-QMC and classic QMC (NONE-QMC). The standard error results are shown in Tables 2 and 3.

Table 2. Standard errors of IS estimators for a 4-dimension binary regression problem.

<i>N</i>	MC	NONE-QMC	LapIS-QMC	ODIS-QMC
512	2.17×10^{-1}	6.86×10^{-2}	1.66×10^{-1}	2.59×10^{-2}
1024	1.40×10^{-1}	2.91×10^{-2}	4.01×10^{-1}	2.41×10^{-2}
2048	1.03×10^{-1}	1.56×10^{-2}	8.39×10^{-1}	6.46×10^{-3}
4096	6.10×10^{-2}	9.86×10^{-3}	2.98×10^{-1}	4.18×10^{-3}
8192	5.39×10^{-2}	4.68×10^{-3}	1.01×10^{-1}	2.29×10^{-3}
16,384	3.39×10^{-2}	2.73×10^{-3}	2.73×10^{-3}	7.13×10^{-4}

Table 3. Standard errors of IS estimators for an 8-dimension binary regression problem.

<i>N</i>	MC	LapIS-QMC	NONE-QMC	ODIS-QMC
512	6.18×10^{-1}	1.81×10^{-1}	4.67×10^{-1}	2.84×10^{-1}
1024	4.37×10^{-1}	2.05×10^{-1}	3.15×10^{-1}	1.94×10^{-1}
2048	3.60×10^{-1}	1.59×10^{-1}	1.93×10^{-1}	9.16×10^{-2}
4096	1.80×10^{-1}	4.00×10^{-1}	1.08×10^{-1}	6.63×10^{-2}
8192	1.54×10^{-1}	1.66×10^{-1}	9.86×10^{-2}	4.74×10^{-2}
16,384	9.54×10^{-2}	1.56×10^{-1}	7.32×10^{-2}	2.57×10^{-2}

Based on the experiments conducted, it is observed that the application of importance sampling in QMC methods does exhibit improvement over the traditional Monte Carlo (MC) methods. However, it was noted that, in this particular problem, the performance of LapIS combined with QMC does not surpass that of optimal ODIS. This may be attributed to the fact that, for this posterior density estimation problem, an increase in sample size could result in the integrand falling outside the RKHS. This aspect is an issue that will be further investigated in our subsequent research.

4.2. Financial Models

4.2.1. Randleman–Bartter Model

We consider the valuation problem of a zero-coupon bond within the framework of the Randleman–Bartter model, as discussed in [20]. A three-dimensional Gaussian integral related to this model was previously analyzed by Caflisch [1]. This general reformulation aims to determine the fair price of a $(d + 1)$ -year zero-coupon bond with a face value of \$1. The valuation is represented as an integral:

$$P = \int_{\mathbb{R}^d} \prod_{k=0}^d \frac{p(z; 0, I_d)}{1 + r_k} dz. \quad (72)$$

The model treats the interest rates as

$$r_k = r_0 \exp(-k\sigma^2/2 + \sigma B_k), \quad k = 1, 2, \dots, d, \quad (73)$$

where the vector $(B_1, B_2, \dots, B_d)^T$ is distributed normally with a zero mean and a covariance matrix C , with entries $C_{ij} = \min(i, j)$. The parameters r_0 and σ indicate the current annual interest rate and volatility, respectively.

The generation of Brownian motion involves finding a matrix A that satisfies $AA^T = C$. The standard construction method for this system is used in our analysis in addition to other widely used methods like the Brownian bridge and principal component analysis, as detailed in [20]. Future studies may include these alternative generation methods.

The standard construction of this system is defined as

$$A := (a_{i,j})_{d \times d} = \begin{pmatrix} 1 & 0 & 0 & \cdots & 0 \\ 1 & 1 & 0 & \cdots & 0 \\ 1 & 1 & 1 & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & 1 & 1 & \cdots & 1 \end{pmatrix}, \quad (74)$$

and Brownian motion is generated as

$$(B_1, B_2, \dots, B_d)^T = A(z_1, z_2, \dots, z_d)^T, \quad (z_1, z_2, \dots, z_d)^T \sim N(0, I_d). \quad (75)$$

Finally, the integral in Equation (72) is rewritten in the standard form of Equation (6), with

$$G(z) = \prod_{k=0}^d (1 + r_k)^{-1}, \quad \mu = 0, \quad \Sigma = I_d. \quad (76)$$

Note that $a_{i,j} = 1$ if and only if $i \geq j$; otherwise, $a_{i,j} = 0$. Therefore,

$$\frac{\partial r_k}{\partial z_i} = \sigma r_k 1\{k \geq i\}. \quad (77)$$

$$\frac{\partial^2 r_k}{\partial z_i \partial z_j} = \sigma^2 r_k 1\{k \geq \max(i, j)\}. \quad (78)$$

We have

$$g(z) = \log G(z) = - \sum_{k=0}^d \log(1 + r_k), \quad H(z) = g(z) - \frac{1}{2} z^T z. \quad (79)$$

Then, the matrix $\nabla^2 H$ is given by $\nabla^2 H = \nabla^2 g - I_d$. Let μ_* be the solution to $\nabla H(z) = 0$ and let $\Sigma_* = (-\nabla^2 H(\mu_*))^{-1}$. The derivatives of g with respect to z_i are

$$\frac{\partial g}{\partial z_i} = - \sum_{k=1}^d \frac{1}{1 + r_k} \frac{\partial r_k}{\partial z_i} = - \sum_{k=i}^d \frac{\sigma r_k}{1 + r_k}. \quad (80)$$

The second derivatives of g are derived as

$$\begin{aligned} \frac{\partial^2 g}{\partial z_i \partial z_j} &= - \sum_{k=i}^d \frac{\sigma [\frac{\partial r_k}{\partial z_j} (1 + r_k) - \frac{\partial r_k}{\partial z_j} r_k]}{(1 + r_k)^2} \\ &= - \sum_{k=i}^d \frac{\sigma \frac{\partial r_k}{\partial z_j}}{(1 + r_k)^2} \\ &= - \sum_{k=i}^d \frac{\sigma^2 r_k 1\{k \geq j\}}{(1 + r_k)^2} \\ &= - \sum_{k=\max(i,j)}^d \frac{\sigma^2 r_k}{(1 + r_k)^2}. \end{aligned} \quad (81)$$

Defining $R_k = \frac{\sigma^2 r_k}{(1+r_k)^2}$ for simplicity, $\nabla^2 g$ is given by

$$\nabla^2 g = \begin{pmatrix} -\sum_{k=1}^d R_k & -\sum_{k=2}^d R_k & -\sum_{k=3}^d R_k & \cdots & -R_d \\ -\sum_{k=2}^d R_k & -\sum_{k=2}^d R_k & -\sum_{k=3}^d R_k & \cdots & -R_d \\ -\sum_{k=3}^d R_k & -\sum_{k=3}^d R_k & -\sum_{k=3}^d R_k & \cdots & -R_d \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ -R_d & -R_d & -R_d & \cdots & -R_d \end{pmatrix}. \quad (82)$$

This matrix is negative definite. A suitable choice for the decomposition matrix H is an upper triangular matrix defined as

$$H = \begin{pmatrix} \sqrt{R_1} & \sqrt{R_2} & \sqrt{R_3} & \cdots & \sqrt{R_d} \\ 0 & \sqrt{R_2} & \sqrt{R_3} & \cdots & \sqrt{R_d} \\ 0 & 0 & \sqrt{R_3} & \cdots & \sqrt{R_d} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & \sqrt{R_d} \end{pmatrix}. \quad (83)$$

Therefore, $-\nabla^2 g = HH^T$. Let $r_* = \min_{k \in 1:d} r_k(\mu_*)$. The weight function $\psi(x) = \exp\left(-\frac{x^2}{2\alpha^2}\right)$ is used to establish the RKHS, with the condition that $\alpha^2 > 2$.

Theorem 4. In the Randleman–Bartter model, if the condition $\alpha^2 < \frac{1}{\kappa} \sqrt{\frac{6}{d^3+2d^2+2d+1}}$ is satisfied, where $\kappa = \frac{\sigma^2}{2+r_*+1/r_*}$, then using randomly shifted lattice rule obtained by the CBC algorithm attains an error bound of $O(N^{-1+1/\alpha^2+\delta})$.

Proof. Assumption 2 is clearly satisfied. Because $\nabla^2 g$ is negative definite, it is sufficient to verify the validity of Assumption 3.

Here, the challenge lies in theoretically computing the maximal eigenvalue of Σ_* and the minimal eigenvalue of $\nabla^2 g(\mu_*)$. Instead of using direct computation, an estimation approach is used that leverages inequalities.

Considering that $-\nabla^2 g$ is a matrix, a classic linear algebra conclusion may be applied via

$$\sum_{k=1}^d v_k^2 = \|-\nabla^2 g\|_F^2, \quad (84)$$

where v_k denotes all eigenvalues of $-\nabla^2 g$ and $\|\cdot\|_F$ denotes the Frobenius norm. Estimating the Frobenius norm with $R_k = \frac{\sigma^2 r_k}{(1+r_k)^2} \geq \kappa$ yields

$$\begin{aligned} \|-\nabla^2 g\|_F^2 &= \sum_{j=1}^d (2j-1) \left(\sum_{k=j}^d R_k\right)^2 \\ &\geq \sum_{j=1}^d (2j-1)(d-j+1)^2 \kappa^2 \\ &= d \frac{(d^3+2d^2+2d+1)\kappa^2}{6} := S_d. \end{aligned} \quad (85)$$

Consequently,

$$\max_{k \in 1:d} v_k \geq \sqrt{S_d/d}. \quad (86)$$

and

$$h = -\max_{k \in 1:d} v_k \leq -\sqrt{\frac{(d^3+2d^2+2d+1)}{6}} \kappa. \quad (87)$$

For $\max_{k \in 1:d} \lambda_k$, a trivial upper bound is used. The proof is completed by applying Theorem 2. $\square$

In Theorem 4, concerns may arise when considering the upper bound of α , as it tends to decrease when the dimension d increases, potentially rendering the condition $\alpha > 2$ impractical in higher dimensions. However, this limitation does not significantly limit the use of the model. First, the estimation of $\max_{k \in \{1, \dots, d\}} \lambda_k$ can be further refined. Second, with $\sigma = 0.1$ and $\kappa \leq \frac{\sigma^2}{4} = 0.025$, a low dimension is both feasible and adequate for most practical scenarios.

In Figure 1, we present the numerical results for dimensions $d = 5$. The parameters employed in this system are $r_0 = 0.1$ and $\sigma = 0.01$. Here, we set the product weights as $\gamma_j = \frac{1}{j^\eta}$, $\eta = 3$. Therefore, these weights satisfy Equation (52) (refer to [21]).

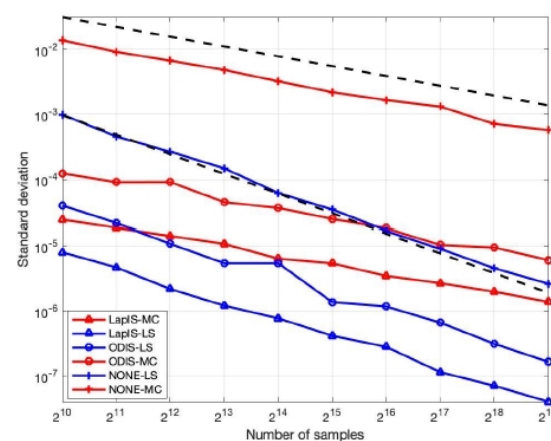


Figure 1. RMSE for the Randleman–Bartter short-term interest model with $d = 5$.

In the RQMC setting, both ODIS and LapIS demonstrate more rapid convergence compared to their performance in the MC context. And we can also find this if we let dimension d grow and set κ not too small. Then, we can obtain a best convergence rate. In the result of the experiment, we can also find that both ODIS- and LapIS-enhanced convergence manifest an error rate approaching $O(N^{-1})$, which agree with our theoretical findings. As the dimensions increase, these findings remain consistent.

Our analysis reveals that employing LapIS yields superior outcomes compared to ODIS, which may be attributed to the condition that G_{IS} belongs to the RKHS when both IS methods are implemented. Consequently, for the integrand belonging to the RKHS, LapIS performs better than ODIS.

4.2.2. Asian Options

The conditions previously discussed are satisfied by the integrands in the above two examples. We now consider a classic financial example, referred to as the Asian option.

Under the Black–Scholes model, the asset price $S(t)$ follows geometric Brownian motion under the risk-neutral measure, given by

$$dS(t) = rS(t)dt + \sigma S(t)dB(t), \quad (88)$$

where r is the risk-free rate, σ is the volatility, and $B(t)$ represents a standard Brownian motion. The solution to this stochastic differential equation is

$$S(t) = S_0 \exp\left(\left(r - \frac{\sigma^2}{2}\right)t + \sigma B(t)\right), \quad (89)$$

with $S_0 = S(0)$ denoting the initial asset price at time zero. The discounted payoff of an Asian call option is

$$e^{-rT} \max(S_A - K, 0), \quad (90)$$

where T is the maturity, K is the strike price, and $S_A = \frac{1}{d} \sum_{k=1}^d S(t_j)$ is the arithmetic average of the asset prices at discrete times $t_j = jT/d$ for $j = 1, 2, \dots, d$. To conduct numerical simulations, we first generate the asset prices at these times. For notational convenience, let $S_j = S(t_j)$ and $\Delta t = T/d$. The asset price vector is then

$$S := (S_1, S_2, \dots, S_d)^T = S_0 \exp \left(\left(r - \frac{\sigma^2}{2} \right) \Delta t (1, 2, \dots, d)^T + \sigma \sqrt{\Delta t} A z \right), \quad (91)$$

where $z = (z_1, z_2, \dots, z_d)^T \sim N(0, I_d)$ and A is the same matrix as in (74), which represents the Cholesky decomposition of the standard Brownian motion covariance matrix $C = (c_{i,j})_{d \times d} = (\min(i, j))_{d \times d}$.

The fair pricing of an Asian option involves integrating Equation (90). Moreover, the integrand is not differentiable, and, therefore, our previous results do not apply. However, we can still observe the effects of importance sampling through numerical experiments. The numerical results for this example are presented in Tables 4 and 5. We use the CBC algorithm with two sets of weights: $\gamma_j = 1/j^3$ and constant weights $\gamma_u \equiv 1$. The parameters are set as $S_0 = 100$, $r = 0.1$, $\sigma = 0.2$, $K = 100$, $T = 1$, dimension $d = 100$, normal weight function $\alpha^2 = 10$, and weight coefficient $\kappa = 0.75$. For clarity, we list the results computed by the MC method as a baseline, followed by those obtained using QMC without IS (NONE), with ODIS, and with LapIS. Errors are reported to two significant figures, and variance reduction factors are rounded to the nearest integers.

Table 4. Standard errors of IS estimators for Asian options with constant weight coefficients.

N	MC	NONE-QMC	ODIS-QMC	LapIS-QMC
1024	9.01×10^{-3}	6.74×10^{-3} (2)	7.07×10^{-4} (162)	3.44×10^{-4} (686)
2048	5.97×10^{-3}	3.05×10^{-3} (4)	3.21×10^{-4} (346)	1.66×10^{-4} (1293)
4096	4.11×10^{-3}	1.25×10^{-3} (11)	1.19×10^{-4} (1193)	7.97×10^{-5} (2659)
8192	2.77×10^{-3}	6.89×10^{-4} (16)	5.53×10^{-4} (2509)	4.00×10^{-5} (4796)
16,384	1.98×10^{-3}	4.79×10^{-4} (17)	2.77×10^{-4} (5109)	2.10×10^{-5} (8890)
32,768	1.32×10^{-3}	2.33×10^{-4} (32)	1.44×10^{-4} (8403)	1.01×10^{-5} (17,081)
65,536	9.74×10^{-4}	1.17×10^{-4} (69)	7.53×10^{-5} (16,731)	4.72×10^{-6} (42,583)

Table 5. Standard errors of IS estimators for Asian options with decreasing weight coefficients.

N	MC	NONE-QMC	ODIS-QMC	LapIS-QMC
1024	9.64×10^{-3}	6.19×10^{-3} (2)	7.32×10^{-4} (173)	3.15×10^{-4} (937)
2048	7.46×10^{-3}	3.03×10^{-3} (6)	3.66×10^{-4} (415)	1.59×10^{-4} (2201)
4096	4.57×10^{-3}	1.33×10^{-3} (12)	1.75×10^{-4} (682)	6.11×10^{-5} (5594)
8192	3.23×10^{-3}	7.99×10^{-4} (16)	8.08×10^{-5} (1596)	2.74×10^{-5} (13,896)
16,384	2.09×10^{-3}	5.25×10^{-4} (16)	4.20×10^{-5} (2476)	1.25×10^{-5} (27,956)
32,768	1.40×10^{-3}	2.56×10^{-4} (30)	2.99×10^{-5} (2192)	5.67×10^{-6} (60,966)
65,536	1.01×10^{-3}	8.55×10^{-5} (140)	1.07×10^{-5} (8909)	3.03×10^{-6} (111,111)

We briefly summarize the performance of the estimator in this case as follows:

1. Using a randomly shifted lattice rule with importance sampling, the error convergence order becomes close to $O(N^{-1})$. The fitting result obtained in the example case is approximately $O(N^{-0.95})$;
2. Importance sampling can be used to reduce calculation errors. Using ODIS can improve the variance reduction factor by hundreds or thousands of times compared to MC, while using LapIS can improve it by thousands of times. In this case, we found

that LapIS still has great potential in financial problems, although it is difficult to study whether this integrand falls in RKHS;

3. Upon changing the weights to constant values, the overall results do not greatly differ. It is worth mentioning that the advantage of LapIS over ODIS is somewhat reduced under this condition. We leave the problem of choosing the weight to future research.

4.3. Estimator Study

In this subsection, we present a comprehensive analysis of the performance of LapIS–QMC and ODIS–QMC across three distinct numerical examples: binary regression, the Randleman–Bartter Model, and Asian options. The effectiveness of these estimators is assessed in terms of error reduction and their suitability for the problems at hand.

In the case of binary regression, we found that our experiments in binary regression demonstrated that the application of importance sampling within QMC methods notably enhances performance compared to traditional Monte Carlo methods. However, it was observed that, in this specific context, the efficacy of LapIS combined with QMC did not exceed that of ODIS. This may be ascribed to the challenges posed by the posterior density estimation problem, where an increased sample size could lead the integrand to fall outside the RKHS. This observation warrants further investigation, which will be pursued in future research.

In examining the Randleman–Bartter Model, it was revealed that employing LapIS yields significantly better outcomes than ODIS. This improvement is likely due to the conditions that the integrand resides within the RKHS and that the growth of the integrand is now larger than second-order exponential function when both importance sampling methods are applied. Therefore, for integrands belonging to the RKHS, LapIS exhibits superior performance over ODIS.

In the context of Asian options, although the integrand is not continuous, we can still find that use of importance sampling has great potential. Importance sampling, particularly using ODIS, significantly enhances variance reduction, potentially by hundreds to thousands of times compared to traditional MC methods. LapIS, in this scenario, showed even greater potential for variance reduction, improving it by thousands of times.

We can also deduce that, if the function belongs to RKHS and the growth of integrand is slower the second-order exponential function, then using LapIS has greater advantage than ODIS. Otherwise, using ODIS is a better choice. Additionally, combining IS with QMC is usually better than the combination of IS with MC.

5. Conclusions

This study investigated the effectiveness of combining randomly shifted lattice rules with importance sampling techniques, specifically focusing on their impact on the convergence rates for integrals. We prove that, for unbounded functions, randomly shifted lattice rules combined with suitably chosen importance densities can achieve a convergence as fast as $O(N^{-1+\epsilon})$ given N samples for arbitrary $\epsilon > 0$ values under certain conditions. We investigated the application of the estimators in statistical and financial models, considering the conditions under which the integrands can meet the aforementioned criteria (in RKHS). Furthermore, we identified the potential benefits of using importance sampling with the lattice rule in option pricing problems. We conclude that both ODIS and LapIS have their own applicable scenarios. Pre-integration methods may be required to verify the theoretical convergence conditions of these systems [22]. Investigating convergence conditions for integrand estimators using pre-integration techniques combined with importance sampling will be a subject in our future research.

Author Contributions: Conceptualization, H.W. and Z.Z.; methodology, H.W.; software, H.W.; validation, H.W.; writing—original draft preparation, Z.Z.; writing—review and editing, H.W.; visualization, H.W.; supervision, H.W.; project administration, H.W.; funding acquisition, H.W. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by NATIONAL SCIENCE FOUNDATION OF CHINA grant number No. 72071119.

Data Availability Statement: The data used in this research are available from the corresponding author upon reasonable request.

Acknowledgments: We would like to express gratitude to Professor Xiaoqun Wang for his guidance and suggestions for revisions on this thesis.

Conflicts of Interest: The authors declare no conflicts of interest.

Appendix A

In this appendix, we deduce the simple form of $\hat{\theta}(h)$. First,

$$\begin{aligned}\hat{\theta}(h) &= \int_0^1 \int_{\Phi^{-1}(x)}^0 \frac{\Phi(t) - x}{\psi^2(t)} \exp(-2\pi i h x) dt dx + \int_0^1 \int_{\Phi^{-1}(1-x)}^0 \frac{\Phi(t) - 1 + x}{\psi^2(t)} \exp(-2\pi i h x) dt dx \\ &= \int_0^1 \int_{\Phi^{-1}(x)}^0 \frac{\Phi(t) - x}{\psi^2(t)} \exp(-2\pi i h x) dt dx + \int_0^1 \int_{\Phi^{-1}(x)}^0 \frac{\Phi(t) - x}{\psi^2(t)} \exp(2\pi i h x) dt dx \\ &= 2 \int_0^1 \int_{\Phi^{-1}(x)}^0 \frac{\Phi(t) - x}{\psi^2(t)} \cos(2\pi h x) dt dx.\end{aligned}$$

The outer integral at $\Phi(0)$ may be split using the Fubini theorem via

$$\begin{aligned}\hat{\theta}(h) &= 2 \int_0^{\Phi(0)} \int_{\Phi^{-1}(x)}^0 \frac{\Phi(t) - x}{\psi^2(t)} \cos(2\pi h x) dt dx + 2 \int_{\Phi(0)}^1 \int_{\Phi^{-1}(x)}^0 \frac{\Phi(t) - x}{\psi^2(t)} \cos(2\pi h x) dt dx \\ &= 2 \int_{-\infty}^0 \frac{1}{\psi^2(t)} \int_0^{\Phi(t)} (\Phi(t) - x) \cos(2\pi h x) dx dt + 2 \int_0^{\infty} \frac{1}{\psi^2(t)} \int_{\Phi(t)}^1 (x - \Phi(t)) \cos(2\pi h x) dx dt \\ &= 2 \int_{-\infty}^0 \frac{1}{\psi^2(t)} \frac{\sin^2(\pi h \Phi(t))}{2\pi^2 h^2} + 2 \int_0^{\infty} \frac{1}{\psi^2(t)} \frac{\sin^2(\pi h \Phi(t))}{2\pi^2 h^2} dt \\ &= \frac{1}{\pi^2 h^2} \int_{-\infty}^{\infty} \frac{\sin^2(\pi h \Phi(t))}{\psi^2(t)} dt \\ &= \frac{1}{\pi^2 h^2} \int_0^1 \frac{\sin^2(\pi h t)}{\psi^2(\Phi^{-1}(t)) \phi(\Phi^{-1}(t))} dt.\end{aligned}$$

References

1. Caflisch, R.E. Monte Carlo and quasi-Monte Carlo methods. *Acta Numer.* **1998**, *7*, 1–49. [CrossRef]
2. Niederreiter, H. *Random Number Generation and Quasi-Monte Carlo Methods*; SIAM: Philadelphia, PA, USA, 1992.
3. Owen, A.B. Scrambled net variance for integrals of smooth functions. *Ann. Stat.* **1997**, *25*, 1541–1562. [CrossRef]
4. Owen, A.B. Monte Carlo variance of scrambled net quadrature. *SIAM J. Numer. Anal.* **1997**, *34*, 1884–1910. [CrossRef]
5. Dick, J.; Pillichshammer, F. *Digital Nets and Sequences: Discrepancy Theory and Quasi-Monte Carlo Integration*; Cambridge University Press: Cambridge, UK, 2010.
6. L'Ecuyer, P.; Lemieux, C. Recent Advances in Randomized Quasi-Monte Carlo Methods. In *Modeling Uncertainty: An Examination of Stochastic Theory, Methods, and Applications*; Kluwer Academic: Boston, MA, USA, 2002; pp. 419–474.
7. He, Z.; Zheng, Z.; Wang, X. On the Error Rate of Importance Sampling with Randomized Quasi-Monte Carlo. *SIAM J. Numer. Anal.* **2023**, *61*, 515–538. [CrossRef]
8. Glasserman, P.; Heidelberger, P.; Shahabuddin, P. Asymptotically optimal importance sampling and stratification for pricing path-dependent options. *Math. Financ.* **1999**, *9*, 117–152. [CrossRef]
9. Rubino, G.; Tuffin, B. *Rare Event Simulation Using Monte Carlo Methods*; John Wiley & Sons: Hoboken, NJ, USA, 2009.
10. Owen, A.B. Monte Carlo Theory, Methods and Examples. Available online: <https://artowen.su.domains/mc/> (accessed on 26 January 2013).
11. Zhang, C.; Wang, X.; He, Z. Efficient importance sampling in quasi-Monte Carlo methods for computational finance. *SIAM J. Sci. Comput.* **2021**, *43*, B1–B29. [CrossRef]
12. Kuo, F.Y.; Dunsmuir, W.T.M.; Sloan, I.H.; Wand, M.P.; Womersley, R.S. Quasi-Monte Carlo for highly structured generalised response models. *Methodol. Comput. Appl. Probab.* **2008**, *10*, 239–275. [CrossRef]
13. Dick, J.; Rudolf, D.; Zhu, H. A weighted discrepancy bound of quasi-Monte Carlo importance sampling. *Stat. Probab. Lett.* **2019**, *149*, 100–106. [CrossRef]

14. Kuo, F.Y.; Wasilkowski, G.W.; Waterhouse, B.J. Randomly shifted lattice rules for unbounded integrands. *J. Complex.* **2006**, *22*, 630–651. [[CrossRef](#)]
15. Nichols, J.A.; Kuo, F.Y. Fast CBC construction of randomly shifted lattice rules achieving $O(N^{-1+\Delta})$ convergence for unbounded integrands over $\mathbb{R}^d$ in weighted spaces with POD weights. *J. Complex.* **2014**, *30*, 444–468. [[CrossRef](#)]
16. Waterhouse, B.J.; Kuo, F.Y.; Sloan, I.H. Randomly shifted lattice rules on the unit cube for unbounded integrands in high dimensions. *J. Complex.* **2006**, *22*, 71–101. [[CrossRef](#)]
17. Sloan, I.H.; Kuo, F.Y.; Joe, S. Constructing randomly shifted lattice rules in weighted Sobolev spaces. *SIAM J. Numer. Anal.* **2002**, *40*, 1650–1665. [[CrossRef](#)]
18. Kuo, F.Y.; Sloan, I.H.; Wasilkowski, G.W.; Waterhouse, B.J. Randomly shifted lattice rules with the optimal rate of convergence for unbounded integrands. *J. Complex.* **2010**, *26*, 135–160. [[CrossRef](#)]
19. Chopin, N.; Ridgway, J. Leave Pima Indians alone: Binary regression as a benchmark for Bayesian computation. *Statist. Sci.* **2017**, *32*, 64–87. [[CrossRef](#)]
20. Hull, J. *Options, Futures and Other Derivatives*; Prentice-Hall: Upper Saddle River, NJ, USA, 2005.
21. Graham, I.G.; Kuo, F.Y.; Nichols, J.A.; Scheichl, R.; Schwab, C.; Sloan, I.H. Quasi-Monte Carlo finite element methods for elliptic PDEs with lognormal random coefficients. *Numer. Math.* **2015**, *131*, 329–368. [[CrossRef](#)]
22. Xiao, Y.; Wang, X. Conditional quasi-Monte Carlo methods and dimension reduction for option pricing and hedging with discontinuous functions. *J. Comput. Appl. Math.* **2018**, *343*, 289–308. [[CrossRef](#)]

Disclaimer/Publisher’s Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.