



HAL
open science

Sufficient Conditions for Central Limit Theorems and Confidence Intervals for Randomized Quasi-Monte Carlo Methods

Marvin K Nakayama, Bruno Tuffin

► **To cite this version:**

Marvin K Nakayama, Bruno Tuffin. Sufficient Conditions for Central Limit Theorems and Confidence Intervals for Randomized Quasi-Monte Carlo Methods. 2021. hal-03196085v1

HAL Id: hal-03196085

<https://inria.hal.science/hal-03196085v1>

Preprint submitted on 12 Apr 2021 (v1), last revised 18 Jul 2021 (v2)

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Sufficient Conditions for Central Limit Theorems and Confidence Intervals for Randomized Quasi-Monte Carlo Methods

Marvin K. Nakayama and Bruno Tuffin

Abstract

Randomized quasi-Monte Carlo methods have been introduced with the main purpose of yielding a computable measure of error for quasi-Monte Carlo approximations through the implicit application of a central limit theorem over independent randomizations. But to increase precision for a given computational budget, the number of independent randomizations is usually set to a small value so that a large number of points are used from each randomized low-discrepancy sequence to benefit from the fast convergence rate of quasi-Monte Carlo. While a central limit theorem has been previously established for a specific but impractical type of randomization, it is also known in general that fixing the number of randomizations and increasing the length of the sequence used for quasi-Monte Carlo can lead to a non-Gaussian limiting distribution. This paper presents sufficient conditions on the relative growth rates of the number of randomizations and the quasi-Monte Carlo sequence length to ensure a central limit theorem and also an asymptotically valid confidence interval. We obtain several results based on the Lindeberg condition and expressed in terms of the regularity of the integrand and the convergence speed of the quasi-Monte Carlo method. We also analyze the resulting estimator's convergence rate.

1 Introduction

Researchers and analysts across diverse fields of science, engineering, business, etc., build stochastic models to study complicated problems, such as molecular dynamics, queueing systems, or pricing of financial instruments. Typically, the goal is to compute the model's mean, which can often be expressed as a (multidimensional) integral, even if not initially expressed in this form. But model complexity usually precludes analytically evaluating the integral, so numerical methods are employed. Monte Carlo (MC) methods are computational algorithms based on random sampling that can be applied to estimate a mean (integral); see, e.g., [1, 14] among the vast literature on the topic. Along with its many other desirable features, MC methods can provide a measure of an estimate's precision through a confidence interval (CI) obtained from an associated central limit theorem (CLT).

Quasi-Monte Carlo (QMC) methods [22, 26] replace the random sampling of MC by a sequence of points *deterministically* chosen to be “evenly dispersed” over the space. The points form a so-called *low-discrepancy sequence*, and thanks to the more rapid coverage of the domain, the resulting approximation is expected to converge faster than MC to the integral's exact value, at least for regular integrands. This improved convergence speed can be shown through several existing error bounds [12], the most famous being the Koksma-Hlawka bound [26]. A drawback of QMC stems from the fact that such error bounds, while theoretically valuable in proving asymptotic properties, are in most cases useless in practice, being not easy to compute and grossly overestimating the error for a fixed sample size; see, e.g., [37, Section 2.2] for a discussion of the issues.

Randomized quasi-Monte Carlo (RQMC) methods have been introduced with the main purpose to solve this problem of assessing error in QMC methods. The principle is to randomly “shake” the low-discrepancy sequence, but without losing its good repartition over the sampling space. From r independently randomized QMC estimators, we then can obtain an approximate CI via a CLT, exploiting the assumed convergence to a normal distribution as $r \rightarrow \infty$. For tutorials on RQMC, see among others [16, 18, 22, 37]. Several types of randomization exist, the main ones being the *random shift* [7], which translates all points of the low-discrepancy sequence by the *same random vector*; the *digital shift* [16], which applies random (digital) shifts to digits of the points of the low-discrepancy sequence, the same digital shift being utilized on digits of the same order for all the points; and the *digital scrambling*, randomly permuting the digits [28, 29].

In implementing RQMC, we apply r independent randomizations of m points from the low-discrepancy sequence of QMC. Given a computation budget that allows for about n evaluations of the integrand, the user then needs to determine how to allocate n to (m, r) so that $mr \approx n$. Choosing m large and r correspondingly small benefits from the faster convergence speed (i.e., the improved precision) of QMC compared to MC. One might try applying a common rule of thumb that suggests a CLT roughly holds for fixed $r \approx 30$, and then lets m be as large as possible to get a smaller error. But this heuristic generally lacks a theoretical basis. For example, [19] establishes that for a single random shift of a lattice, the limiting error distribution as $m \rightarrow \infty$ typically follows a non-normal distribution; thus, for RQMC with any fixed $r \geq 1$ in this setting, the error’s limiting distribution will not be Gaussian, so the common suggestion of specifying a small r rests on shaky ground. One CLT, though, has been verified [23] (as $m \rightarrow \infty$ for fixed $r \geq 1$) in the case of digital scrambling of a particular type called *nested uniform scrambling*. But this scrambling is computationally prohibitive, which has limited its adoption in practice.

The question we therefore aim at answering here is the following: for RQMC, how should (m, r) comparatively increase to ensure a CLT and yield an *asymptotically valid CI* (AVCI)? To analyze this issue, we formulate the problem using a so-called *triangular array*, for which a CLT can be obtained under the Lindeberg condition [2, Theorem 27.2]. More precisely, we provide sufficient conditions on the relative increase of m and r under three types of assumptions on properties of the integrand. We initially establish CLTs in terms of the true variance from a single randomization, but this quantity can hardly be known in general. Thus, we also derive similar conditions for convergence to a normal distribution when the variance is estimated, which secures the AVCI property. Comparing our conditions on the integrand, we show that weaker assumptions correspond to stronger requirements on the number of randomizations (i.e., higher relative proportion for r). Similarly, the faster the convergence of the underlying QMC method, the more weight should be placed on randomizations. But in all cases when $m \rightarrow \infty$, we further verify that under our assumptions, RQMC always outperforms MC in terms of convergence speed of the *root-mean-square error* (RMSE) of their estimators.

The rest of the paper progresses as follows. Section 2 presents the notation used throughout the paper and reviews MC, QMC and RQMC methods. It also introduces the different assumptions on the relative increase of m and r and on the properties of the integrand, along with related useful lemmas. Section 3 applies the conditions of Lindeberg and Lyapounov [2, Theorem 27.3] for triangular-array CLTs in our RQMC framework. We further specialize the CLTs to obtain corollaries exploiting specific properties of the integrand, and it is shown that the stronger the condition on the integrand, the more weight can be put on the QMC component. Section 4 presents similar results when the variance is estimated, the “real-life” situation, to obtain conditions for AVCI. Section 5 expresses the results under the special context when $m \equiv n^c$ and $r \equiv n^{1-c}$ for $c \in (0, 1)$, helping to gain deeper insights on the conditions and their implications. Section 6 gives some numerical results illustrating the practical

validity and utility of the theoretical results. Section 7 provides concluding remarks. In [25], we present many of the theorems, corollaries and propositions, but all without proofs.

2 Notation/Framework

Our goal is to estimate

$$\mu = \mathbb{E}[h(U)] = \int_{[0,1]^s} h(u) \, du,$$

where $h : [0, 1]^s \rightarrow \mathbb{R}$ is a given function (integrand) for some fixed $s \geq 1$, random vector $U \sim \mathcal{U}[0, 1]^s$ with $\mathcal{U}[0, 1]^s$ the uniform distribution on the s -dimensional unit hypercube $[0, 1]^s$, \sim means “is distributed as”, and \mathbb{E} denotes the expectation operator. Integrating over $[0, 1]^s$ is the standard QMC setting. We can think of integrand h as being a complicated simulation program that converts s independent univariate uniform random numbers into observations from specified input distributions (possibly with dependencies and different marginals), which are used to produce an output of the stochastic model. We next describe methods for estimating μ via MC, QMC, and RQMC.

2.1 Monte Carlo

The MC estimator of μ is $\hat{\mu}_n^{\text{MC}} = \frac{1}{n} \sum_{i=1}^n h(U_i)$, with $U_i, i = 1, 2, \dots, n$, as *independent and identically distributed (i.i.d.)* $\mathcal{U}[0, 1]^s$ random vectors. Suppose that $0 < \psi^2 \equiv \text{Var}[h(U)] < \infty$, where $\text{Var}[\cdot]$ denotes the variance operator. The MC estimator’s estimator’s root-mean-square error then satisfies

$$\text{RMSE} [\hat{\mu}_n^{\text{MC}}] = \frac{\psi}{\sqrt{n}} \quad (1)$$

because $\hat{\mu}_n^{\text{MC}}$ is unbiased (i.e., $\mathbb{E}[\hat{\mu}_n^{\text{MC}}] = \mu$), as will be the case for all estimators of μ that we consider. Moreover the MC estimator $\hat{\mu}_n^{\text{MC}}$ obeys a (Gaussian) CLT (e.g., [2, Theorem 27.1]): $\sqrt{n}[\hat{\mu}_n^{\text{MC}} - \mu]/\psi \Rightarrow \mathcal{N}(0, 1)$ as $n \rightarrow \infty$, where \Rightarrow denotes weak convergence [2, Section 25], and $\mathcal{N}(a, b^2)$ is a normal random variable with mean a and variance b^2 . The CLT provides a simple way to construct a (probabilistic) measure of the estimator’s error through a confidence interval for μ . Specifically, we define an MC CI as $I_{n,\gamma}^{\text{MC}} \equiv [\hat{\mu}_n^{\text{MC}} \pm z_\gamma \hat{\psi}_n / \sqrt{n}]$, where $\hat{\psi}_n^2 = [1/(n-1)] \sum_{i=1}^n [h(U_i) - \hat{\mu}_n^{\text{MC}}]^2$ is an unbiased estimator of ψ^2 , $0 < \gamma < 1$ is the specified confidence level (e.g., $\gamma = 0.95$), z_γ satisfies $\Phi(z_\gamma) = 1 - (1 - \gamma)/2$, and Φ is the $\mathcal{N}(0, 1)$ *cumulative distribution function* (CDF). Then $I_{n,\gamma}^{\text{MC}}$ is an asymptotically valid CI for μ in the sense that $\lim_{n \rightarrow \infty} P(\mu \in I_{n,\gamma}^{\text{MC}}) = \gamma$ [4, Example 10.4.4].

2.2 Quasi-Monte Carlo

QMC replaces MC’s random i.i.d. uniforms $(U_i)_{i \geq 1}$ with a deterministic low-discrepancy sequence $\Xi = (\xi_i)_{i \geq 1}$ (e.g., a *digital net*, such as a Sobol’ sequence, or a *lattice* [22, Chapter 5]), leading QMC to approximate μ via $\hat{\mu}_n^{\text{Q}} = \frac{1}{n} \sum_{i=1}^n h(\xi_i)$ based on n evaluations of the integrand h . The QMC error is often expressed as controlled by the *Koksma-Hlawka bound* (e.g., [22, Section 5.6.1])

$$|\hat{\mu}_n^{\text{Q}} - \mu| \leq V_{\text{HK}}(h) D_n^*(\Xi) \quad (2)$$

for all n , where $V_{\text{HK}}(h)$ is the Hardy-Krause variation of the integrand h , and $D_n^*(\Xi)$ is the star-discrepancy of the first n terms of Ξ , which satisfy $V_{\text{HK}}(h) \geq 0$ and $0 \leq D_n^*(\Xi) \leq 1$. In (2), $V_{\text{HK}}(h)$

measures the “roughness” of h , while $D_n^*(\Xi)$ quantifies the “non-uniformity” of Ξ . We typically have that

$$D_n^*(\Xi) = O(n^{-1}(\ln n)^s), \quad \text{as } n \rightarrow \infty, \quad (3)$$

where $f(n) = O(g(n))$ as $n \rightarrow \infty$ for functions f and g means that there exist positive constants a_0 and n_0 such that $|f(n)| \leq a_0|g(n)|$ for all $n > n_0$. Thus, when $V_{\text{HK}}(h) < \infty$, putting (3) into (2) shows that $|\widehat{\mu}_n^{\text{Q}} - \mu| = O(n^{-1}(\ln n)^s)$, so the QMC error has an asymptotically faster convergence rate than $\widehat{\mu}_n^{\text{MC}}$ does (as measured, e.g., by the MC estimator’s RMSE in (1)). But we have no practical error bound for QMC because computing $V_{\text{HK}}(h)$ is at least as difficult as computing μ and is the sum of $2^s - 1$ Vitali variations [26], therefore potentially very large even for moderate dimensions. Similarly, (3) provides only an upper bound for the rate at which $D_n^*(\Xi)$ decreases and can be quite loose for moderate values of n . Other error bounds similar to (2) exist (see among others [11, 12, 15, 21, 26]), but each encounters similar computational issues.

2.3 Randomized Quasi-Monte Carlo

RQMC randomizes the low-discrepancy sequence, but without losing its good repartition property, and computes an estimator from the randomized point set. More precisely, let $(U'_i)_{i \geq 1}$ be a randomized low-discrepancy sequence, such that each U'_i is uniformly distributed over $[0, 1]^s$ but $(U'_i)_{i \geq 1}$ are correlated and preserve the low-discrepancy property. RQMC repeats this $r \geq 1$ independent times, computing an estimator from each randomization. Specifically, using the first m points from a low-discrepancy sequence Ξ , RQMC applies $r \geq 1$ independent randomizations (indexed by $j = 1, 2, \dots, r$) to the m points. Let $U'_{i,j}$ be the i -th point of the j -th randomization ($1 \leq i \leq m$ and $1 \leq j \leq r$). The RQMC estimator is then

$$\widehat{\mu}_{m,r}^{\text{RQ}} = \frac{1}{r} \sum_{j=1}^r X_j, \quad \text{where} \quad X_j = \frac{1}{m} \sum_{i=1}^m h(U'_{i,j}), \quad (4)$$

with X_j as the estimator from randomization $j = 1, 2, \dots, r$. The independence across the r randomizations ensure that X_j , $j = 1, 2, \dots, r$, are i.i.d. From their sample variance $\widehat{\sigma}_{m,r}^2 = \sum_{j=1}^r (X_j - \widehat{\mu}_{m,r}^{\text{RQ}})^2 / (r - 1)$ when $r \geq 2$, we then obtain a potential γ -level CI for μ as $I_{m,r,\gamma}^{\text{RQ}} \equiv [\widehat{\mu}_{m,r}^{\text{RQ}} \pm z_\gamma \widehat{\sigma}_{m,r}]$. The hope is that as m or r (or both) grows large, the overall RQMC estimator $\widehat{\mu}_{m,r}^{\text{RQ}}$ obeys a Gaussian CLT (see Section 3) and $I_{m,r,\gamma}^{\text{RQ}}$ is an AVCI (see Section 4). We next describe some possible randomizations and the associated existing theoretical results.

2.3.1 Random Shift

A *random shift* of a low-discrepancy sequence generates a single uniformly distributed point $U \sim \mathcal{U}[0, 1]^s$ and adds it to each point of Ξ , modulo 1, coordinate-wise [7]. Formally, using the first m points of the low-discrepancy sequence Ξ and $r \geq 1$ independent randomizations leads to

$$U'_{i,j} = \langle U_j + \xi_i \rangle \quad (5)$$

in (4), where across randomizations, U_1, U_2, \dots, U_r are i.i.d. $\mathcal{U}[0, 1]^s$ and with $\langle x \rangle$ the modulo-1 operator applied to each coordinate of $x \in \mathbb{R}^s$. It is simple to show that each $U'_{i,j} \sim \mathcal{U}[0, 1]^s$. Within each randomization j , we have that $\langle U_j + \xi_i \rangle$, $i = 1, 2, \dots, m$, are dependent as they share the same U_j . The RQMC estimator of μ from random shifts is then as in (4) with $U'_{i,j}$ from (5).

Each randomized sequence $\Xi_{U_j} \equiv (\langle U_j + \xi_i \rangle)_{i \geq 1}$ satisfies

$$D_m^*(\Xi_{U_j}) \leq 4^s D_m^*(\Xi), \quad (6)$$

as shown in [35, Theorem 2]. As a consequence, if h is a function with bounded Hardy-Krause variation, the standard deviation of each X_j in (4) from a point set of size m is

$$\text{RMSE}[X_j] = O(m^{-1}(\log m)^s), \quad (7)$$

faster than the $\Theta(m^{-1/2})$ rate in (1) for MC with the same number m of calls to function h , where the notation $f_1(m) = \Theta(f_2(m))$ as $m \rightarrow \infty$ for functions f_1 and f_2 means that there exist positive constants a_0, a_1 , and m_0 such that $a_0|f_2(m)| \leq |f_1(m)| \leq a_1|f_2(m)|$ for all $m > m_0$.

The convergence speed can even be faster for special classes of functions and specific sequences Ξ called *lattice rules* [36, 17] for which the random shift preserves the lattice structure. Let \mathbb{Z} denote the set of all integers, and for $g = (g_1, g_2, \dots, g_s) \in \mathbb{Z}^s$, define $t(g) = \prod_{i=1}^s \max(1, |g_i|)$. For a periodic function $f : \mathbb{R}^s \rightarrow \mathbb{R}$ with period 1 over each coordinate, define its Fourier coefficient of rank $g \in \mathbb{Z}^s$ as $\hat{f}(g) = \int_{[0,1]^s} f(x) e^{-2\pi\sqrt{-1}g \cdot x} dx$, where $v \cdot y = \sum_{i=1}^s v_i y_i$ is the inner product of $v = (v_1, \dots, v_s) \in \mathbb{R}^s$ and $y \in (y_1, \dots, y_s) \in \mathbb{R}^s$. For $\alpha > 1$ and $C > 0$, let $E_\alpha^s(C)$ be the set of such periodic functions $f : \mathbb{R}^s \rightarrow \mathbb{R}$ for which

$$|\hat{f}(g)| \leq Ct(g)^{-\alpha} \text{ for all } g \in \mathbb{Z}^s.$$

Then, for each $\alpha > 1$, $C > 0$ and $m \geq 1$, there exists a lattice rule $\Xi^{[m]} = (\xi_i^{[m]})_{i \geq 1}$ (depending on m , as well as α , C , and s) such that for $U \sim \mathcal{U}[0, 1]^s$,

$$\max_{f \in E_\alpha^s(C)} \text{RMSE} \left[\frac{1}{m} \sum_{i=1}^m f(\langle U + \xi_i^{[m]} \rangle) \right] = O(m^{-\alpha} (\ln m)^{\alpha s}).$$

Thus, the convergence speed has an asymptotic upper bound that is even better than what we get from the Koksma-Hlawka bound (2) and (3).

For a randomly-shifted low-discrepancy sequence, the RQMC estimator may not obey a Gaussian CLT for a fixed r as $m \rightarrow \infty$, as shown for lattices in [19], which further analyzes the error. In particular, for $r = 1$, the limiting error distribution in dimension $s = 1$ is uniform over a bounded interval if the integrand is non-periodic, and has a square root form over a bounded interval if the integrand is periodic. In higher dimensions (still for $r = 1$), [19] argues that characterizing the error distribution is not practical. Thus, for any fixed $r \geq 1$, the limit distribution as $m \rightarrow \infty$ is generally non-normal, thereby motivating our goal of defining rules on (m, r) that ensure a Gaussian CLT.

2.3.2 Scrambled Digital Nets

Owen [27, 28] introduces *scrambled digital nets* as another form of RQMC. The idea is to scramble the digits of special low-discrepancy sequences, namely, digital nets in base b [26]. The method applies random permutations to the digits in a way that preserves the low-discrepancy property. We do not provide the full description of the construction of the low-discrepancy sequence (see [26] for details), but rather focus on the randomization. For a digital net $\Xi = (\xi_i)_{i \geq 1}$ in base b and dimension s , we express the k -th coordinate ($1 \leq k \leq s$) of the i -th point $\xi_i = (\xi_i^{(1)}, \dots, \xi_i^{(s)})$ as $\xi_i^{(k)} = \sum_{\ell=1}^{\infty} \xi_i^{(k,\ell)} b^{-\ell}$ with each $\xi_i^{(k,\ell)} \in \{0, \dots, b-1\}$. If we let $U_i' = (U_i^{(1),\ell}, \dots, U_i^{(s),\ell})$ denote the i -th randomized point in a

generic randomization (therefore omitting the index j for the j -th randomization), its k -th coordinate is defined by

$$U_i^{(k),\prime} = \sum_{\ell=1}^{\infty} U_i^{(k,\ell),\prime} b^{-\ell}$$

with

$$U_i^{(k,\ell),\prime} = \begin{cases} \pi_0^k(\xi_i^{(k,1)}) & \text{if } \ell = 1 \\ \pi_{0, U_i^{(k,1),\prime}, \dots, U_i^{(k,\ell-1),\prime}}^k(\xi_i^{(k,\ell)}) & \text{if } \ell > 1. \end{cases}$$

where π_0^k , π_0 , $\pi_{0, U_i^{(k,1),\prime}, \dots, U_i^{(k,\ell-1),\prime}}^k$, $\forall k, \ell$, are independent and uniformly distributed on the set of $b!$ permutations of $\{0, \dots, b-1\}$. In other words, the digits are randomly permuted, with independent permutations. This randomization is called *nested uniform scrambling*. For a digital net, the development in base b of ξ_i is finite, meaning that the required number of permutations is finite, even if large. Other more practical forms of scrambling appear in [13, 16, 24], including the *linear matrix scramble*.

One advantage of scrambled digital nets is that they keep the discrepancy property of the original digital net. Specifically, let Ξ_{Π} denote the scrambling of a digital net Ξ in base b . If there is a constant $C > 0$ such that $D_m^*(\Xi) \leq Cm^{-1}(\ln m)^s$ for all m , then the scrambling also satisfies [16, 27]

$$D_m^*(\Xi_{\Pi}) \leq Cm^{-1}(\ln m)^s, \quad (8)$$

so the estimator X_j in (4) from a single randomization has the same convergence speed as (7) when the integrand h is a function with bounded Hardy-Krause variation. For special classes of functions [29], the RMSE of the quadrature rule based on nested uniform scrambling can be as small as $O(m^{-3/2}(\log m)^{(s-1)/2})$. For RQMC via digital nets and the full nested scrambling of Owen, [23] establishes a Gaussian CLT for $r = 1$ as $m \rightarrow \infty$. (Note that then the number of involved independent permutations also tends to infinity.) But the CLT of [23] is limited to this specific and impractical scrambling; more general CLTs are needed for other forms of RQMC.

2.3.3 Digital Shift

The *digital shift* randomization is a third possibility, which also applies a random-shift principle but specifically designed for digital nets, with the idea of preserving its digital-net structure. Formally, consider again a digital net $\Xi = (\xi_i)_{i \geq 1}$ in base b and dimension s , and recall the notation $\xi_i = (\xi_i^{(1)}, \dots, \xi_i^{(s)})$ for the i -th point, whose k -th coordinate ($1 \leq k \leq s$) is $\xi_i^{(k)} = \sum_{l=1}^{\infty} \xi_i^{(k,l)} b^{-l}$ with each $\xi_i^{(k,l)} \in \{0, \dots, b-1\}$. The j -th randomization employs a uniform $U_j = (U_j^{(1)}, \dots, U_j^{(s)}) \sim \mathcal{U}[0, 1]^s$, and we write the development in base b of its k -th coordinate as $U_j^{(k)} = \sum_{\ell=1}^{\infty} U_j^{(k,\ell)} b^{-\ell}$ (where $U_j^{(k,\ell)} \in \{0, \dots, b-1\}$). For the i -th randomized point $U_{i,j}^{\prime} = (U_{i,j}^{\prime(1)}, \dots, U_{i,j}^{\prime(s)})$ from the j -th randomization, its k -th coordinate is defined by

$$U_{i,j}^{\prime(k)} = \sum_{\ell=1}^{\infty} U_{i,j}^{\prime(k,\ell)} b^{-\ell}$$

where

$$U_{i,j}^{\prime(k,\ell)} = \sum_{l=1}^{\infty} \left[(\xi_i^{(k,l)} + U_j^{(k,\ell)}) \bmod b \right] b^{-l}.$$

As with the scrambling procedure of Section 2.3.2, the digital shift applied to a digital net retains the low-discrepancy property of the original sequence [16]. Specifically, for a digital net Ξ , let $\Xi_{U_j}^{\text{Dig}}$

be its digital shift based on $U_j \sim \mathcal{U}[0, 1]^s$. Then there exists a constant $C > 0$ such that when $D_m^*(\Xi) \leq Cm^{-1}(\ln m)^s$ for all m , we also have

$$D_m^*(\Xi_{U_j}^{\text{Dig}}) \leq Cm^{-1}(\ln m)^s. \quad (9)$$

2.4 Assumptions and Preliminary Results

Recall that for a given computation budget of about n integrand evaluations, we defined the RQMC estimator in (4) with (m, r) such that $mr \approx n$, where r is the number of randomizations and m is the number of points used from each randomized sequence. To study the asymptotic behavior as $n \rightarrow \infty$, we take $r \equiv r_n \geq 1$ and $m \equiv m_n \geq 1$ as functions of n satisfying the following:

Assumption 1.A. $m_n r_n \leq n$ for each $n \geq 1$, with $m_n r_n / n \rightarrow 1$ and $r_n \rightarrow \infty$ as $n \rightarrow \infty$.

Under Assumption 1.A, the RQMC estimator of μ in (4) becomes

$$\hat{\mu}_{m_n, r_n}^{\text{RQ}} = \frac{1}{r_n} \sum_{j=1}^{r_n} X_{n,j}, \quad \text{where} \quad X_{n,j} = \frac{1}{m_n} \sum_{i=1}^{m_n} h(U_{i,j}'), \quad (10)$$

so $X_{n,j}$ is the estimator from randomization $j = 1, 2, \dots, r_n$, of m_n points. Our goal in Section 3 is to determine conditions on the behavior of (m_n, r_n) as n grows that ensure $\hat{\mu}_{m_n, r_n}^{\text{RQ}}$ obeys a CLT (with Gaussian limit) as $n \rightarrow \infty$, and Section 4 derives conditions for an asymptotically valid CI for μ . Other papers, e.g., [8, 10], adopt a framework similar to Assumption 1.A to study simulation methods.

Assumption 1.A requires that $r_n \rightarrow \infty$ as $n \rightarrow \infty$ because otherwise, a Gaussian CLT may not hold. Indeed, as noted earlier, [19] shows that when applying RQMC using a lattice rule and the random shift, the resulting estimator can obey a limit theorem with non-Gaussian limit as $m \rightarrow \infty$ for fixed $r \geq 1$ (Section 2.3.1), and the only known Gaussian CLT being for the (impractical) nested digital scrambling [23] (Section 2.3.2). But while Assumption 1.A specifies that $r_n \rightarrow \infty$ as $n \rightarrow \infty$, it does not require that $m_n \rightarrow \infty$ as $n \rightarrow \infty$, and we sometimes take $m_n = m_0$ for some fixed $m_0 \geq 1$, where $m_0 = 1$ corresponds to MC.

Section 5 will also consider the following special case of Assumption 1.A:

Assumption 1.B. $m_n = n^c$ and $r_n = n^{1-c}$ with $c \in (0, 1)$.

Under Assumption 1.B, both $r_n, m_n \rightarrow \infty$ as $n \rightarrow \infty$ because $c \in (0, 1)$. As m_n and r_n need to be integers, we should define, e.g., $m_n = \lfloor n^c \rfloor$ and $r_n = \lfloor n^{1-c} \rfloor$, where $\lfloor \cdot \rfloor$ denotes the floor function. But to simplify the notation and analysis, we will assume that n^c and n^{1-c} are both integers. For those $c \in (0, 1)$ ensuring that $\hat{\mu}_{m_n, r_n}^{\text{RQ}}$ satisfies a Gaussian CLT or that AVCI holds, Section 5 determines the optimal c that results in the best possible convergence rate for $\text{RMSE}[\hat{\mu}_{m_n, r_n}^{\text{RQ}}]$ as $n \rightarrow \infty$.

Let Ξ' be a generic randomized low-discrepancy sequence. It is $\Xi' = \Xi_{U_j}$ in case of a random shift, $\Xi' = \Xi_{\Pi}$ for digital scrambling, and $\Xi' = \Xi_{U_j}^{\text{Dig}}$ when applying a digital shift. As we will explain below, each of these randomization methods satisfy the following condition, which will be helpful when analyzing RQMC estimators with integrands h having $V_{\text{HK}}(h) < \infty$.

Assumption 2. For the RQMC method used, there exists a constant $0 < w'_0 < \infty$ such that each randomized sequence Ξ' satisfies $D_m^*(\Xi') \leq w'_0 m^{-1}(\ln m)^s$ for all $m > 1$, where w'_0 depends on the RQMC method but not on the randomization's realization (e.g., of the random uniforms or permutations).

To see why the RQMC methods in Section 2.3 satisfy Assumption 2, note that by (3), there exists a constant $0 < w_0 < \infty$ such that $D_m^*(\Xi) \leq w_0 m^{-1} (\ln m)^s$ whenever $m > 1$. Hence, Assumption 2 holds with $w'_0 = 4^s w_0$ for a random shift by (6) (whatever the considered low-discrepancy sequence), and with $w'_0 = C$ in (8) and (9) for scrambling and a digital shift.

Our results will further depend on properties of the integrand h . We will consider four alternative conditions on h , which are presented in order of decreasing strength, as will be established in Proposition 1 below.

Assumption 3.A. *The integrand h is of bounded Hardy-Krause variation on $[0, 1]^s$, i.e., $V_{\text{HK}}(h) < \infty$.*

Assumption 3.B. *The integrand h is bounded; i.e., there exists a constant $t_0 < \infty$ such that $h(u) \leq t_0$ for all $u \in [0, 1]^s$.*

Assumption 3.C. *There exists $b > 0$ such that*

$$\mathbb{E} \left[|h(U) - \mu|^{2+b} \right] < \infty, \quad (11)$$

where $U \sim \mathcal{U}[0, 1]^s$.

Assumption 3.D. *The variance of $h(U)$, $U \sim \mathcal{U}[0, 1]^s$, is finite; i.e., (11) holds for $b = 0$.*

Assumption 3.A is a condition on the roughness of the integrand h . In contrast, Assumptions 3.B, 3.C and 3.D restrict how slowly the tails of the distribution of $h(U)$ can decrease, with 3.B being an extreme case of no tails.

Proposition 1. *Assumption 3.A is strictly stronger than Assumption 3.B, itself strictly stronger than Assumption 3.C, which in turn is strictly stronger than Assumption 3.D.*

Proof. As shown in [31], functions of bounded Hardy-Krause variation are Riemann integrable, and therefore bounded, which [30] also uses. Thus, Assumption 3.A is stronger than Assumption 3.B. To show that the relation is strict, for any dimension $s \geq 2$, consider h to be an indicator function with discontinuities not lining up with the axes: it is bounded but has $V_{\text{HK}}(h) = \infty$. An explicit example [31] is $h(u) = \mathcal{I}(\sum_{i=1}^s u_i \leq 1)$, for $u = (u_1, \dots, u_s) \in [0, 1]^s$, with $\mathcal{I}(\cdot)$ the indicator function.

We next verify that Assumption 3.B is stronger than Assumption 3.C. For a bounded h , Assumption 3.B implies $|h(u) - \mu| \leq 2t_0$ for all $u \in [0, 1]^s$. It follows that $\mathbb{E}[|h(U) - \mu|^q] \leq (2t_0)^q$ for all $q > 0$, so all of its absolute central moments in (11) are finite. To show the converse is not true, consider $h(u) = \Phi^{\leftarrow}(u)$ for $s = 1$ (recall Φ is the $\mathcal{N}(0, 1)$ CDF), where $F^{\leftarrow}(u) = \inf\{x : F(x) \geq u\}$ is the inverse of a CDF F , so h is not bounded but $h(U)$ has finite absolute central moments of all orders.

Assumption 3.C implies Assumption 3.D by Lyapounov's inequality [2, pp. 81 and 277]. To show that the relation is strict, we modify an example of [9, p. 366]. For $s = 1$, let $h(u) = F^{\leftarrow}(u)$, where F is the CDF of the density function $f(x) = k_0 \mathcal{I}(x \geq e) / [x^3 (\ln x)^2]$ with $k_0 = (\int_e^\infty dx / [x^3 (\ln x)^2])^{-1} \doteq 26.64$. As $h(U) \sim F$, we have that $\mathbb{E}[h(U)^2] = \int_e^\infty \frac{k_0 x^2}{x^3 (\ln x)^2} dx = \int_e^\infty \frac{k_0}{x (\ln x)^2} dx = \left[-\frac{k_0}{\ln x} \right]_{x=e}^\infty = k_0$, so $\text{Var}[h(U)] < \infty$. But for any fixed $b > 0$, we get $\mathbb{E}[h(U)^{2+b}] = \int_e^\infty \frac{k_0 x^{2+b}}{x^3 (\ln x)^2} dx = k_0 \int_e^\infty \frac{1}{x^{1-b} (\ln x)^2} dx \geq (k_0/k_b) \int_e^\infty \frac{1}{x^{1-b/2}} dx = \infty$ for $k_b = [4/(be)]^2$ because $(\ln x)^2 \leq k_b x^{b/2}$ for all $x \geq e$. \square

We next give some bounds on absolute central moments of the estimator $X_{n,1}$ in (10) from a single randomization, which will be useful when establishing a CLT or AVCI. An argument analogous to [35, Theorem 2] yields

Lemma 1. *Under Assumptions 1.A, 2, and 3.A, for any $q > 0$ and for all n such that $m_n > 1$,*

$$\eta_{n,q} \equiv \mathbb{E}[|X_{n,1} - \mu|^q] \leq \mathbb{E}[(V_{\text{HK}}(h)D_{m_n}^*(\Xi'))^q] \leq \left(\frac{w'_0 V_{\text{HK}}(h)(\ln m_n)^s}{m_n}\right)^q < \infty \quad (12)$$

By (12), when $V_{\text{HK}}(h) < \infty$ and $m_n \rightarrow \infty$ as $n \rightarrow \infty$, the order- q absolute central moment of $X_{n,j}$ in (10) shrinks as $O([\ln m_n]^s/m_n)^q$ as $n \rightarrow \infty$, so $\eta_{n,q} = O(m_n^{-q+\epsilon})$ for each $\epsilon > 0$. But assuming $V_{\text{HK}}(h) < \infty$ is restrictive, as noted in Proposition 1. Even if we do not have $V_{\text{HK}}(h) < \infty$, we can still bound $\eta_{n,q}$ as follows when assuming (11) holds with $q \geq 1$ replacing $2 + b$.

Lemma 2. *Under Assumption 1.A, for any $q \geq 1$, if $\mathbb{E}[|h(U) - \mu|^q] < \infty$ for $U \sim \mathcal{U}[0, 1]^s$, then $\eta_{n,q} \leq \mathbb{E}[|h(U) - \mu|^q]$ for every n .*

Proof. As seen by (10), $X_{n,1}$ averages m_n dependent terms, which we will handle through Minkowski's inequality [2, eq. (5.40)]:

$$\begin{aligned} \mathbb{E}[|X_{n,1} - \mu|^q] &= \mathbb{E}\left[\left|\sum_{i=1}^{m_n} \frac{h(U'_{i,1}) - \mu}{m_n}\right|^q\right] \leq \left[\sum_{i=1}^{m_n} \left(\mathbb{E}\left[\left|\frac{h(U'_{i,1}) - \mu}{m_n}\right|^q\right]\right)^{1/q}\right]^q \\ &= \left[m_n \left(\mathbb{E}\left[\left|\frac{h(U) - \mu}{m_n}\right|^q\right]\right)^{1/q}\right]^q = \mathbb{E}[|h(U) - \mu|^q], \end{aligned}$$

where the third step holds because each $U'_{i,1}$ $i = 1, 2, \dots, m_n$, is distributed as $U \sim \mathcal{U}[0, 1]^s$. \square

With RQMC, we typically have $\sigma_m \equiv \left(\text{Var}[\frac{1}{m} \sum_{i=1}^m h(U'_{i,1})]\right)^{1/2} = O(m^{-\alpha})$ as $m \rightarrow \infty$ with $\alpha > 1/2$ (e.g., see (12) when $V_{\text{HK}}(h) < \infty$). In this case, the RQMC standard deviation (and RMSE) for a single randomization of a low-discrepancy sequence of length m has a better rate of convergence than MC, which in comparison has $\text{RMSE}[\hat{\mu}_m^{\text{MC}}] = \sqrt{\text{Var}[\hat{\mu}_m^{\text{MC}}]} = \psi/\sqrt{m}$ by (1). Let

$$\alpha_* = -\lim_{m \rightarrow \infty} \frac{\ln(\sigma_m)}{\ln(m)}, \quad (13)$$

where we assume the limit exists, so α_* is the only constant such that the rate (as $m \rightarrow \infty$) at which σ_m decreases is strictly faster than $m^{-\alpha_*+\epsilon}$ and strictly slower than $m^{-\alpha_*-\epsilon}$ for all $\epsilon > 0$; i.e., $\sigma_m = o(m^{-\alpha_*+\epsilon})$ and $\sigma_m = \omega(m^{-\alpha_*-\epsilon})$ as $m \rightarrow \infty$ for any $\epsilon > 0$, where $f(m) = o(g(m))$ as $m \rightarrow \infty$ means that $f(m)/g(m) \rightarrow 0$ as $m \rightarrow \infty$, and $f(m) = \omega(g(m))$ as $m \rightarrow \infty$ means that $f(m)/g(m) \rightarrow \infty$ as $n \rightarrow \infty$ (i.e., $g(n) = o(f(n))$). By (12), we see that

$$\alpha_* \geq 1 \quad \text{for RQMC when } V_{\text{HK}}(h) < \infty, \quad (14)$$

as in Assumption 3.A, and we assume in general (as is typical of RQMC) that

$$\alpha_* > \frac{1}{2}. \quad (15)$$

The value of α_* depends on the particular integrand h and the method used to construct the randomized low-discrepancy sequence $(U'_{i,1})_{i \geq 1}$, but not on how (m_n, r_n) or c is specified in Assumptions 1.A and 1.B.

Similarly, for a randomized sequence $\Xi' = (U'_i)_{i \geq 1}$, let $E_m(\Xi') = \frac{1}{m} \sum_{i=1}^m h(U'_i) - \mu$ be the error of the estimator using the first m points. Then define its exponential rate as

$$\alpha(\Xi') = - \lim_{m \rightarrow \infty} \frac{\ln(|E_m(\Xi')|)}{\ln(m)}, \quad (16)$$

assuming such a limit always exists,

$$\alpha' = \inf_{\Xi'} \alpha(\Xi'). \quad (17)$$

Thus, for any randomized sequence Ξ' , we see that $|E_m(\Xi')| = o(m^{-\alpha+\epsilon})$ and $|E_m(\Xi')| = \omega(m^{-\alpha-\epsilon})$ as $m \rightarrow \infty$ for each $\epsilon > 0$, and $\alpha' \leq \alpha_*$ always holds.

Assumption 4. *The convergence rate exponent α' of the worst-case error among all randomizations is the same as the standard deviation rate exponent α_* ; i.e., $\alpha' = \alpha_*$.*

As we will see later in Section 5.1, when Assumption 4 holds, the RQMC estimator $\hat{\mu}_{m_n, r_n}^{\text{RQ}}$ will obey a CLT for $(m_n, r_n) = (n^c, n^{1-c})$ for any $c \in (0, 1)$ (Assumption 1.B). However, establishing Assumption 4 may be difficult in practice, so much of our paper focuses on providing other more verifiable sufficient conditions that secure a CLT.

For a given total number n of integrand evaluations, a common suggestion when using RQMC methods is to let m_n be as large as possible to benefit from the superior convergence speed of QMC (compared to MC), but we still want r_n to be big enough so that a Gaussian CLT roughly holds (see the discussion at the end of Section 2.3.1) and an asymptotically valid CI for μ can be constructed. As in (12), when $V_{\text{HK}}(h) < \infty$ and $m_n \rightarrow \infty$ as $n \rightarrow \infty$, we have that

$$\text{RMSE}[\hat{\mu}_{m_n, r_n}^{\text{RQ}}] \leq \frac{[w'_0 V_{\text{HK}}(h) (\ln m_n)^s / m_n]}{\sqrt{r_n}} = \Theta \left(\frac{(c \ln n)^s}{n^{(1+c)/2}} \right)$$

by replacing m_n and r_n by n^c and n^{1-c} , respectively, as in Assumption 1.B. Hence, larger c leads to faster convergence. Taking $c = 1$ is optimal in this respect but then a Gaussian CLT may not be guaranteed (see [19]), as noted earlier.

The following sections establish various conditions that ensure $\hat{\mu}_{m_n, r_n}^{\text{RQ}}$ obeys a CLT and when we can obtain an asymptotically valid confidence interval for μ . Sections 3 and 4 derive these conditions when (m_n, r_n) satisfy Assumption 1.A, whereas Section 5 instead adopts Assumption 1.B, which permits simpler and more intuitive analysis.

3 General Conditions for a Central Limit Theorem

When analyzing $\hat{\mu}_{m_n, r_n}^{\text{RQ}}$ in (10) as $n \rightarrow \infty$, we encounter a complication: the distribution of the averaged terms $X_{n,1}, X_{n,2}, \dots, X_{n,r_n}$ changes with n . A theoretical framework for handling this under Assumption 1.A uses that the $(X_{n,j})_{n=1,2,\dots;j=1,2,\dots,r_n}$ in (10) form a *triangular array* [2, p. 359] (which other authors, e.g., [32, Section 1.9.3] and [6, p. 205], refer to as a *double array*). In a triangular array, the r_n variables within each row n are independent, but variables in different rows may be dependent. Let $\mu_{n,j} = \mathbb{E}[X_{n,j}] = \mu$ and

$$\sigma_{n,j}^2 = \text{Var}[X_{n,j}] \equiv \sigma_{m_n}^2,$$

where both μ and $\sigma_{m_n}^2$ do not depend on j . Indeed, we have

$$X_{n,1}, X_{n,2}, \dots, X_{n,r_n} \text{ are i.i.d., each with some distribution } F_n. \quad (18)$$

This setup allows for the distribution F_n to change with n , as is the case in (10).

Although (18) has that the r_n random variables are i.i.d. for each n , the general setting for a triangular array, as in [2, p. 359], assumes that they are only independent but not necessarily identically distributed (nor that they have the same mean and variance). Specifically, recall the Lindeberg CLT [2, Theorem 27.2] for triangular arrays:

For each n , assume that $X_{n,j}$, $j = 1, \dots, r_n$, are independent (but not necessarily identically distributed), with $E[X_{n,j}] = \mu_{n,j}$ and $\text{Var}[X_{n,j}] = \sigma_{n,j}^2 < \infty$, and let $s_n^2 = \sigma_{n,1}^2 + \dots + \sigma_{n,r_n}^2$. Let $G_{n,j}$ denote the distribution of $Y_{n,j} = X_{n,j} - \mu_{n,j}$ and

$$\tau_{n,j}^2(t) = \int_{|y| > ts_n} y^2 dG_{n,j}(y)$$

for $t \geq 0$. Also, let $\bar{X}_n = (1/r_n) \sum_{j=1}^{r_n} X_{n,j}$ and $\mu_n = (1/r_n) \sum_{j=1}^{r_n} \mu_{n,j}$. Then under Assumption 1.A,

$$\frac{\bar{X}_n - \mu_n}{\sqrt{\text{Var}[\bar{X}_n]}} \Rightarrow \mathcal{N}(0, 1), \quad \text{as } n \rightarrow \infty \quad (19)$$

provided that the Lindeberg condition holds:

$$\frac{\tau_{n,1}^2(t) + \dots + \tau_{n,r_n}^2(t)}{\sigma_{n,1}^2 + \dots + \sigma_{n,r_n}^2} \rightarrow 0, \quad \text{as } n \rightarrow \infty, \quad \forall t > 0. \quad (20)$$

The Lindeberg condition (20) implies that $\max_{1 \leq j \leq r_n} \sigma_{n,j}^2 / s_n^2 \rightarrow 0$ as $n \rightarrow \infty$ (e.g., [20, p. 588]), so it ensures that the contribution of any single $X_{n,j}$, $1 \leq j \leq r_n$, to their sum's variance s_n^2 is negligible for large n . This then precludes the possibility that, e.g., $X_{n,j} \equiv 0$, $2 \leq j \leq r_n$, so the left side of (19) reduces to $(X_{n,1} - \mu_{n,1}) / \sigma_{n,1}$, which can have any distribution with mean 0 and variance 1.

We next apply the Lindeberg condition to analyze the RQMC estimator $\hat{\mu}_{m_n, r_n}^{\text{RQ}}$ in (10). By (18), we get

$$s_n^2 = \sigma_{n,1}^2 + \dots + \sigma_{n,r_n}^2 = r_n \sigma_{m_n}^2. \quad (21)$$

Denote the distribution of $Y_{n,j} = X_{n,j} - \mu$ by G_n , which does not depend on j by (18). Note that $\sigma_{m_n}^2 = \int_{y \in \mathfrak{R}} y^2 dG_n(y)$, and let

$$\tau_n^2(t) = \int_{|y| > ts_n} y^2 dG_n(y), \quad \text{for } t > 0. \quad (22)$$

We will impose another assumption to preclude uninteresting cases. Without the following, the exact result is eventually always returned by the RQMC estimator.

Assumption 5. $\sigma_{m_n}^2 > 0$ for all n large enough.

Then for RQMC, we can specialize the Lindeberg condition (20) as (23) below.

Theorem 1. *Suppose that Assumptions 1.A and 5 hold. If the Lindeberg condition*

$$\frac{\tau_n^2(t)}{\sigma_{m_n}^2} \rightarrow 0, \quad \text{as } n \rightarrow \infty, \quad \forall t > 0 \quad (23)$$

holds, then the RQMC estimator in (10) satisfies the CLT

$$\frac{\hat{\mu}_{m_n, r_n}^{\text{RQ}} - \mu}{\sigma_{m_n} / \sqrt{r_n}} \Rightarrow \mathcal{N}(0, 1), \quad \text{as } n \rightarrow \infty. \quad (24)$$

From [2, p. 361] or [20, p. 572], condition (23) is even necessary and sufficient for the CLT (24) since for all j , $\sigma_{n,j}^2/s_n^2 = 1/r_n \rightarrow 0$ as $n \rightarrow \infty$ by Assumption 1.A. The Lindeberg condition (23), which imposes restrictions on the tail behavior of G_n through (22), holds under a Lyapounov (moment) condition [2, Theorem 27.3], which (25) below specializes for our RQMC setting.

Theorem 2. *Suppose that Assumptions 1.A and 5 hold. Further suppose that for each n , there exists $b > 0$ such that $\mathbb{E} \left[|X_{n,1} - \mu|^{2+b} \right] < \infty$ and that*

$$\frac{\mathbb{E} \left[|X_{n,1} - \mu|^{2+b} \right]}{r_n^{b/2} \sigma_{m_n}^{2+b}} \rightarrow 0, \quad \text{as } n \rightarrow \infty. \quad (25)$$

Then the Lindeberg condition (23) and CLT (24) hold.

Proof. Fix any $t > 0$ for $\tau_n^2(t)$ in (22), and we will bound $\tau_n^2(t)$ as in the proof of the Lyapounov CLT in [2, Theorem 27.3]. Specifically, the condition $|y| > ts_n$ in (22) implies that $|y|^b/(t^b s_n^b) > 1$, so

$$\tau_n^2(t) \leq \frac{1}{t^b s_n^b} \int_{|y| > ts_n} y^2 |y|^b dG_n(y) \leq \frac{1}{t^b s_n^b} \mathbb{E} \left[|Y_{n,1}|^{2+b} \right] = \frac{1}{t^b r_n^{b/2} \sigma_{m_n}^b} \mathbb{E} \left[|X_{n,1} - \mu|^{2+b} \right] \quad (26)$$

by (21), with $\mathbb{E} \left[|X_{n,1} - \mu|^{2+b} \right] < \infty$ by assumption, which ensures $\sigma_{m_n}^2 < \infty$. Thus, as in (23), dividing throughout (26) by $\sigma_{m_n}^2$, which is strictly positive for all n large enough (Assumption 5), shows that (25) guarantees (23) because $t > 0$ is fixed and arbitrary, so CLT (24) holds by Theorem 1. \square

3.1 Corollaries of Theorems 1 and 2

We now develop various properties that imply the assumptions of Theorems 1 or 2 hold to secure CLT (24). We first derive a corollary of Theorem 1 based on Assumption 4, which imposes constraints on both the integrand and RQMC sequence.

Corollary 1. *Suppose that Assumptions 1.A, 4, and 5 hold. If $m_n \rightarrow \infty$ as $n \rightarrow \infty$ and*

$$\frac{r_n}{m_n^{\lambda'}} \rightarrow \infty, \quad \text{as } n \rightarrow \infty \quad (27)$$

for some $\lambda' > 0$, which can be chosen arbitrarily small, then the Lindeberg condition (23) and CLT (24) hold.

Proof. For any $t > 0$, let $A_n(t) = \{y : |y| > t\sqrt{r_n}\sigma_{m_n}\}$ be the domain of integration of $\tau_n^2(t)$ in (22). Under Assumption 5 and because $m_n \rightarrow \infty$ as $n \rightarrow \infty$, (13) implies that $\sigma_{m_n} = \omega(m_n^{-\alpha_* - \lambda'})$ as $n \rightarrow \infty$, so there exists $0 < k_0 < \infty$ such that $\sigma_{m_n} \geq k_0 m_n^{-\alpha_* - \lambda'}$ for all large enough n . Thus, the lower bound in $A_n(t)$ satisfies $t\sqrt{r_n}\sigma_{m_n} \geq tk_0\sqrt{r_n}m_n^{-\alpha_* - \lambda'}$ for all n large enough. Similarly, for any randomized sequence $\Xi' = (U'_i)_{i \geq 1}$, (16) and (17) ensure that the error $E_{m_n} \equiv \frac{1}{m_n} \sum_{i=1}^{m_n} h(U'_i) - \mu$ has $|E_{m_n}| = o(m_n^{-\alpha_* + \lambda'})$ as $n \rightarrow \infty$, so there exists $0 < k_1 < \infty$ such that $|E_{m_n}| \leq k_1 m_n^{-\alpha_* + \lambda'} = k_1 m_n^{-\alpha_* + \lambda'}$ for all n sufficiently large by Assumption 4. Hence, we get $A_n(t) \subseteq \{y : tk_0\sqrt{r_n}m_n^{-\alpha_* - \lambda'} < |y| \leq k_1 m_n^{-\alpha_* + \lambda'}\} \equiv A'_n(t)$ for all large enough n . Comparing the lower and upper bounds for $|y|$ in $A'_n(t)$ shows that $A'_n(t) = \emptyset$ for all n sufficiently large when $\sqrt{r_n}m_n^{-\alpha_* - \lambda'}/m_n^{-\alpha_* + \lambda'} \rightarrow \infty$ as $n \rightarrow \infty$, which is equivalent to $r_n/m_n^{4\lambda'} \rightarrow \infty$ as $n \rightarrow \infty$. But as $\lambda' > 0$ can be chosen arbitrarily small, this last condition is the same as (27), so we see that $A_n(t) \subseteq A'_n(t) = \emptyset$ for all n large enough, which guarantees that the Lindeberg condition (23) holds, yielding CLT (24) by Theorem 1. \square

Condition (27) in Corollary 1 specifies that $r_n \rightarrow \infty$ does not grow too slowly (as $n \rightarrow \infty$) compared to m_n . For example, (27) does not hold for $(m_n, r_n) = (\lfloor n/\ln n \rfloor, \lfloor \ln n \rfloor)$. But as we will see later in Section 5.1, (27) is always true under Assumption 1.B. Thus, we can choose $(m_n, r_n) = (n^{1-\epsilon}, n^\epsilon)$ for arbitrarily small $\epsilon > 0$, which enables exploiting QMC's superior convergence rate.

As establishing Assumption 4 may be difficult in practice, we next provide other conditions that are more readily verifiable to ensure CLT (24). We will prove corollaries corresponding to each of our restrictions on the integrand h in Assumptions 3.A–3.D, which are in decreasing order of strength (Proposition 1).

Corollary 2. *Suppose that Assumption 1.A holds with $m_n > 1$ for all n large enough, along with Assumptions 2, 3.A ($V_{\text{HK}}(h) < \infty$), and 5. Also, suppose that*

$$r_n^{1-\lambda} \left(\frac{m_n \sigma_{m_n}}{(\ln m_n)^s} \right)^2 \rightarrow \infty, \quad \text{as } n \rightarrow \infty, \quad (28)$$

for some constant $\lambda \in (0, 1)$, which can be chosen arbitrarily small. Then the Lyapounov condition (25) and CLT (24) hold.

Proof. We will verify the conditions of Theorem 2 for $b = 2(1 - \lambda)/\lambda > 0$. By assumption, we have that $m_n > 1$ for all n sufficiently large, and consider any such n for which Assumption 5 also holds. Because $V_{\text{HK}}(h) < \infty$, Lemma 1 guarantees that $\mathbb{E}[|X_{n,1} - \mu|^{2+b}] < \infty$ by (12) with $q = 2 + b$, so $0 < \sigma_{m_n}^2 < \infty$. Also, applying (12) of Lemma 1 ensures that the left side of (25) satisfies

$$\frac{\mathbb{E}[|X_{n,1} - \mu|^{2+b}]}{r_n^{b/2} \sigma_{m_n}^{2+b}} \leq \frac{1}{r_n^{b/2}} \left(\frac{w'_0 V_{\text{HK}}(h) (\ln m_n)^s}{m_n \sigma_{m_n}} \right)^{2+b}. \quad (29)$$

Raising the right side of (29) to the $2/(2 + b)$ power shows that it vanishes as $n \rightarrow \infty$ if and only if

$$r_n^{b/(2+b)} \left(\frac{m_n \sigma_{m_n}}{(\ln m_n)^s} \right)^2 \rightarrow \infty, \quad \text{as } n \rightarrow \infty \quad (30)$$

But $b/(2 + b) = 1 - \lambda$, so (28) and (30) imply (25), and (23) and (24) follow from Theorem 2. \square

The following corollary of Theorem 1 considers the case when the integrand h is a bounded function.

Corollary 3. *Suppose that Assumptions 1.A, 3.B (h is bounded), and 5 hold. If*

$$s_n^2 = r_n \sigma_{m_n}^2 \rightarrow \infty, \quad \text{as } n \rightarrow \infty, \quad (31)$$

then the Lindeberg condition (23) and CLT (24) hold.

Proof. As h is bounded, Assumption 3.B implies that $|Y_{n,j}| \leq 2t_0$ for all n and j , ensuring $\tau_n^2(t) = 0$ whenever $ts_n > 2t_0$. Given any $t > 0$, it follows from (31) that there exists n_0 such that $ts_n > 2t_0$ for all $n > n_0$, so the numerator of (23) is zero for all $n > n_0$. Thus, Theorem 1 implies (24). \square

As with Corollary 2, the next corollary follows from Theorem 2, but it does not require $V_{\text{HK}}(h) < \infty$, precluding the application of Lemma 1. It instead employs Lemma 2, so it assumes a moment condition on $h(U)$ (Assumption 3.C).

Corollary 4. *Suppose that Assumptions 1.A, 3.C (finite absolute central moment of order $2 + b$ for some $b > 0$), and 5 hold. Also, suppose that*

$$r_n^{b/(2+b)} \sigma_{m_n}^2 \rightarrow \infty, \quad \text{as } n \rightarrow \infty. \quad (32)$$

Then the Lyapounov condition (25) and CLT (24) hold.

Proof. Because $b > 0$, raising (32) to the $(2 + b)/2$ power is equivalent to

$$r_n^{b/2} \sigma_{m_n}^{2+b} \rightarrow \infty, \quad \text{as } n \rightarrow \infty. \quad (33)$$

By (11), Lemma 2 with $q = 2 + b$ then implies that the left side of (25) satisfies

$$\frac{\mathbb{E} \left[|X_{n,1} - \mu|^{2+b} \right]}{r_n^{b/2} \sigma_{m_n}^{2+b}} \leq \frac{\mathbb{E} \left[|h(U) - \mu|^{2+b} \right]}{r_n^{b/2} \sigma_{m_n}^{2+b}} \rightarrow 0, \quad \text{as } n \rightarrow \infty \quad (34)$$

by (33), so (25) holds. Hence, (23) and (24) follow from Theorem 2. \square

While Assumption 1.A specifies that $r_n \rightarrow \infty$ as $n \rightarrow \infty$, it does not require that $m_n \rightarrow \infty$ as $n \rightarrow \infty$. All of the previous corollaries allow $m_n \rightarrow \infty$ as $n \rightarrow \infty$, although this was not required. The following result specializes Theorem 1 to consider the case when m_n is fixed and only r_n grows as n increases.

Corollary 5. *Suppose that Assumption 3.D ($\text{Var}[h(U)] < \infty$) holds. If*

$$(m_n, r_n) = \left(m_0, \left\lfloor \frac{n}{m_0} \right\rfloor \right) \text{ for all } n, \text{ for some fixed } m_0 \geq 1 \text{ with } \sigma_{m_0}^2 > 0, \quad (35)$$

which implies Assumptions 1.A and 5, then the Lindeberg condition (23) and CLT (24) hold.

Proof. For $n \geq m_0$, we have by (35) that $X_{n,j} = X_j$ for $j = 1, 2, \dots, r_n$, with $X_j = (1/m_0) \sum_{i=1}^{m_0} h(U'_{i,j})$ not depending on n because $m_n = m_0$ is fixed. Thus, X_1, X_2, \dots form a single i.i.d. sequence with mean μ and variance $\sigma_{m_0}^2 = \text{Var}[X_1] = \text{Var}[X_{n,1}]$, which Lemma 2 with $q = 2$ ensures is finite by Assumption 3.D. For G as the distribution of $X_1 - \mu$, (23) reduces to $\lim_{n \rightarrow \infty} \frac{1}{\sigma_{m_0}^2} \int_{|y| > t\sigma_{m_0}\sqrt{r_n}} y^2 dG(y) \rightarrow 0$ as $n \rightarrow \infty$ for each $t > 0$, which holds because $\{|y| > t\sigma_{m_0}\sqrt{r_n}\} \rightarrow \emptyset$ as $n \rightarrow \infty$ since $0 < \sigma_{m_0} < \infty$ and $r_n \rightarrow \infty$. Hence, Theorem 1 yields (24). \square

3.2 Comparison of Conditions for CLT

We now want to compare Corollaries 2–5 from Section 3.1, each of which ensures the CLT (24). Proposition 1 establishes that these corollaries impose successively weaker restrictions on the integrand h . We next show that in many settings, the corollaries require correspondingly stronger conditions on (m_n, r_n) . (Sections 5.8 and 5.10 will provide further comparisons under Assumption 1.B.)

Proposition 2. *If Assumption 1.A holds, then condition (35) of Corollary 5 implies condition (32) of Corollary 4, and (32) in turn implies condition (31) of Corollary 3. If in addition*

$$\frac{1}{r_n^\lambda} \left(\frac{m_n}{(\ln m_n)^s} \right)^2 \rightarrow d_0 \in (0, \infty], \quad \text{as } n \rightarrow \infty, \text{ for some } \lambda \in (0, 1), \quad (36)$$

which holds, e.g., under Assumption 1.B, then (31) implies condition (28) of Corollary 2.

Proof. We first compare the restrictions on (m_n, r_n) in Corollaries 4 and 5. If (35) holds, then $r_n^{b/(2+b)} \sigma_{m_n}^2 = (\lfloor n/m_0 \rfloor)^{b/(2+b)} \sigma_{m_0}^2 \rightarrow \infty$ as $n \rightarrow \infty$ because $\sigma_{m_0}^2 > 0$, so (35) implies (32).

We next analyze the requirements on (m_n, r_n) in Corollaries 3 and 4. Expressing $r_n \sigma_{m_n}^2 = r_n^{2/(2+b)} r_n^{b/(2+b)} \sigma_{m_n}^2$ shows that (32) implies (31) since $r_n^{2/(2+b)} \rightarrow \infty$ as $n \rightarrow \infty$ by Assumption 1.A.

We finally compare condition (28) of Corollary 2 with condition (31) of Corollary 3. Writing the left side of (28) as

$$r_n^{1-\lambda} \left(\frac{m_n \sigma_{m_n}}{(\ln m_n)^s} \right)^2 = [r_n \sigma_{m_n}^2] \left[\frac{1}{r_n^\lambda} \left(\frac{m_n}{(\ln m_n)^s} \right)^2 \right]$$

makes clear that choosing (m_n, r_n) to satisfy (31) ensures that (28) is satisfied when (36) holds. To verify (36) under Assumption 1.B, observe that $(m_n, r_n) = (n^c, n^{1-c})$ for $c \in (0, 1)$ leads to the left side of (36) becoming $\frac{1}{n^{\lambda(1-c)}} \left(\frac{n^c}{(\ln n^c)^s} \right)^2 = n^{2c-\lambda(1-c)}/(c \ln n)^{2s}$, which grows to ∞ as $n \rightarrow \infty$ for any $\lambda \in (0, \lambda_0)$, where $\lambda_0 = \min(2c/(1-c), 1) > 0$. \square

The condition (36) in Proposition 2 specifies that r_n does not grow too quickly (as $n \rightarrow \infty$) compared to m_n . While (36) is always true under Assumption 1.B, it does not hold, e.g., for fixed $m_n \equiv m_0$, as in condition (35) of Corollary 5, or for $(m_n, r_n) = (\lfloor \ln n \rfloor, \lfloor n/\ln n \rfloor)$.

4 Asymptotically Valid Confidence Interval

We now want to leverage the CLTs in Section 3 to construct an asymptotically valid CI for μ under the framework of Assumption 1.A. Suppose that $r_n \geq 2$, which Assumption 1.A ensures holds for all n sufficiently large. For each such n , the $X_{n,j}$, $j = 1, 2, \dots, r_n$, are i.i.d. by (18), and their sample variance

$$\widehat{\sigma}_{m_n, r_n}^2 = \frac{1}{r_n - 1} \sum_{j=1}^{r_n} (X_{n,j} - \widehat{\mu}_{m_n, r_n}^{\text{RQ}})^2 \quad (37)$$

is an unbiased estimator of $\sigma_{m_n}^2 = \text{Var}[X_{n,1}]$. For a given desired confidence level $100\gamma\%$, with $0 < \gamma < 1$, we then consider a CI for μ as

$$I_{m_n, r_n, \gamma}^{\text{RQ}} \equiv [\widehat{\mu}_{m_n, r_n}^{\text{RQ}} \pm z_\gamma \widehat{\sigma}_{m_n, r_n} / \sqrt{r_n}]. \quad (38)$$

Our goal now is to provide conditions ensuring that $I_{n, \gamma}$ is an AVCI, i.e., (40) below holds.

Theorem 3. *Suppose that Assumptions 1.A and 5 hold. If the CLT (24) holds and if*

$$\frac{\widehat{\sigma}_{m_n, r_n}^2}{\sigma_{m_n}^2} \Rightarrow 1, \quad \text{as } n \rightarrow \infty, \quad (39)$$

then

$$\lim_{n \rightarrow \infty} P(\mu \in I_{m_n, r_n, \gamma}^{\text{RQ}}) = \gamma. \quad (40)$$

Proof. Note that (39) implies that $\sigma_{m_n} / \widehat{\sigma}_{m_n, r_n} \Rightarrow 1$ as $n \rightarrow \infty$ by the continuous-mapping theorem (e.g., [2, Theorem 25.7]). Thus, (24) ensures that

$$\frac{\widehat{\mu}_{m_n, r_n}^{\text{RQ}} - \mu}{\widehat{\sigma}_{m_n, r_n} / \sqrt{r_n}} = \left(\frac{\sigma_{m_n}}{\widehat{\sigma}_{m_n, r_n}} \right) \frac{\widehat{\mu}_{m_n, r_n}^{\text{RQ}} - \mu}{\sigma_{m_n} / \sqrt{r_n}} \Rightarrow \mathcal{N}(0, 1), \quad \text{as } n \rightarrow \infty, \quad (41)$$

by Slutsky's theorem [32, Theorem 1.5.4]. Hence,

$$P(\mu \in I_{m_n, r_n, \gamma}^{\text{RQ}}) = P\left(-z_\gamma \leq \frac{\widehat{\mu}_{m_n, r_n}^{\text{RQ}} - \mu}{\widehat{\sigma}_{m_n, r_n} / \sqrt{r_n}} \leq z_\gamma\right) \rightarrow \gamma, \quad \text{as } n \rightarrow \infty,$$

by the portmanteau theorem [2, Theorem 25.8], establishing (40). \square

We now want a sufficient condition ensuring that (39) holds so that we can have AVCI (40).

Theorem 4. *Suppose that Assumptions 1.A and 5 hold. If $\mathbb{E}[(X_{n,1} - \mu)^4] < \infty$ and*

$$\frac{\mathbb{E}[(X_{n,1} - \mu)^4]}{r_n \sigma_{m_n}^4} \rightarrow 0, \quad \text{as } n \rightarrow \infty, \quad (42)$$

then the CLT (24), (39), and AVCI (40) hold.

Proof. Note that (42) is equivalent to (25) for $b = 2$, so Theorem 2 guarantees (24) because we assumed that $\mathbb{E}[(X_{n,1} - \mu)^4] < \infty$. Thus, if (39) holds, then Theorem 3 will imply (40).

We will prove (39) by establishing that $p_n \equiv P\left(\left|\frac{\widehat{\sigma}_{m_n, r_n}^2}{\sigma_{m_n}^2} - 1\right| > v\right) \rightarrow 0$ as $n \rightarrow \infty$ for each fixed $v > 0$. Assume that n is large enough so that $r_n \geq 2$ (see Assumption 1.A). By Chebyshev's inequality [2, eq. (21.13)] and because $\mathbb{E}[\widehat{\sigma}_{m_n, r_n}^2] = \sigma_{m_n}^2$ [32, p. 173], we get

$$\begin{aligned} p_n &\leq \frac{1}{v^2} E\left[\left(\frac{\widehat{\sigma}_{m_n, r_n}^2}{\sigma_{m_n}^2} - 1\right)^2\right] = \frac{1}{v^2 \sigma_{m_n}^4} E\left[\left(\widehat{\sigma}_{m_n, r_n}^2 - \sigma_{m_n}^2\right)^2\right] = \frac{1}{v^2 \sigma_{m_n}^4} \text{Var}\left[\widehat{\sigma}_{m_n, r_n}^2\right] \\ &= \frac{1}{v^2 r_n \sigma_{m_n}^4} \left(\mathbb{E}[(X_{n,1} - \mu)^4] - \frac{r_n - 3}{r_n - 1} \sigma_{m_n}^4\right), \end{aligned} \quad (43)$$

where (43) follows from [32, p. 184], with $\mathbb{E}[(X_{n,1} - \mu)^4] < \infty$ by assumption. Because $r_n \rightarrow \infty$ as $n \rightarrow \infty$ by Assumption 1.A, (43) vanishes as $n \rightarrow \infty$ by (42) since $v > 0$ is fixed, thus verifying (39). \square

4.1 Specific Sufficient Conditions for AVCI

We next establish AVCI (40) (often through Theorem 4) by separately considering Assumptions 3.A, 3.C, and 3.D.

Corollary 6. *Suppose that Assumption 1.A holds with $m_n > 1$ for all n large enough, along with Assumptions 2, 3.A (h is of bounded Hardy-Krause variation), and 5. If*

$$r_n \left(\frac{m_n \sigma_{m_n}}{(\ln m_n)^s}\right)^4 \rightarrow \infty, \quad \text{as } n \rightarrow \infty, \quad (44)$$

then the CLT (24), (39), and AVCI (40) hold.

Proof. Taking the square-root of (44) shows that it is equivalent to (28) with $\lambda = 1/2$, so (44) guarantees CLT (24) by Corollary 2. We will next establish (39) and (40) by verifying the conditions of Theorem 4. By assumption, we have that $m_n > 1$ for all n sufficiently large, and consider any such

n for which Assumption 5 also holds. Because $V_{\text{HK}}(h) < \infty$ from Assumption 3.A, Lemma 1 implies that $\mathbb{E}[(X_{n,1} - \mu)^4] < \infty$ by (12) with $q = 4$. Also, (12) ensures that the left side of (42) satisfies

$$\frac{\mathbb{E}[(X_{n,1} - \mu)^4]}{r_n \sigma_{m_n}^4} \leq \frac{1}{r_n} \left(\frac{w'_0 V_{\text{HK}}(h) (\ln m_n)^s}{m_n \sigma_{m_n}} \right)^4. \quad (45)$$

Using (44) in (45) verifies (42), so (39) and (40) follow from Theorem 4. \square

The following result instead assumes a moment condition for $h(U)$.

Corollary 7. *Suppose that Assumptions 1.A and 5 hold, along with Assumption 3.C (finite absolute central moment) for $b = 2$. If*

$$r_n \sigma_{m_n}^4 \rightarrow \infty, \quad \text{as } n \rightarrow \infty, \quad (46)$$

then the CLT (24), (39), and AVCI (40) hold.

Proof. Taking the square-root of (46) shows it is equivalent to (32) with $b = 2$, so (46) ensures CLT (24) by Corollary 4 because we further assumed that (11) holds for $b = 2$. We will next establish (39) and (40) by verifying the conditions of Theorem 4. By Lemma 2 with $q = 4$, we see that the numerator in (42) is bounded by $\mathbb{E}[(h(U) - \mu)^4]$, which does not depend on n and is finite under the assumed validity of (11) for $b = 2$. Thus, (46) implies (42), so Theorem 4 yields (39) and (40). \square

The next result considers the case when $m_n \equiv m_0$ is fixed, as in condition (35) of Corollary 5.

Corollary 8. *Under the same conditions as in Corollary 5, (39) and AVCI (40) hold.*

Proof. Corollary 5 ensures the CLT (24) is true. Also, as each $X_{n,j} = X_j$ in (18) does not depend on n , the triangular array reduces to a single i.i.d. sequence. Consequently, standard arguments (e.g., [32, Theorem 2.2.3A] or [20, Example 2.1.3]) show that (39) holds, so Theorem 3 implies (40). \square

Corollaries 5 and 8 assume the same conditions, where the former establishes the CLT (24), and the latter proves that (38) is AVCI (40). Thus, when $m_n = m_0$ is fixed, AVCI does not require any additional conditions beyond that for a CLT. Corollaries 6 and 7 also allow for $m_n = m_0$, but further permit $m_n \rightarrow \infty$. But if $m_n = m_0$ is fixed, Corollary 8 ensures AVCI (40) more broadly than Corollaries 6 and 7, as the latter impose additional restrictions on integrand h .

4.2 Remarks on Comparing Conditions for CLT and AVCI

We now compare Corollaries 6, 7, and 8, all of which ensure AVCI (40). Recall that Proposition 1 establishes that Assumption 3.A in Corollary 6 is strictly stronger than Assumption 3.C in Corollary 7, and the latter restriction is strictly stronger than Assumption 3.D. We next compare the corollaries' constraints on (m_n, r_n) .

Proposition 3. *If Assumption 1.A holds, then condition (35) used in Corollary 8 implies condition (46) of Corollary 7. If in addition $m_n > 1$ for all n sufficiently large, then (46) implies condition (44) of Corollary 6.*

Proof. If (35) holds, then $r_n \sigma_{m_n}^4 = \lfloor n/m_0 \rfloor \sigma_{m_0}^4 \rightarrow \infty$ as $n \rightarrow \infty$ because $\sigma_{m_0}^4 > 0$, so (35) implies (46). To compare (44) and (46), write the left side of the former as $r_n \left(\frac{m_n \sigma_{m_n}}{(\ln m_n)^s} \right)^4 = r_n \sigma_{m_n}^4 \left(\frac{m_n}{(\ln m_n)^s} \right)^4$, so (44) is weaker than (46). \square

Comparing the conditions securing CLT (24) and AVCI (40) provides further insights. Corollaries 5 and 8 assume the same conditions, so when $m_n = m_0$ is fixed, AVCI (40) does not require any additional conditions beyond that for the CLT (24). Theorem 4 and Corollaries 6 and 7 provide sufficient conditions for AVCI, which also imply CLT (24). Ideally, we would have that AVCI is true whenever the CLT holds without any additional restrictions. To study this, we could compare (the conditions of) Theorem 2 to Theorem 4, compare Corollary 2 to Corollary 6, and compare Corollary 4 to Corollary 7. However, as such a comparison is long, we instead will carry out the analysis only under Assumption 1.B in Sections 5.8–5.10, which will lead to simpler and more intuitive results.

5 Analysis When $(m_n, r_n) = (n^c, n^{1-c})$ (Assumption 1.B)

Recall that Assumption 1.B specializes Assumption 1.A so that $(m_n, r_n) = (n^c, n^{1-c})$ for some $c \in (0, 1)$. We now want to utilize the results of Sections 3 and 4 to determine what values of c ensure CLT (24) or AVCI (40), and which of those c lead to the best rates of convergence for $\text{RMSE}[\hat{\mu}_{m_n, r_n}^{\text{RQ}}]$ as $n \rightarrow \infty$. Proposition 4 at the end of this section summarizes the results under Assumption 1.B.

For $m_n = n^c$, we get from (13) that (as c is bounded)

$$\sigma_{m_n} = \omega(n^{-c\alpha_* - \epsilon}), \quad \text{as } n \rightarrow \infty, \text{ for any } \epsilon > 0. \quad (47)$$

Therefore, using $r_n = n^{1-c}$ and (10), (18), and (47) lead to

$$\text{RMSE}[\hat{\mu}_{m_n, r_n}^{\text{RQ}}] = \frac{\sigma_{m_n}}{\sqrt{r_n}} = \omega\left(n^{-v(\alpha_*, c) - \epsilon}\right), \quad \text{as } n \rightarrow \infty, \text{ for any } \epsilon > 0, \quad (48)$$

where $v(\alpha_*, c) \equiv c[\alpha_* - \frac{1}{2}] + \frac{1}{2}$. We similarly have $\text{RMSE}[\hat{\mu}_{m_n, r_n}^{\text{RQ}}] = O(n^{-v(\alpha_*, c) + \epsilon})$ as $n \rightarrow \infty$ for all $\epsilon > 0$. As ϵ can be arbitrarily small, $\text{RMSE}[\hat{\mu}_{m_n, r_n}^{\text{RQ}}]$ then decreases at about rate (ignoring leading coefficients and lower-order terms) $n^{-v(\alpha_*, c)}$ as $n \rightarrow \infty$. Our assumption (15) ensures $v(\alpha_*, c) > 1/2$ for every $c \in (0, 1)$. Consequently, for *any* choice of c in Assumption 1.B, the RQMC estimator's RMSE shrinks faster than the MC estimator's RMSE, which decreases at rate $n^{-v_{\text{MC}}}$ as $n \rightarrow \infty$, where

$$v_{\text{MC}} \equiv \frac{1}{2} \quad (49)$$

by (1). Moreover, for any fixed α_* satisfying (15), $v(\alpha_*, c)$ is strictly increasing in c , so choosing larger c leads to the RQMC RMSE shrinking faster. We next want to see how large $c \in (0, 1)$ can be chosen and still ensure CLT (24) or AVCI (40).

As will be shown in Sections 5.1–5.6, the corollaries guaranteeing CLT (24) or AVCI (40) in Sections 3 and 4 will typically lead to imposing restrictions on c of the form

$$c < c_k(\alpha_*) \quad (50)$$

for some $0 < c_k(\alpha_*) \leq 1$ depending on the particular Corollary k considered. (The only exceptions to constraints as in (50) are Corollaries 5 and 8, which essentially have $c = 0$ because they assume that $n^c = m_n = m_0$ is fixed in (35); see Section 5.7 for more details.) We will see that except for $k = 1$, each upper bound $c_k(\alpha_*)$ in (50) is strictly decreasing in α_* . Thus, as α_* gets larger (i.e., better RQMC convergence rate for a single randomization), the choices for c ensuring CLT (24) or AVCI (40) shrink in most cases, so the length $m_n = n^c$ of the low-discrepancy sequence needs to grow more slowly and the number $r_n = n^{1-c}$ of independent randomizations must increase more quickly in n .

The “largest” possible c satisfying (50) is $c = c_k(\alpha_*) - \delta$ for $\delta > 0$ infinitesimally small. For this choice of c , what is the RMSE convergence rate of the RQMC estimator? The exponent of n in (48) then becomes $-[(c_k(\alpha_*) - \delta)(\alpha_* - 1/2) + 1/2] - \epsilon = -v_k(\alpha_*) - t_1(\epsilon, \delta)$, where

$$v_k(\alpha_*) \equiv c_k(\alpha_*) \left(\alpha_* - \frac{1}{2} \right) + \frac{1}{2} \quad (51)$$

and $t_1(\epsilon, \delta) \equiv \epsilon - \delta[\alpha_* - 1/2]$. For each (arbitrarily small) $\epsilon > 0$, we take any $\delta \in (0, \delta_0(\epsilon))$ for $\delta_0(\epsilon) = \epsilon/(\alpha_* - 1/2)$ so that $t_1(\epsilon, \delta) > 0$ under our assumption (14). We then get by (48) that

$$\text{RMSE} [\hat{\mu}_{m_n, r_n}^{\text{RQ}}] = \omega \left(n^{-v_k(\alpha_*) - t_1(\epsilon, \delta)} \right) \quad \text{as } n \rightarrow \infty, \text{ for any } \epsilon > 0 \text{ and any } \delta \in (0, \delta_0(\epsilon)). \quad (52)$$

By (13), we also have that $\sigma_m = O(m^{-\alpha_* + \epsilon'})$ as $m \rightarrow \infty$ for all arbitrarily small $\epsilon' > 0$, so under Assumption 1.B, we get $\sigma_{m_n} = O(n^{-c\alpha_* + \epsilon})$ as $n \rightarrow \infty$ for all $\epsilon > 0$. This then leads to $\text{RMSE} [\hat{\mu}_{m_n, r_n}^{\text{RQ}}] = \sigma_{m_n} / \sqrt{r_n} = O(n^{-[c(\alpha_* - 1/2) + 1/2] + \epsilon})$ as $n \rightarrow \infty$, and using $c = c_k(\alpha_*) - \delta$ results in

$$\text{RMSE} [\hat{\mu}_{m_n, r_n}^{\text{RQ}}] = O \left(n^{-v_k(\alpha_*) + t_2(\epsilon, \delta)} \right) \quad \text{as } n \rightarrow \infty, \text{ for any } \epsilon > 0 \text{ and any } \delta \in (0, \delta_0(\epsilon)), \quad (53)$$

where $t_2(\epsilon, \delta) \equiv \epsilon + \delta[\alpha_* - 1/2] > 0$. Therefore, because $t_1(\epsilon, \delta) > 0$ and $t_2(\epsilon, \delta) > 0$ in (52) and (53) can be made arbitrarily small by taking $\epsilon \rightarrow 0$ (which also ensures $\delta \rightarrow 0$), the optimal rate (ignoring constant coefficients and lower-order terms) at which $\text{RMSE} [\hat{\mu}_{m_n, r_n}^{\text{RQ}}]$ decreases under (50) is about

$$\text{RMSE} [\hat{\mu}_{m_n, r_n}^{\text{RQ}}] \approx n^{-v_k(\alpha_*)} \quad \text{as } n \rightarrow \infty, \quad (54)$$

for $v_k(\alpha_*)$ from (51).

We see by (51) that $v_k(\alpha_*) > 1/2$ under our assumption that $\alpha_* > 1/2$ in (15) because $c_k(\alpha_*) > 0$ always holds in (50). Hence, the optimal rate at which the RQMC estimator’s RMSE shrinks is faster than MC’s rate, which has exponent $v_{\text{MC}} = 1/2$ from (49). For any Corollaries k and k' , (51) also implies that under our assumption (15),

$$v_k(\alpha_*) > v_{k'}(\alpha_*) \quad \text{if and only if} \quad c_k(\alpha_*) > c_{k'}(\alpha_*), \quad (55)$$

so expanding the range of valid values for c in (50) corresponds to better optimal RMSE convergence rate by (54). If the constraint (50) has $c_k(\alpha_*) = 1$, which is the largest possible upper bound for c in Assumption 1.B, then (51) shows that $v_k(\alpha_*) = \alpha_*$; thus in this case, even with multiple randomizations, $\text{RMSE} [\hat{\mu}_{m_n, r_n}^{\text{RQ}}]$ decreases at about the rate of a single randomization of a low-discrepancy sequence of full length $m = n$.

The next few subsections will specialize $c_k(\alpha_*)$ in (50) and $v_k(\alpha_*)$ in (51) for Corollaries $k = 1, 2, 3, 4, 6$, and 7 . We will compare the resulting values analytically and graphically in Sections 5.8 and 5.10, respectively.

5.1 CLT Conditions Under Corollary 1

Corollary 1 ensures CLT (24) under condition (27) and the key Assumption 4. In the setting of Assumption 1.B, note that (27) becomes $n^{1-c}/n^{\lambda'c} \rightarrow \infty$ as $n \rightarrow \infty$ for some $\lambda' > 0$, or equivalently, that $1 - c - \lambda'c > 0$, which reduces to $c < 1/(1 + \lambda')$. But as $\lambda' > 0$ can be chosen arbitrarily small in Corollary 1, we can define $c_1(\alpha_*)$ in the constraint (50) and $v_1(\alpha_*)$ in (51) as

$$c_1(\alpha_*) = 1 \quad \text{and} \quad v_1(\alpha_*) \equiv c_1(\alpha_*) \left(\alpha_* - \frac{1}{2} \right) + \frac{1}{2} = \alpha_*.$$

5.2 CLT Conditions When $V_{\text{HK}}(h) < \infty$ Under Corollary 2

We now derive a constraint on c as in (50) to ensure CLT (24) through Corollary 2, which requires $V_{\text{HK}}(h) < \infty$ (Assumption 3.A) and that $(m_n, r_n) = (n^c, n^{1-c})$ satisfies (28) for some (small) $\lambda \in (0, 1)$. Under Assumption 1.B, the left side of (28) becomes $n^{(1-c)(1-\lambda)} \left(\frac{n^c \sigma_{m_n}}{(c \ln n)^s} \right)^2$. Thus, squaring (47) implies that (28) holds when $(1-c)(1-\lambda) + 2c - 2c\alpha_* > 0$ for some $\lambda \in (0, 1)$. This is equivalent to (for some $\lambda \in (0, 1)$)

$$c < \frac{1-\lambda}{2\alpha_* - 1 - \lambda} \equiv c_2(\alpha_*), \quad \text{where} \quad 0 < c_2(\alpha_*) \leq 1 \quad (56)$$

since $\alpha_* \geq 1$ by (14). If $\alpha_* = 1$, then $c_2(\alpha_*) = 1$, making (56) the weakest possible constraint on c under Assumption 1.B. For $\alpha_* > 1$, we get

$$c_2(\alpha_*) \approx \frac{1}{2\alpha_* - 1} \quad (57)$$

because Corollary 2 allows taking $\lambda \in (0, 1)$ as arbitrarily small in (28) and (56).

By (51), under Corollary 2, the optimal rate at which $\text{RMSE}[\widehat{\mu}_{m_n, r_n}^{\text{RQ}}]$ decreases (as $n \rightarrow \infty$) is about $n^{-v_2(\alpha_*)}$, with

$$v_2(\alpha_*) \equiv c_2(\alpha_*) \left(\alpha_* - \frac{1}{2} \right) + \frac{1}{2} = \frac{2\alpha_* - 1 - \lambda\alpha_*}{2\alpha_* - 1 - \lambda}, \quad \text{where} \quad \frac{1}{2} < v_2(\alpha_*) \leq 1 \quad (58)$$

by (56) and (14), so $v_2(\alpha_*) > v_{\text{MC}} = 1/2$ by (49). If $\alpha_* = 1$, then $v_2(\alpha_*) = 1$. For $\alpha_* > 1$, putting (57) into (58) yields

$$v_2(\alpha_*) \approx 1. \quad (59)$$

5.3 CLT Conditions for Bounded h and Corollary 3

We now derive a constraint on c as in (50) to ensure CLT (24) through Corollary 3, which requires that the integrand h is bounded (Assumption 3.B) and that condition (31) holds. Under Assumption 1.B, (47) implies that $(m_n, r_n) = (n^c, n^{1-c})$ satisfies (31) if $r_n \sigma_{m_n}^2 = \omega(n^{1-c} n^{-2c\alpha_* - 2\epsilon}) \rightarrow \infty$ for all sufficiently small $\epsilon > 0$ as $n \rightarrow \infty$, which is true when $1 - c - 2c\alpha_* > 0$, or equivalently,

$$c < \frac{1}{2\alpha_* + 1} \equiv c_3(\alpha_*), \quad \text{where} \quad 0 < c_3(\alpha_*) < \frac{1}{2} \quad (60)$$

under the assumption that $\alpha_* > 1/2$ in (15). If $\alpha_* \geq 1$, as (14) ensures when $V_{\text{HK}}(h) < \infty$, which is not required by Corollary 3, then $0 < c_3(\alpha_*) \leq 1/3$.

By (51) and (60), the optimal rate at which $\text{RMSE}[\widehat{\mu}_{m_n, r_n}^{\text{RQ}}]$ decreases (as $n \rightarrow \infty$) under Corollary 3 is about $n^{-v_3(\alpha_*)}$ with

$$v_3(\alpha_*) \equiv c_3(\alpha_*) \left(\alpha_* - \frac{1}{2} \right) + \frac{1}{2} = \frac{2\alpha_*}{2\alpha_* + 1}, \quad \text{where} \quad \frac{1}{2} < v_3(\alpha_*) < 1 \quad (61)$$

when $\alpha_* > 1/2$, as assumed in (15). (If $\alpha_* \geq 1$, as in (14), then $2/3 \leq v_3(\alpha_*) < 1$.) Hence, under Corollary 3, the RQMC estimator's RMSE shrinks faster than the MC estimator's RMSE by (49).

5.4 CLT Conditions Under Moment Conditions of Corollary 4

We now derive a constraint on c as in (50) to ensure CLT (24) through Corollary 4, which assumes that (11) (from Assumption 3.C) and (32) both hold for some $b > 0$. In (32) under Assumption 1.B, we have $r_n^{b/(2+b)} = n^{(1-c)b/(2+b)}$ and $\sigma_{m_n} = \omega(n^{-c\alpha_* - \epsilon})$ for any $\epsilon > 0$ by (47). Thus, $(m_n, r_n) = (n^c, n^{1-c})$ satisfies (32) if $r_n^{b/(2+b)} \sigma_{m_n}^2 = \omega\left(n^{(1-c)\frac{b}{2+b} - 2c\alpha_* - 2\epsilon}\right) \rightarrow \infty$ as $n \rightarrow \infty$, which holds when $(1-c)\frac{b}{2+b} - 2c\alpha_* > 0$, or equivalently,

$$c < \frac{1}{2\alpha_*(1 + \frac{2}{b}) + 1} \equiv c_4(\alpha_*, b), \quad \text{where} \quad 0 < c_4(\alpha_*, b) < \frac{1}{2} \quad (62)$$

for each $b > 0$ when $\alpha_* > 1/2$, as assumed in (15). If $\alpha_* \geq 1$, as (14) ensures when $V_{\text{HK}}(h) < \infty$ (Assumption 3.A), which is not required by Corollary 4, then $0 < c_4(\alpha_*, b) \leq 1/3$. In general, for any fixed $\alpha_* > 1/2$, $c_4(\alpha_*, b)$ is strictly increasing in b , so the more absolute central moments of $h(U)$ that are finite (as required by (11)), the larger we can choose c in (62).

By (51), under Corollary 4, the optimal rate at which $\text{RMSE}[\hat{\mu}_{m_n, r_n}^{\text{RQ}}]$ decreases (as $n \rightarrow \infty$) is about $n^{-v_4(\alpha_*, b)}$, where for any $b > 0$,

$$v_4(\alpha_*, b) \equiv c_4(\alpha_*, b) \left(\alpha_* - \frac{1}{2} \right) + \frac{1}{2} = \frac{2\alpha_*(1 + \frac{1}{b})}{2\alpha_*(1 + \frac{2}{b}) + 1}, \quad \text{where} \quad \frac{1}{2} < v_4(\alpha_*, b) < 1 \quad (63)$$

when $\alpha_* > 1/2$, as in (15). Hence, we get that $v_4(\alpha_*, b) > v_{\text{MC}} = 1/2$ by (49) for any $b > 0$. Also, $v_4(\alpha_*, b)$ is strictly increasing in b because $c_4(\alpha_*, b)$ has this property and $\alpha_* > 1/2$, so the optimal convergence rate of $\text{RMSE}[\hat{\mu}_{m_n, r_n}^{\text{RQ}}]$ under Corollary 4 improves as $h(U)$ has more finite absolute central moments.

5.5 AVCI Conditions When $V_{\text{HK}}(h) < \infty$ Under Corollary 6

For $I_{m_n, r_n, \gamma}^{\text{RQ}}$ in (38) to be AVCI (40), we assumed, in addition to the CLT in (24), that (39) holds. Corollary 6 ensures (39) is satisfied when $V_{\text{HK}}(h) < \infty$ (Assumption 3.A) and condition (44) holds, which is the same as the square of CLT condition (28) for $\lambda = 1/2$. In the setting of Assumption 1.B, (28) holds for $\lambda = 1/2$ if c satisfies (56) with $\lambda = 1/2$, so AVCI condition (44) is true when

$$c < \frac{1}{4\alpha_* - 3} \equiv c_6(\alpha_*), \quad \text{where} \quad 0 < c_6(\alpha_*) \leq 1 \quad (64)$$

because $\alpha_* \geq 1$ by (14). Note that $c_6(\alpha_*) = 1$ when $\alpha_* = 1$, and $c_6(\alpha_*) < 1$ for $\alpha_* > 1$.

By (51), under Corollary 6, the optimal rate at which $\text{RMSE}[\hat{\mu}_{m_n, r_n}^{\text{RQ}}]$ decreases (as $n \rightarrow \infty$) is about $n^{-v_6(\alpha_*)}$, with

$$v_6(\alpha_*) \equiv c_6(\alpha_*) \left(\alpha_* - \frac{1}{2} \right) + \frac{1}{2} = \frac{3\alpha_* - 2}{4\alpha_* - 3}, \quad \text{where} \quad \frac{3}{4} < v_6(\alpha_*) \leq 1 \quad (65)$$

since $\alpha_* \geq 1$ by (14). Note that $v_6(\alpha_*) = 1$ when $\alpha_* = 1$, and $v_6(\alpha_*) < 1$ for $\alpha_* > 1$. Also, (65) implies that $v_6(\alpha_*) > v_{\text{MC}} = 1/2$ by (49).

5.6 AVCI Conditions Under Moment Conditions of Corollary 7

By Corollary 7, which does not require $V_{\text{HK}}(h) < \infty$, the combination of (11) holding for $b = 2$ (Assumption 3.C) and condition (46) implies (39), then ensuring AVCI (40). Under Assumption 1.B, we have in (46) that $\sigma_{m_n}^4 = \omega(n^{-4c\alpha_* - 4\epsilon})$ as $n \rightarrow \infty$ for any $\epsilon > 0$ by (47) and $r_n = n^{1-c}$. Thus, (46) holds if $r_n \sigma_{m_n}^4 = \omega(n^{1-c} n^{-4c\alpha_* - 4\epsilon}) \rightarrow \infty$ as $n \rightarrow \infty$, which is true if $1 - c - 4c\alpha_* > 0$, or equivalently,

$$c < \frac{1}{4\alpha_* + 1} \equiv c_7(\alpha_*), \quad \text{where} \quad 0 < c_7(\alpha_*) < \frac{1}{3} \quad (66)$$

when $\alpha_* > 1/2$, as assumed in (15). If $\alpha_* \geq 1$, as (14) ensures when $V_{\text{HK}}(h) < \infty$ (Assumption 3.A) which is not required by Corollary 7, then $0 < c_7(\alpha_*) \leq 1/5$.

By (51), under Corollary 7, the optimal rate at which $\text{RMSE}[\hat{\mu}_{m_n, r_n}^{\text{RQ}}]$ decreases (as $n \rightarrow \infty$) is about $n^{-v_7(\alpha_*)}$, where

$$v_7(\alpha_*) \equiv c_7(\alpha_*) \left(\alpha_* - \frac{1}{2} \right) + \frac{1}{2} = \frac{3\alpha_*}{4\alpha_* + 1}, \quad \text{where} \quad \frac{1}{2} < v_7(\alpha_*) < \frac{3}{4} \quad (67)$$

when $\alpha_* > 1/2$, as assumed in (15). Thus, we have $v_7(\alpha_*) > v_{\text{MC}} = 1/2$ by (49).

5.7 Remarks on the Case $m_n \equiv m_0$ in Corollaries 5 and 8

As noted before after (50), Corollaries 5 and 8 essentially have $c = 0$ as they assume that $n^c = m_n = m_0$ is fixed in (35). Thus, under Corollaries 5 and 8, the exact rate at which $\text{RMSE}[\hat{\mu}_{m_n, r_n}^{\text{RQ}}]$ decreases (as $n \rightarrow \infty$) is $n^{-1/2}$, which can be seen through (51) and (54) by setting $c_k(\alpha_*) = 0$. Hence, the RMSE for this case of RQMC shrinks at the same rate as for MC by (49), but RQMC, which can be viewed as a variance-reduction technique (e.g., see [16, 17]), typically has a smaller leading constant.

5.8 Analytical Comparisons of the $c_k(\alpha_*)$ and the $v_k(\alpha_*)$

We next compare the values of $c_k(\alpha_*)$ and $v_k(\alpha_*)$ for the various values of k from the previous subsections. The case $k = 1$ is omitted from the discussion since Assumption 4 is different from the assumptions used in the other corollaries, combining properties of the RQMC method and the integrand. Assumption 4 may also be difficult to verify in practice. Note nevertheless that $c_1(\alpha_*) \geq c_k(\alpha_*)$ and $v_1(\alpha_*) = \alpha_* \geq v_k(\alpha_*) \forall k \geq 2$. As shown before in Propositions 1 and 2 under Assumption 1.A, we will see trade-offs in the conditions that ensure CLT (24) under Assumption 1.B: stronger conditions on the integrand h (through Assumptions 3.A–3.C) lead to looser constraints on c from larger $c_k(\alpha_*)$ in (50). A similar situation will also hold for guaranteeing AVCI (40), as we saw before in Proposition 3 under Assumption 1.A. Also, when imposing comparable conditions on h , the value of $c_k(\alpha_*)$ is always no larger (and often strictly smaller) to ensure AVCI than for the CLT, so making sure AVCI holds typically requires restricting c more than for a CLT.

5.8.1 $k = 2$ vs. $k = 3$ and $k = 4$

To compare $c_2(\alpha_*)$ in (56) of Corollary 2 with $c_3(\alpha_*)$ from (60) for Corollary 3, we can choose $\lambda \in (0, 1)$ in (56) to be arbitrarily small by Corollary 2. Taking any $\lambda \in (0, 1/\alpha_*)$ leads to $c_2(\alpha_*) > c_3(\alpha_*)$, which can also be seen from (57) and (60); also see Proposition 2. Similarly, the same choices for λ results in $v_2(\alpha_*) > v_3(\alpha_*)$ by (55). But recall that Corollary 2 required that $V_{\text{HK}}(h) < \infty$ (Assumption 3.A),

which is stronger than restricting h to be bounded (Assumption 3.B), as Corollary 3 imposed; see Proposition 1.

We now compare $c_2(\alpha_*)$ in (56) from Corollary 2 to $c_4(\alpha_*, b)$ from (62) of Corollary 4 when $V_{\text{HK}}(h) < \infty$ (Assumption 3.A), as required by Corollary 2 but not by Corollary 4. We can choose $\lambda \in (0, 1)$ in (56) to be arbitrarily small, and we have that $c_2(\alpha_*) > c_4(\alpha_*, b)$ for all $\lambda \in (0, \lambda_0(b))$ with $\lambda_0(b) \equiv (2\alpha_* + b)/(2\alpha_* + \alpha_*b) > 0$. Thus, condition (56) is (substantially) less restrictive on our choices for c than (62) for each $b > 0$. This further implies that $v_2(\alpha_*) > v_4(\alpha_*, b)$ by (55).

5.8.2 $k = 3$ vs. $k = 4$

Observe that $c_3(\alpha_*) > c_4(\alpha_*, b)$ for each $b > 0$ by (60) and (62), so the condition (62) for Corollary 4 restricts our choices for c more than condition (60) from Corollary 3, but (60) was obtained under a stronger assumption (h is bounded, i.e., Assumption 3.B) than requiring that condition (11) holds for some $b > 0$ (Assumption 3.C), used to get (62) for Corollary 4. Because $c_4(\alpha_*, b)$ is strictly increasing in b , the constraint on c from condition (62) loosens as b grows. As the condition (11) of Corollary 4 stipulates that the order- $(-2 + b)$ absolute central moment of $h(U)$ is finite, we see that the more absolute central moments that $h(U)$ has, the faster the length $m = n^c$ of the low-discrepancy sequence can grow with n , according to (62). Similarly, by (55), the exponent $v_4(\alpha_*, b)$ from (63) governing the optimal rate at which $\text{RMSE}[\hat{\mu}_{m_n, r_n}^{\text{RQ}}]$ decreases as $n \rightarrow \infty$ under Corollary 4 is strictly worse than the exponent $v_3(\alpha_*)$ from (61) under Corollary 3.

Moreover, note that $\lim_{b \rightarrow \infty} c_4(\alpha_*, b) = c_3(\alpha_*)$, the latter being the upper bound for c in (60) when the integrand h is bounded. Similarly, we have that $\lim_{b \rightarrow \infty} v_4(\alpha_*, b) = v_3(\alpha_*)$. Thus, the tradeoffs of the conditions of Corollaries 3 and 4 disappear as $b \rightarrow \infty$. While this did not necessarily have to happen (e.g., consider $h(u) = \Phi^{\leftarrow}(u)$, as in the proof of Proposition 1), it is reasonable: we can think of bounded h as being an extreme special case as $b \rightarrow \infty$ of the order- $(-2 + b)$ absolute central moment of $h(U)$ being finite.

5.8.3 $k = 6$ vs. $k = 2$ and $k = 7$

We first compare $c_6(\alpha_*)$ from (64) and $c_7(\alpha_*)$ in (66), each of which is an upper bound for c to ensure AVCI (40). It is clear that $c_6(\alpha_*) > c_7(\alpha_*)$, so condition (64), obtained under $V_{\text{HK}}(h) < \infty$ (Assumption 3.A), is a strictly weaker restriction on the choice of c than the condition (66), derived under moment conditions (Assumption 3.C) but without requiring $V_{\text{HK}}(h) < \infty$; also see Proposition 3. Thus, assuming $V_{\text{HK}}(h) < \infty$ allows us to expand the values of c that ensure AVCI. Moreover, we have that $v_6(\alpha_*) > v_7(\alpha_*)$ by (55), so when $V_{\text{HK}}(h) < \infty$, the rate at which $\text{Var}[\hat{\mu}_{m_n, r_n}^{\text{RQ}}]$ converges is faster by choosing c to optimally satisfy Corollary 6 rather than Corollary 7.

We now want to see if the upper bound $c_6(\alpha_*)$ in AVCI condition (64) is more restrictive than the upper bound $c_2(\alpha_*)$ from CLT condition (56), where both were obtained under the assumption that $V_{\text{HK}}(h) < \infty$. When $\alpha_* = 1$, we have $c_6(\alpha_*) = c_2(\alpha_*) = 1$. For $\alpha_* > 1$, we can choose $\lambda \in (0, 1)$ in (56) to be arbitrarily small, and taking any $\lambda \in (0, 1/2)$ leads to $c_6(\alpha_*) < c_2(\alpha_*)$, which can also be seen by comparing (57) and (64). Thus, when $\alpha_* = 1$, the same values of c guarantee both CLT (24) and AVCI (40). But for $\alpha_* > 1$, ensuring AVCI (40) requires restricting c more than what is needed to make sure CLT (24) holds.

5.8.4 $k = 7$ vs. $k = 3$ and $k = 4$

Note that $c_7(\alpha_*) < c_3(\alpha_*)$ for all $\alpha_* > 0$ by (60) and (66), so AVCI condition (66) obtained from Corollary 7 restricts our choices for c more than the CLT condition (60) under Corollary 3. Hence, ensuring AVCI (40) under Corollary 7 requires further constraining c compared to obtaining CLT (24) under Corollary 3 (bounded h). Also, by (55), it follows that $v_7(\alpha_*) < v_3(\alpha_*)$.

By Corollary 4, CLT (24) alternatively holds when (62) and (11) hold for some $b > 0$, as discussed in Section 5.4. To compare the AVCI upper bound $c_7(\alpha_*)$ in (66) to CLT bound $c_4(\alpha_*)$ in (62), which has $c_4(\alpha_*, b) = 1/[1 + 2\alpha_*(1 + \frac{2}{b})]$, note that $c_4(\alpha_*, b) \leq c_7(\alpha_*)$ if and only if $b \leq 2$. Thus, if we select $c < c_4(\alpha_*, b)$ for some $b \leq 2$ to ensure the CLT (24) under Corollary 4, then the same c also yields AVCI (40). But if c was originally chosen so that it satisfies the CLT condition (62) for some $b > 2$ but not for $b \leq 2$, then we further need to restrict c so that (66) holds to guarantee AVCI (40). Hence, ensuring both the CLT and AVCI leads to the constraint $c < \min[c_4(\alpha_*, b), c_7(\alpha_*)] \equiv c_{4,7}(\alpha_*, b)$. In this case, the optimal RMSE convergence rate is about $n^{-v_{4,7}(\alpha_*, b)}$ with $v_{4,7}(\alpha_*, b) = \min[v_4(\alpha_*, b), v_7(\alpha_*)]$, where $v_{4,7}(\alpha_*, b) > 1/2$.

5.9 Summary of Results Under Assumption 1.B

In this section, we therefore have proved the following main results.

Proposition 4. *Under Assumption 1.B that $(m_n, r_n) = (n^c, n^{1-c})$ for some $c \in (0, 1)$, define $\alpha_* > 1/2$ as in (13) and (15), which is the exponential decreasing rate of the standard deviation of an estimator from a single randomization; $c_k(\alpha_*)$ as in (50), which is the supremum of values of c (depending on the particular Corollary k considered) to guarantee a CLT and/or AVCI; and $v_k(\alpha_*)$ in (54), which is the corresponding maximum RMSE convergence rate from r_n i.i.d. randomizations. Then the following properties hold:*

- Under Assumption 4, a condition for CLT under Corollary 4 is $c < c_1(\alpha_*) = 1$ for which $v_1(\alpha_*) = \alpha_*$.
- When $V_{\text{HK}}(h) < \infty$, a condition for CLT under Corollary 2 is $c < c_2(\alpha_*) \approx \frac{1}{2\alpha_* - 1}$ for which $v_2(\alpha_*) \approx 1$, and a condition for AVCI under Corollary 6 is $c < c_6(\alpha_*) = \frac{1}{4\alpha_* - 3}$ for which $v_6(\alpha_*) = \frac{3\alpha_* - 2}{4\alpha_* - 3}$.
- When h is bounded, Corollary 3 leads to supremum values $c_3(\alpha_*) = \frac{1}{2\alpha_* + 1}$ and $v_3(\alpha_*) = \frac{2\alpha_*}{2\alpha_* + 1}$.
- Under the finite $2 + b$ absolute central moment condition of Corollary 4, a CLT is obtained for $c < c_4(\alpha_*, b) = \frac{1}{2\alpha_*(1 + \frac{2}{b}) + 1}$ for which $v_4(\alpha_*, b) = \frac{2\alpha_*(1 + \frac{1}{b})}{2\alpha_*(1 + \frac{2}{b}) + 1}$, while AVCI is obtained for $b = 2$ under Corollary 7 for $c < c_7(\alpha_*) = \frac{1}{4\alpha_* + 1}$ for which $v_7(\alpha_*) = \frac{3\alpha_*}{4\alpha_* + 1}$.
- $\forall k \in \{2, 3, 4, 6, 7\}$, $v_k(\alpha_*) > 1/2$, that is, we obtain a CLT with a better convergence rate than for MC.
- The stricter the condition on h in Corollary k , the larger the corresponding $c_k(\alpha_*)$.
- For a particular condition on h , ensuring AVCI (40) has no larger (and often strictly smaller) $c_k(\alpha_*)$ compared to the corresponding condition to secure CLT (24).
- $v_k(\alpha_*) > v_{k'}(\alpha_*)$ if and only if $c_k(\alpha_*) > c_{k'}(\alpha_*)$: the stronger the requirement on the integrand, the faster the convergence rate.

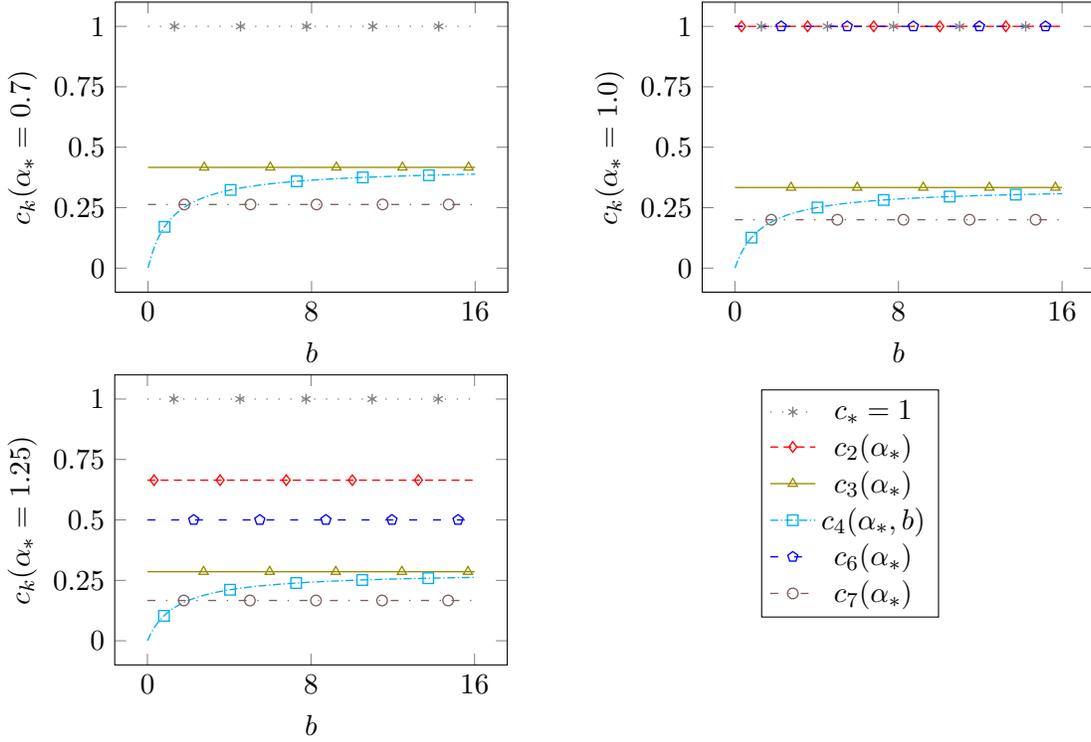


Figure 1: Plots of the upper bounds $c_k(\alpha_*)$ in (50) of c in Assumption 1.B for different Corollaries k . The plots display the $c_k(\alpha_*)$ as functions of b for different fixed values of α_* . The upper left panel does not include $c_2(\alpha_*)$ and $c_6(\alpha_*)$ because these require $V_{\text{HK}}(h) < \infty$, which then implies $\alpha_* \geq 1$ by (14). The plots show that stronger restrictions on the integrand h lead to larger $c_k(\alpha_*)$.

5.10 Graphical Comparisons of the $c_k(\alpha_*)$ and the $v_k(\alpha_*)$

Figure 1 shows the upper bounds $c_k(\alpha_*)$ in (50) for c in Assumption 1.B, where $c_k(\alpha_*)$ is given by $c_2(\alpha_*)$ in (56) for Corollary 2 (with $\lambda = 0.01$), $c_3(\alpha_*)$ in (60) for Corollary 3, $c_4(\alpha_*, b)$ in (62) for Corollary 4, $c_6(\alpha_*)$ in (64) for Corollary 6, and $c_7(\alpha_*)$ in (66) for Corollary 7. The plots display, for various fixed values of α_* , the upper bounds as functions of $b > 0$ from Assumption 3.C. We further plot the upper bound $c_* = 1$ for reference because Assumption 1.B requires $c \in (0, 1)$, which is also seen through (56), (60), (62), (64), and (66). Note that it also corresponds to $c_1(\alpha_*)$. We do not include $v_1(\alpha_*) = \alpha_*$ in the graphs to better see the differences between the other corollaries.

Figure 1 corroborates the results obtained in Section 5.8 and summarized in Proposition 4. We can readily check that:

- $c_4(\alpha_*, b) < c_3(\alpha_*) < c_2(\alpha_*)$ ($c_2(\alpha_*)$ being valid only when $\alpha_* \geq 1$), illustrating that the stricter the condition of the integrand h (see Proposition 1), the larger the possible value of c to ensure CLT (24) (Proposition 2).

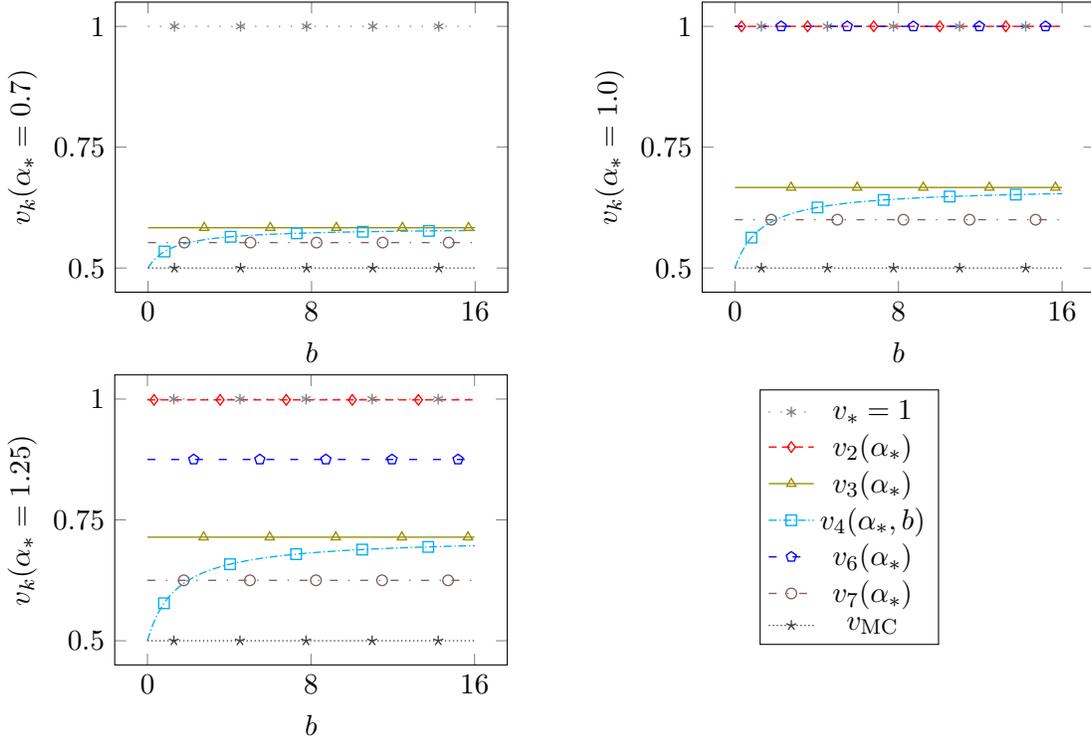


Figure 2: Plots of the negative exponent $v_k(\alpha_*)$ of the optimal rate at which the estimator RMSE decreases as functions of b for different values of α_* . The upper left panel does not include $v_2(\alpha_*)$ and $v_6(\alpha_*)$ because these require $V_{\text{HK}}(h) < \infty$, which then implies $\alpha_* \geq 1$ by (14). The plots show that stronger restrictions on the integrand h lead to larger $v_k(\alpha_*)$.

- $c_7(\alpha_*) < c_6(\alpha_*)$, so we similarly see that a stronger condition on h leads to larger range of values of c that ensure AVCI (40); also see Proposition 3.
- $c_4(\alpha_*, b)$ approaches $c_3(\alpha_*)$ as b grows large, which agrees with the principle that finite absolute central moments of order $2 + b$ as $b \rightarrow \infty$ is “close” to meaning a bounded integrand.
- As b increases (i.e., more finite absolute central moments), the upper bound $c_4(\alpha_*, b)$ grows, so more effort can be put on the QMC part (i.e., $m_n = n^c$ can be larger) when establishing a CLT through the moment conditions of Corollary 4.
- Ensuring AVCI (40) often (but not always) entails restricting c more than what guarantees a CLT, which can be seen from $c_7(\alpha_*) < c_3(\alpha_*)$ and $c_6(\alpha_*) \leq c_2(\alpha_*)$.

Figure 2 plots, as functions of $b > 0$ from Assumption 3.C, the (negative) exponent $v_k(\alpha_*)$ from (51) of the optimal rate at which the RQMC estimator’s RMSE decreases, given by $v_2(\alpha_*)$ in (58) for Corollary 2, $v_3(\alpha_*)$ in (61) for Corollary 3, $v_4(\alpha_*, b)$ in (63) for Corollary 4, $v_6(\alpha_*)$ in (65) for Corollary 6, and $v_7(\alpha_*)$ in (67) for Corollary 7. By (58), (61), (63), (65), and (67), each of these

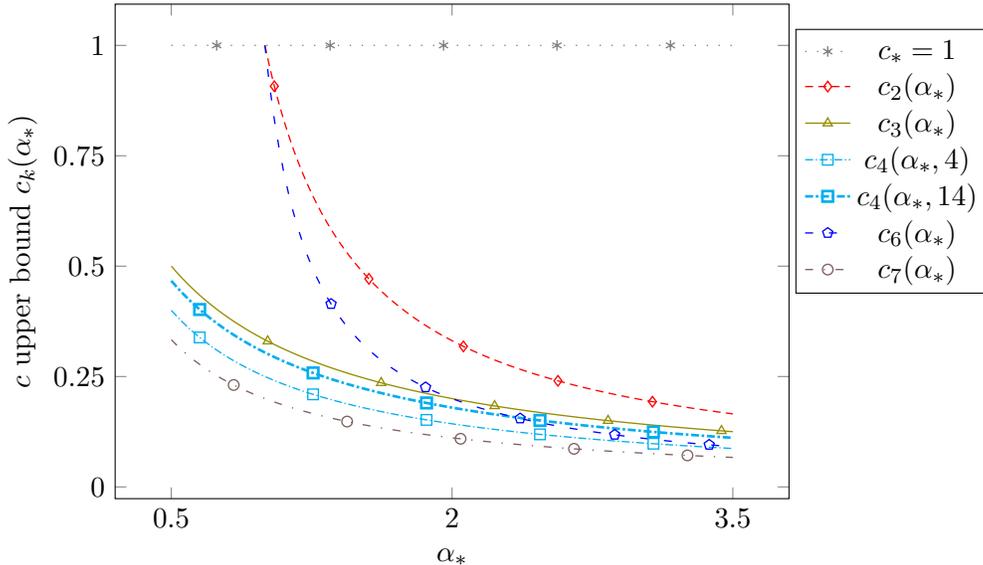


Figure 3: Plots of the upper bounds $c_k(\alpha_*)$ of c as functions of α_* . Functions $c_2(\alpha_*)$ and $c_6(\alpha_*)$ require $V_{\text{HK}}(h) < \infty$, so they are shown for only $\alpha_* \geq 1$ because of (14).

exponents is at most $v_* = 1$, which is also plotted for reference. The figure further includes $v_{\text{MC}} = 1/2$ from (49) for comparison, and all $v_k(\alpha_*)$ are strictly greater than v_{MC} , so RQMC always has that the optimal rate at which its RMSE decreases is better than for MC. By (55), the comparisons and ordering of RMSE rate exponents $v_k(\alpha_*)$ are the same as the ones obtained before for the $c_k(\alpha_*)$.

Comparing across the different panels in Figure 1, we see that each upper bound $c_k(\alpha_*)$ in (50) on c decreases as α_* increases. To investigate this further, Figures 3 and 4 plot the $c_k(\alpha_*)$ and $v_k(\alpha_*)$, respectively, as functions of α_* . We can then see the differences between the various corollaries and assumptions as the QMC method improves (i.e., α_* increases).

Figure 3 more clearly illustrates that as α_* grows, each upper bound $c_k(\alpha_*)$ in (50) on c decreases, so guaranteeing a CLT (24) or AVCI (40) requires putting more effort on the MC part (i.e., $r_n = n^{1-c}$ grows as c decreases) and correspondingly less on the QMC (i.e., $m_n = n^c$ shrinks as c gets smaller). The tradeoff could potentially harm RQMC's optimal RMSE convergence rate because of the diminished benefits (decreasing $c_k(\alpha_*)$) of QMC's improved convergence rates from larger α_* . However, Figure 4 shows that this is not the case for all CLT and most AVCI conditions: the RMSE optimal convergence speed determined by $v_k(\alpha_*)$, as in (51), is generally increasing in α_* . The one exception is $v_6(\alpha_*)$, which determines the optimal RMSE convergence rate to ensure AVCI when $V_{\text{HK}}(h) < \infty$. In this case, Figure 3 shows that as α_* increases, $c_6(\alpha_*)$ drops off quickly, so the number of randomizations must grow rapidly as α_* increases to ensure AVCI when $V_{\text{HK}}(h) < \infty$. But even so, we still have $v_6(\alpha_*) > v_7(\alpha_*)$ for all α_* , where $v_7(\alpha_*)$ in (67) is the optimal RMSE rate exponent for AVCI obtained under the moment conditions of Corollary 7.

6 Numerical results

The goal of this section is to illustrate the accuracy of the results presented in this paper. We aim to see

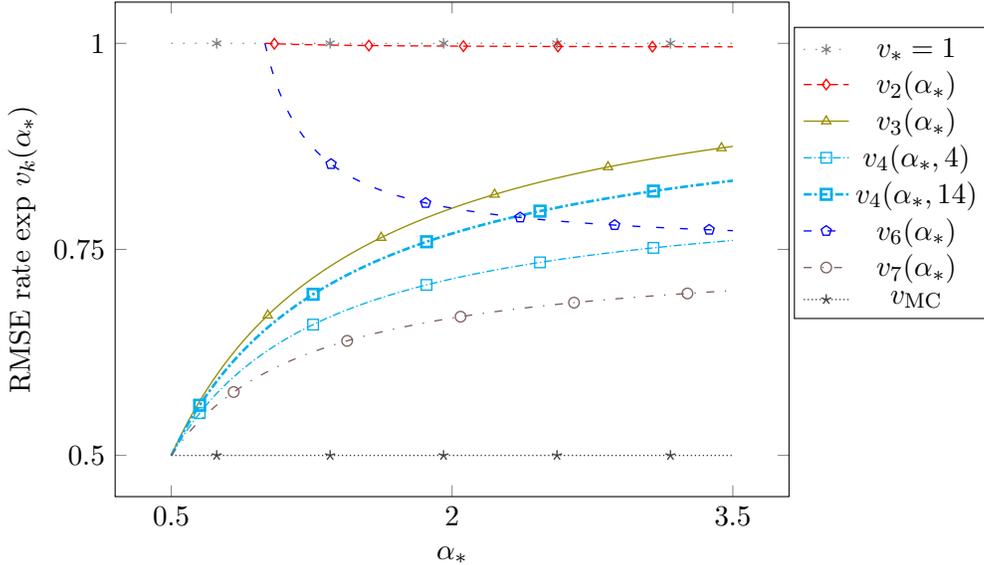


Figure 4: Plots of the negative exponent $v_k(\alpha_*)$ of the optimal rate at which the estimator RMSE decreases as functions of α_* . Functions $v_2(\alpha_*)$ and $v_6(\alpha_*)$ require $V_{\text{HK}}(h) < \infty$, so they are shown for only $\alpha_* \geq 1$ because of (14).

for various values of $c \in (0, 1)$ from Assumption 1.B if a CLT seems to actually hold as $n \rightarrow \infty$, where n is the total number of integrand evaluations. As with any empirical study of asymptotic behavior, our analysis encounters inherent limitations because we can check for approximate normality for only finite values of n . But in spite of this, looking at a range of c in $(0, 1)$ can help to see how tight our conditions are to guarantee convergence to a normal distribution. Indeed, our corollaries provide only *sufficient* conditions, and it may be possible that values of c larger than a particular $c_k(\alpha_*)$ from Section 5 still lead to a CLT. We will investigate this point.

We implemented our experiments in python, with the RQMC sequences generated using the QMCPy library [5]. We consider three different randomized sequences: Sobol' sequences with digital shift (DS) [16], scrambled Sobol' sequence with linear matrix scrambling (LMS) [13, 24], and Lattice rules with random shift (RS) [7, 36, 17]. In comparing the three RQMC methods, we take m_n to be various powers of 2 as the lengths of the RQMC sequence to benefit from the digital net structure of Sobol' sequences.

Recall that Proposition 1 established a strict ordering of the restrictions in Assumption 3 on the integrand h . We now consider three functions associated with the three cases of Assumption 3.

1. For $u = (u_1, \dots, u_s) \in [0, 1]^s$, function $h_v(u) = \prod_{l=1}^s (u_l - \frac{1}{2})^2$ is of bounded Hardy-Krause variation (as a product of functions of bounded variation [3, Section 9.3]), so Assumption 3.A holds.
2. Consider the unidimensional function $g_b(u_1) = \sin(1/u_1)$ for $u_1 \in (0, 1]$ and $g_b(0) = 0$. This function is known to have $V_{\text{HK}}(g_b) = \infty$ and to be bounded. The product function for $u = (u_1, \dots, u_s) \in [0, 1]^s$

$$h_b(u) = \prod_{l=1}^s g_b(u_l)$$

also have $V_{\text{HK}}(h_b) = \infty$ but is bounded, so Assumption 3.B holds but not Assumption 3.A.

3. Consider now $g_{\text{ub},\theta}(u_1) = u_1^{-\theta}$ for $u_1 \in (0, 1]$ and $g_{\text{ub},\theta}(0) = 0$, with $\theta > 0$. Define

$$h_{\text{ub},\theta}(u) = \prod_{l=1}^s g_{\text{ub},\theta}(u_l)$$

for $u \in [0, 1]^s$. Function $h_{\text{ub},\theta}$ is unbounded and the moment of order $2 + b$ of $h_{\text{ub},\theta}(U)$ for $U \sim U[0, 1]^s$ is $\mathbb{E}[(h_{\text{ub},\theta}(U))^{2+b}] = \int_{[0,1]^s} (h_{\text{ub},\theta}(u))^{2+b} du = \left[\int_0^1 (u_1)^{-\theta(2+b)} du_1 \right]^s$, which is finite if and only if $\theta(2 + b) < 1$. In this case, Assumption 3.C holds but not Assumption 3.B.

To test whether a CLT roughly holds, we generated 100 independent values of $\hat{\mu}_{m_n, r_n}^{\text{RQ}}$. We then applied a statistical test on the resulting sample to test for normality. Among existing tests, we chose to use the Shapiro-Wilk test [33], which is specifically designed for checking the goodness-of-fit of a normal distribution to data and is known to have the best power for a given significance among such tests [34]. The null hypothesis H_0 of our test is that the population is normally distributed, and we provide the resulting p -value: a very low p -value indicates that the normal property (and the CLT) can clearly be rejected.

6.1 Estimation of α_*

For checking if the values $c_k(\alpha_*)$ from Section 5 provide appropriate thresholds on values of c under Assumption 1.B to secure a CLT, we need to estimate α_* in (13). A standard procedure for convergence rate estimation of QMC and RQMC methods applies log-log regression. Assuming $\sigma_m \approx \beta m^{-\alpha_*}$ is equivalent to

$$\ln(\sigma_m) \approx \ln(\beta) - \alpha_* \ln(m).$$

To estimate the unknowns β and α_* , we generated data for K values $m^{(i)}$, $1 \leq i \leq K$, of m . For each i , we estimate $\sigma_{m^{(i)}}$ through the sample standard deviation of r_0 independent estimates of μ from single randomizations of $m^{(i)}$ points. Then, using the simplifying notation $\ell_i = \ln(m^{(i)})$ and $\nu_i = \ln(\sigma_{m^{(i)}})$ $\forall 1 \leq i \leq K$, standard linear regression yields an estimator of α_* as

$$\hat{\alpha}_* = - \frac{\sum_{i=1}^K (\nu_i - \hat{\nu}_K)(\ell_i - \bar{\ell}_K)}{\sum_{i=1}^K (\ell_i - \bar{\ell}_K)^2} \quad (68)$$

where

$$\hat{\nu}_K = \frac{1}{K} \sum_{i=1}^K \nu_i \quad \text{and} \quad \bar{\ell}_K = \frac{1}{K} \sum_{i=1}^K \ell_i.$$

For our estimations, we used $r_0 = 100$, $K = 17$, and $m^{(i)} = 2^{i+5}$ for $i \in \{1, \dots, 17\}$. Table 1 gives the estimated α_* for the three considered RQMC methods (Sobol' sequence with DS, Sobol' sequence with LMS, and Lattice with RS) and the three integrands (h_v , h_b and $h_{\text{ub},0.35}$) in dimensions $s = 3$ and $s = 4$. We can see estimated values of α_* larger than 1, the rate in Koksma-Hlawka bound, for function h_v which is of bounded Hardy-Krause variation. For functions h_b and $h_{\text{ub},0.35}$, which have infinite Hardy-Krause variation, estimated values of α_* are smaller than 1 but larger than 1/2 (see (15)), the exponential rate when using MC methods.

Integrand	Dimension s	Sobol' DS	Sobol' LMS	Lattice RS
h_v	3	1.3420	1.3040	1.4218
h_b	3	0.6843	0.6735	0.6173
$h_{ub,0.35}$	3	0.6103	0.6217	0.5612
h_v	4	1.2620	1.1282	1.1528
h_b	4	0.6856	0.6412	0.6035
$h_{ub,0.35}$	4	0.6115	0.5624	0.5437

Table 1: Value $\hat{\alpha}_*$ estimated from a log-log regression

6.2 Analysis for a fixed value of n

In this subsection, we set the dimension as $s = 4$. Recall that n is the total number of integrand evaluations, which is distributed among $r_n = n^{1-c}$ independent randomizations of $m_n = n^c$ points from a low-discrepancy sequence. We proceed as follows: we fix $n = 2^{14} = 16384$, and choose $(m_n, r_n) = (2^t, 2^{14-t})$ for $t \in \{2, \dots, 12\}$ to study a wide range of values of $c = t/14$. Figure 5 displays, as c increases, the p -values of the Shapiro-Wilk test for h_v , h_b and $h_{ub,0.35}$ and the three types of randomization. Recall that the p -values are random values, with $\mathcal{U}[0, 1]$ distribution under H_0 , and a very small value provides strong evidence that the data come from a non-normal distribution.

- For function h_v , the p -values are reasonably high (and plausibly uniform), up to $c \approx 0.8$ for Sobol' with DS and LMS. For the randomly shifted lattice, we do not see any suspect behavior for all considered c . Theoretically, as h_v satisfies Assumption 3.A, Corollaries 2 and 6 apply; thus, taking $\alpha_* = 1.35$ as an approximation from Table 1, we get $c_2(\alpha_*) \approx 0.588$ and $c_6(\alpha_*) \approx 0.417$, smaller than the 0.8 from the numerical experiments.
- For function h_b , which is bounded (so Assumption 3.B holds) but has infinite Hardy-Krause variation, the Shapiro-Wilk test rejects H_0 only once (at the 5% significance level), for Sobol' with LMS when $c = 9/14$, but not for larger values of c , so it seems reasonable to attribute the one small value to regular statistical variability. Theoretically, as Assumption 3.B holds, Corollaries 3 and 7 apply, and taking $\alpha_* = 0.65$ (approximately what Table 1 shows) results in $c_3(\alpha_*) \approx 0.435$ and $c_7(\alpha_*) \approx 0.278$ by (60) and (66). Using a larger value of n would be of interest, but the computations already take hours for $n = 2^{14}$.
- For function $h_{ub,0.35}$, p -values are generally very low, except for some c for the lattices with RS. This generally indicates non-normal behavior. The function $h_{ub,0.35}$ is unbounded but with a finite absolute central moment of order $2 + b$ for $0.35(2 + b) < 1$, that is, $b < 1/0.35 - 2 \approx 0.857$, so Assumption 3.C holds for these values of b . Corollary 4 then applies, and using $\alpha_* = 0.5437$ from Table 1 (for Lattice with RS but using other values do not significantly change the results) gives $c_4(0.5437, 1/0.35 - 2) \approx 0.216$ from (62). This seems corroborated by the observed low p -values for most values of c . (As Assumption 3.C does not hold for $b = 2$, Corollary 7 is not applicable.)

As further illustration, we provide the Q-Q plots in Figure 6 for several values of c when $n = 2^{14}$ for function $h_{ub,0.35}$ and lattice points with random shift. A Q-Q plot compares two probability distributions by plotting their quantiles against each other. The values correspond to the 100 estimates used for the Shapiro-Wilk tests with p -values displayed in Figure 5. In Figure 6, all but the upper left

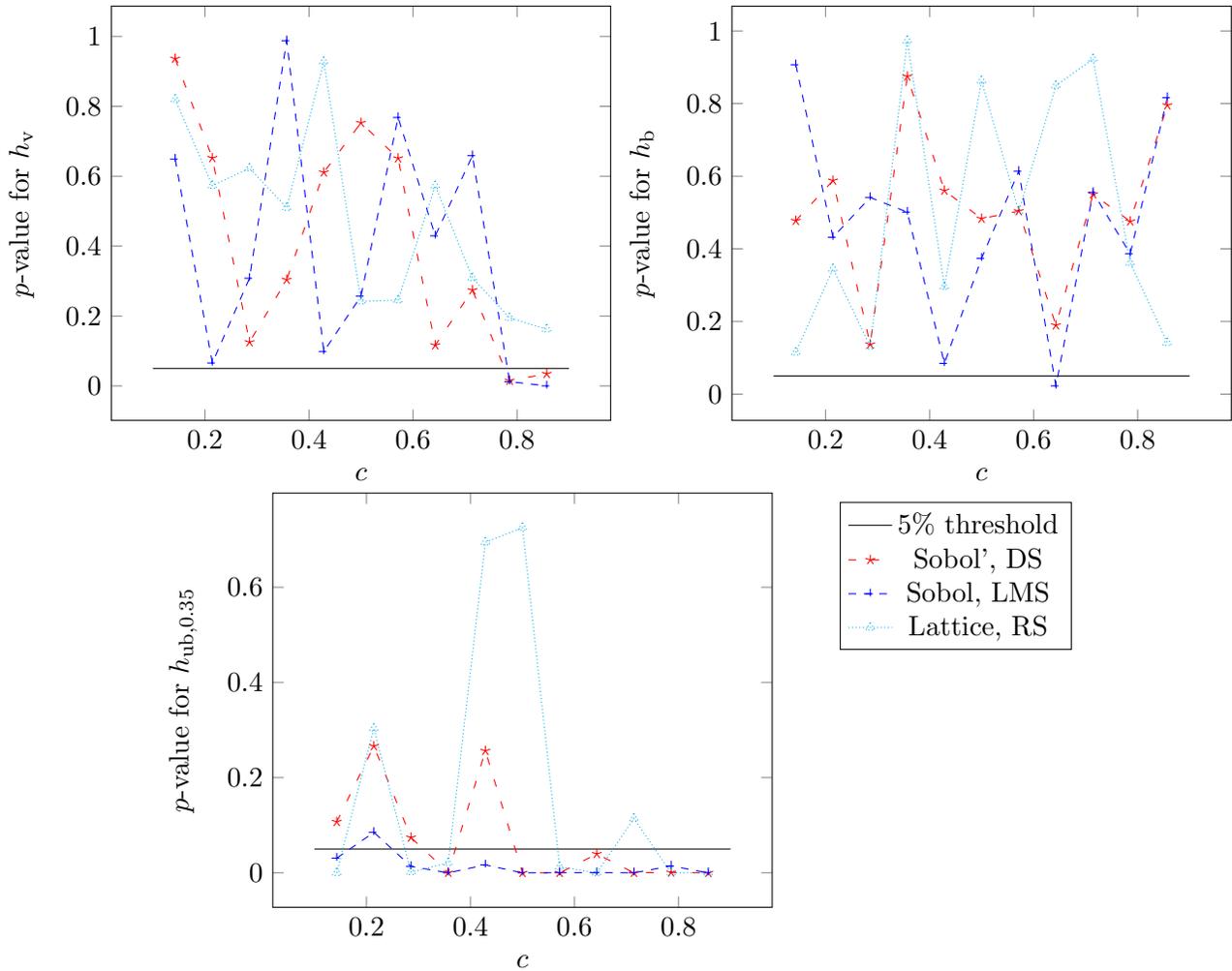
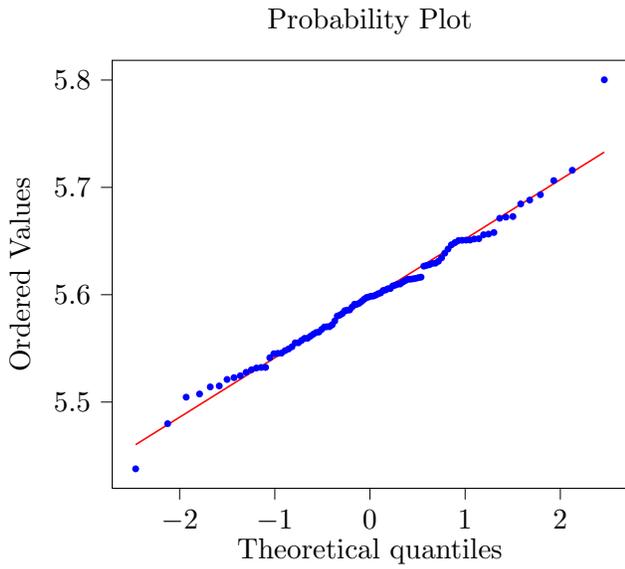
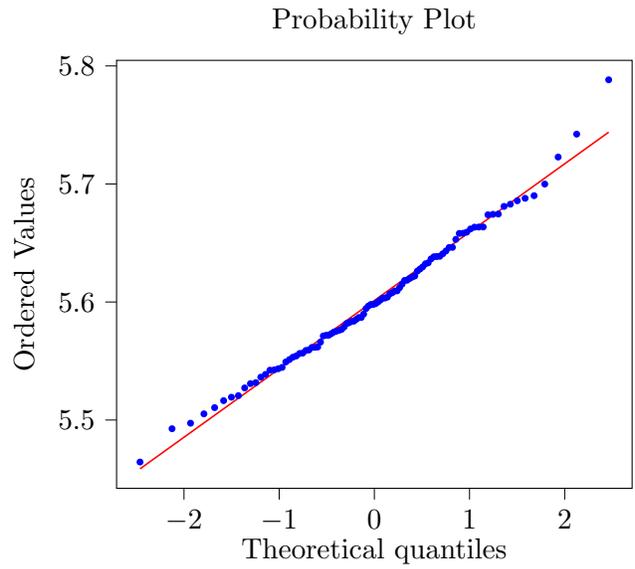


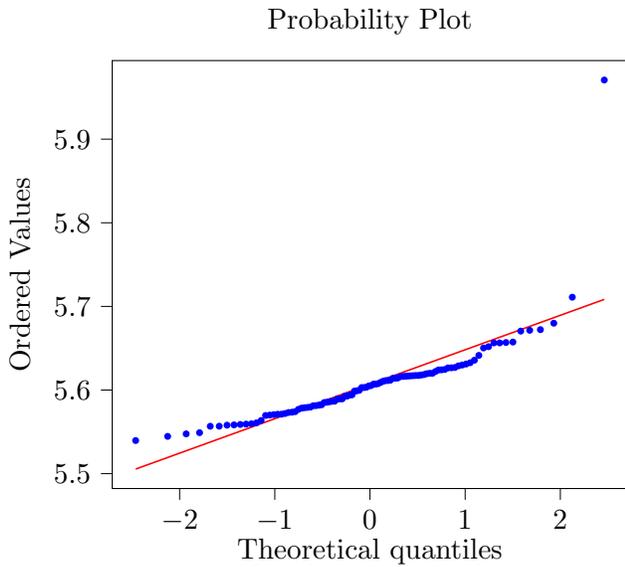
Figure 5: p -values of the Shapiro-Wilk test for various values of $c = t/14$ for $2 \leq t \leq 12$ and for three RQMC techniques with $n = 2^{14}$, from 100 independent estimations. The considered integrand is h_v (top left), h_b (top right) and $h_{ub,0.35}$ (bottom left), each for dimension $s = 4$.



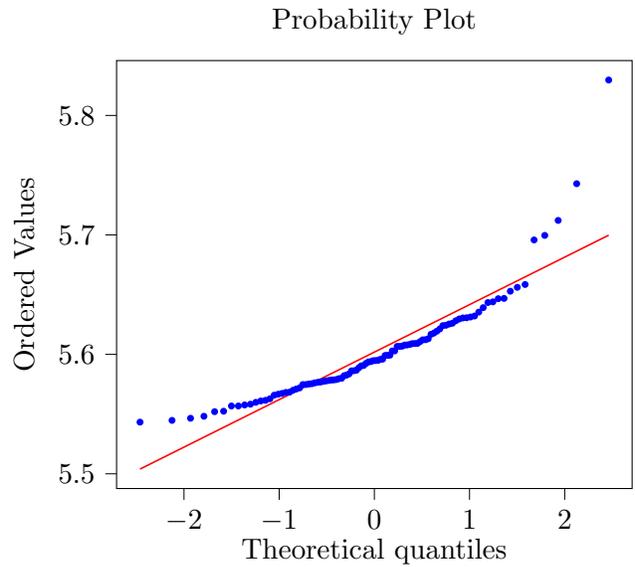
(a) $c = 3/14$



(b) $c = 7/14$



(c) $c = 11/14$



(d) $c = 12/14$

Figure 6: Q-Q plots for $h_{ub,0.35}$ in dimension 4, with Lattice points, random shift, and $n = 2^{14}$

panel correspond to when the Shapiro-Wilk test rejects H_0 . The curvature in the Q-Q plots, especially for large c , indicate positive skewness, so the data have a heavy right tail.

7 Conclusions

RQMC methods are powerful simulation methods accelerating the convergence rate with respect to MC. A standard way to estimate the error in practice is by applying a CLT over r independent randomizations, but typically limiting this number so that more weight can be put on the low-discrepancy size m to gain from the superior convergence rate of QMC. The only existing CLT result as m increases is for scrambled digital nets with nested uniform scrambling, but this RQMC technique is not practical. It was proved that otherwise, increasing m for any fixed number of randomizations may lead to a non-normal distribution. To our knowledge, no theoretical result has ever been published in the literature guaranteeing in general a CLT. We have provided sufficient conditions on (m, r) and their relative increase under Lindeberg's Condition framework for triangular arrays. The conditions depend on the properties of the integrand and the convergence speed of the RQMC standard deviation from a single randomization. We have also given conditions for AVCI, when the standard deviation is estimated. We have presented several properties over the conditions and the convergence speed of the resulting estimators.

In the framework of Assumption 1.B (i.e., $(m_n, r_n) = (n^c, n^{1-c})$) in Section 5.1, Corollary 1 establishes a CLT for any $c \in (0, 1)$ when Assumption 4 holds (the worst-case error for a single randomization and the standard deviation converge at the same rate). We are currently investigating alternatives to Assumption 4. As directions for other future research, we also plan to provide a guide for practitioners on how to choose under Assumption 1.B a value of c as large as possible to satisfy a CLT. As this may entail estimating α_* in (13), which is the exponent defining the rate at which the standard deviation decreases for a single randomization, we would need to account for the statistical error in our estimate of α_* . Given that we provide *sufficient* conditions, we also aim to see if they can be weakened. More numerical investigations can also be worthwhile towards this goal.

Acknowledgment

The authors would like to thank Art Owen for pointing out references, used in the proof of Proposition 1, showing that if the integrand h has bounded Hardy-Krause variation, then h is bounded.

References

- [1] S. Asmussen and P. Glynn. *Stochastic Simulation: Algorithms and Analysis*. Springer, New York, 2007.
- [2] P. Billingsley. *Probability and Measure*. John Wiley and Sons, New York, 3rd edition, 1995.
- [3] S. Breneis. Functions of bounded variation in one and multiple dimensions. Master's thesis, Johannes Kepler Universitat Linz, Austria, 2020.
- [4] G. Casella and R. L. Berger. *Statistical Inference*. Duxbury, Pacific Grove, Calif., 2nd edition, 2002.

- [5] S.-C. T. Choi, F. J. Hickernell, M. McCourt, and A. Sorokin. QMCPy: A quasi-Monte Carlo Python library, 2020+.
- [6] K. L. Chung. *A Course in Probability Theory*. Academic Press, San Diego, third edition, 2001.
- [7] R. Cranley and T. N. L. Patterson. Randomization of number theoretic methods for multiple integration. *SIAM Journal on Numerical Analysis*, 13(6):904–914, 1976.
- [8] H. Damerdj. Strong consistency of the variance estimator in steady-state simulation output analysis. *Mathematics of Operations Research*, 19:494–512, 1994.
- [9] R. Durrett. *Probability: Theory and Examples*. Cambridge University Press, Cambridge, UK, 5th edition, 2019.
- [10] P. W. Glynn. Limit theorems for the method of replication. *Stochastic Models*, 3(3):343–355, 1987.
- [11] F. J. Hickernell. A generalized discrepancy and quadrature error bound. *Mathematics of Computation*, 67(221):299–322, 1998.
- [12] F. J. Hickernell. Error analysis for quasi-Monte Carlo methods. In P. W. Glynn and A. B. Owen, editors, *Monte Carlo and Quasi-Monte Carlo Methods: MCQMC 2016*, pages 3–27, Berlin, 2018. Springer.
- [13] H. S. Hong and F. H. Hickernell. Algorithm 823: Implementing scrambled digital sequences. *ACM Transactions on Mathematical Software*, 29:95–109, 2003.
- [14] D. P. Kroese, T. Taimre, and Z. I. Botev. *Handbook of Monte Carlo Methods*. John Wiley and Sons, New York, 2011.
- [15] L. Kuipers and H. Niederreiter. *Uniform Distribution of Sequences*. John Wiley, New York, 1974.
- [16] P. L’Ecuyer. Randomized quasi-Monte Carlo: An introduction for practitioners. In P. W. Glynn and A. B. Owen, editors, *Monte Carlo and Quasi-Monte Carlo Methods: MCQMC 2016*, pages 29–52, Berlin, 2018. Springer.
- [17] P. L’Ecuyer and C. Lemieux. Variance reduction via lattice rules. *Management Science*, 46(9):1214–1235, 2000.
- [18] P. L’Ecuyer and C. Lemieux. Recent advances in randomized quasi-Monte Carlo methods. In M. Dror, P. L’Ecuyer, and F. Szidarovszky, editors, *Modeling Uncertainty: An Examination of Stochastic Theory, Methods, and Applications*, pages 419–474. Kluwer Academic, Boston, 2002.
- [19] P. L’Ecuyer, D. Munger, and B. Tuffin. On the distribution of integration error by randomly-shifted lattice rules. *Electronic Journal of Statistics*, 4:950–993, 2010.
- [20] E. L. Lehmann. *Elements of Large-Sample Theory*. Springer, New York, 1999.
- [21] C. Lemieux. Quasi-random number techniques. In S. G. Henderson and B. L. Nelson, editors, *Simulation*, Handbooks in Operations Research and Management Science, pages 351–380. Elsevier, Amsterdam, The Netherlands, 2006. Chapter 12.

- [22] C. Lemieux. *Monte Carlo and Quasi-Monte Carlo Sampling*. Series in Statistics. Springer, New York, 2009.
- [23] W.-L. Loh. On the asymptotic distribution of scrambled net quadrature. *Annals of Statistics*, 31(4):1282–1324, 2003.
- [24] J. Matousěk. On the L_2 -discrepancy for anchored boxes. *J. of Complexity*, 14:527–556, 1998.
- [25] M. K. Nakayama and B. Tuffin. Sufficient conditions for a central limit theorem to assess the error of randomized quasi-Monte Carlo methods. Submitted to the 2021 Winter Simulation Conference.
- [26] H. Niederreiter. *Random Number Generation and Quasi-Monte Carlo Methods*, volume 63. SIAM, Philadelphia, 1992.
- [27] A. B. Owen. Randomly permuted (t,m,s)-nets and (t,s)-sequences. In *Monte Carlo and Quasi-Monte Carlo Methods in Scientific Computing: Lecture Notes in Statistics*, volume 106, pages 299–317. Springer, 1995.
- [28] A. B. Owen. Monte Carlo variance of scrambled net quadrature. *SIAM Journal of Numerical Analysis*, 34:1884–1910, 1997.
- [29] A. B. Owen. Scrambled net variance for integrals of smooth functions. *Annals of Statistics*, 25(4):1541–1562, 1997.
- [30] A.B. Owen. Multidimensional variation for quasi-monte carlo. In *Contemporary Multivariate Analysis and Design of Experiments*, pages 49–74. World Scientific, 2005.
- [31] A.B. Owen and D. Rudolf. A strong law of large numbers for scrambled net integration. <https://arxiv.org/abs/2002.07859>, 2020.
- [32] R. J. Serfling. *Approximation Theorems of Mathematical Statistics*. John Wiley and Sons, New York, 1980.
- [33] S. S. Shapiro and M. B. Wilk. An analysis of variance test for normality (complete samples). *Biometrika*, 52(3/4):591–611, Dec. 1965.
- [34] S. S. Shapiro, M. B. Wilk, and H. J. Chen. A comparative study of various tests of normality. *Journal of the American Statistical Association*, 63:1343–1372, 1968.
- [35] B. Tuffin. Variance reduction applied to product-form multi-class queuing networks. *ACM Transactions on Modeling and Computer Simulation*, 7(4):478–500, 1997.
- [36] B. Tuffin. Variance reduction order using good lattice points in Monte Carlo methods. *Computing*, 61(4):371–378, 1998.
- [37] B. Tuffin. Randomization of quasi-monte carlo methods for error estimation: Survey and normal approximation. *Monte Carlo Methods and Applications*, 10(3-4):617–628, 2004.