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

Marvin K. Nakayama
Department of Computer Science, New Jersey Institute of Technology, Newark, NJ 07102, marvin@njit.edu

Bruno Tuffin
Inria, Univ Rennes, CNRS, IRISA, Campus de Beaulieu, 263 Avenue Général Leclerc, 35042 Rennes, FRANCE, bruno.tuffin@inria.fr

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 computationally expensive 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.

Key words: Randomized quasi-Monte Carlo, central limit theorems, confidence intervals

History:

---

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., Asmussen and Glynn (2007), Kroese et al. (2011) 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 (Lemieux 2009, Niederreiter 1992) replace the random sampling of (standard) 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 can 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 (Hickernell 2018), the most famous being the Koksma-Hlawka inequality (Niederreiter 1992). 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., (Tuffin 2004, 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 $r \to \infty$. For tutorials on RQMC, see among others, L'Ecuyer (2018), L'Ecuyer and Lemieux (2002), Lemieux (2009), and Tuffin (2004). Several types of randomization exist, the main ones being the *random shift* (Cranley and Patterson 1976), which translates all points of the low-discrepancy sequence by the *same random vector*; the *digital shift* (L'Ecuyer 2018), 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*, which randomly permutes the digits (Owen 1997a,b).

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 taking m to be as large as possible to get a smaller error. But this heuristic generally lacks a theoretical basis. For example, L'Ecuyer et al. (2010) establish that for a single (i.e., $r = 1$) random shift of a lattice, the limiting error distribution as $m \to \infty$ 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 by Loh (2003) (as $m \to \infty$ for fixed $r \geq 1$) in the case of digital scrambling of a particular type called *nested uniform scrambling*. But this scrambling is computationally expensive, which has limited its adoption in practice.

The question we thus aim at answering here is, 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 (Billingsley 1995, Theorem 27.2). More precisely, we provide sufficient conditions on the relative increase of m and r under various alternative 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 AVCI conditions when the variance is estimated. Comparing our conditions on the integrand shows 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 $m \to \infty$, we further verify that under our assumptions, RQMC always outperforms MC in terms of convergence speed of the estimators' *root-mean-square errors* (RMSEs).

The rest of the paper unfolds as follows. Section 2 presents the notation 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 (Billingsley 1995, 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 show that stronger conditions on the integrand allow more weight to 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 provides concluding remarks. An e-companion contains all proofs (Section EC.1), as well as additional analysis when $(m,r) = (n^c, n^{1-c})$ (Sections EC.2 and EC.3) and numerical results (Section EC.4). In Nakayama and Tuffin (2021), we present many of the theorems, corollaries and propositions (but all without proofs), as well as Figures EC.1, EC.3, and EC.4, which are in Section EC.3 of the current paper.

2. Notation/Framework

Our goal is to estimate

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

where $h : [0,1]^s \to \mathbb{R}$ 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, and means of many stochastic models may be expressed in this way, e.g., through a change of variables. 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 (standard) MC estimator of μ is $\widehat{\mu}_n^{\mathrm{MC}} = \frac{1}{n}\sum_{i=1}^n h(U_i)$, with U_i, $i = 1, 2, \ldots, n$, as *independent and identically distributed (i.i.d.)* $\mathcal{U}[0,1]^s$ random vectors. Suppose that $\psi^2 \equiv \mathrm{Var}[h(U)] \in (0, \infty)$, where $\mathrm{Var}[\cdot]$ denotes the variance operator. The MC estimator's root-mean-square error then satisfies

$$\mathrm{RMSE}\left[\widehat{\mu}_n^{\mathrm{MC}}\right] = \frac{\psi}{\sqrt{n}} \tag{1}$$

because $\widehat{\mu}_n^{\mathrm{MC}}$ is unbiased (i.e., $\mathbb{E}[\widehat{\mu}_n^{\mathrm{MC}}] = \mu$), as will be true for all estimators of μ that we consider. Moreover, the MC estimator $\widehat{\mu}_n^{\mathrm{MC}}$ obeys a (Gaussian) CLT (e.g., Billingsley 1995, Theorem 27.1): $\sqrt{n}[\widehat{\mu}_n^{\mathrm{MC}} - \mu]/\psi \Rightarrow \mathcal{N}(0,1)$ as $n \to \infty$, where $\Rightarrow$ denotes weak convergence, and $\mathcal{N}(a, d^2)$ is a normal random variable with mean a and variance d^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 $I_{n,\gamma}^{\mathrm{MC}} \equiv [\widehat{\mu}_n^{\mathrm{MC}} \pm z_\gamma \widehat{\psi}_n/\sqrt{n}]$, where $\widehat{\psi}_n^2 = [1/(n-1)]\sum_{i=1}[h(U_i) - \widehat{\mu}_n^{\mathrm{MC}}]^2$ 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 $I_{n,\gamma}^{\mathrm{MC}}$ is an asymptotically valid CI for μ in the sense that $\lim_{n\to\infty} \mathbb{P}(\mu \in I_{n,\gamma}^{\mathrm{MC}}) = \gamma$ (Casella and Berger 2002, 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* (Lemieux 2009, Chapter 5)), leading QMC to approximate μ via $\widehat{\mu}_n^{\mathrm{Q}} = \frac{1}{n}\sum_{i=1}^n h(\xi_i)$ based on n evaluations of the integrand h. The *Koksma-Hlawka inequality* (e.g., (Lemieux 2009, Section 5.6.1)) provides an error bound

$$|\widehat{\mu}_n^{\mathrm{Q}} - \mu| \leq V_{\mathrm{HK}}(h) D_n^*(\Xi) \tag{2}$$

for all n, where $V_{\mathrm{HK}}(h)$ 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 $V_{\mathrm{HK}}(h) \geq 0$ and $0 \leq D_n^*(\Xi) \leq 1$. In (2), $V_{\mathrm{HK}}(h)$ measures the "roughness" of h, while $D_n^*(\Xi)$ quantifies the "non-uniformity" of Ξ. We typically have

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

where $f(n) = O(g(n))$ as $n \to \infty$ 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 $V_{\mathrm{HK}}(h) < \infty$, putting (3) into (2) shows that $|\widehat{\mu}_n^{\mathrm{Q}} - \mu| = O(n^{-1}(\ln n)^s)$, so the QMC error has an asymptotically faster convergence rate than $\widehat{\mu}_n^{\mathrm{MC}}$ does (as measured, e.g., by the MC estimator's RMSE in (1)). But we have no practical error bound for QMC because computing $V_{\mathrm{HK}}(h)$ is at least as difficult as computing μ and is the sum of $2^s - 1$ Vitali variations (Niederreiter 1992), 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, Hickernell (1998, 2018), Kuipers and Niederreiter (1974), Lemieux (2006), Niederreiter (1992)), but each encounters similar computational issues.

2.3. Randomized Quasi-Monte Carlo

RQMC randomizes the low-discrepancy sequence, 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 constructed from Ξ, 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 of the original sequence. RQMC repeats this $r \geq 1$ i.i.d. times, computing an estimator from each randomization. Specifically, let $U_{i,j}'$ be the i-th point of the j-th randomization ($i \geq 1$ and $1 \leq j \leq r$). The RQMC estimator is then

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

with X_j as the estimator from randomization $j = 1, 2, \ldots, r$, of m points. (Section 2.4 discusses how (m, r) are chosen.) The independence across the r randomizations ensures that X_j, $j = 1, 2, \ldots, r$, are i.i.d. From their sample variance $\widehat{\sigma}^2_{m,r} = \sum_{j=1}^{r}(X_j - \widehat{\mu}^{\mathrm{RQ}}_{m,r})^2/(r-1)$ when $r \geq 2$, we then obtain a potential γ-level CI for μ as $I^{\mathrm{RQ}}_{m,r,\gamma} \equiv [\widehat{\mu}^{\mathrm{RQ}}_{m,r} \pm z_\gamma \widehat{\sigma}_{m,r}]$. The hope is that as m or r (or both) grows large, the overall RQMC estimator $\widehat{\mu}^{\mathrm{RQ}}_{m,r}$ obeys a Gaussian CLT (see Section 3) and $I^{\mathrm{RQ}}_{m,r,\gamma}$ is an AVCI (see Section 4). We next describe some possible randomizations and 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 (Cranley and Patterson 1976). Formally, using the first m points of the low-discrepancy sequence Ξ and $r \geq 1$ independent randomizations leads to

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

in (4), where across randomizations, $U_1, U_2, \ldots, U_r$ 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 $\langle U_j + \xi_i \rangle$, $i = 1, 2, \ldots, m$, 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

$$D^*_m(\Xi_{U_j}) \leq 4^s D^*_m(\Xi), \tag{6}$$

as shown in (Tuffin 1997, 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

$$\mathrm{RMSE}[X_j] = O(m^{-1}(\ln m)^s) \tag{7}$$

as $m \to \infty$, 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 $m \to \infty$ 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* (Tuffin 1998, L'Ecuyer and Lemieux 2000) for which the random shift preserves the lattice structure. Let $\mathbb{Z}$ denote the set of all integers, and for $g = (g_1, g_2, \ldots, g_s) \in \mathbb{Z}^s$, define $t(g) = \prod_{i=1}^s \max(1, |g_i|)$. For a periodic function $f : \mathbb{R}^s \to \mathbb{R}$ with period 1 over each coordinate, define its Fourier coefficient of rank $g \in \mathbb{Z}^s$ as $\hat{f}(g) = \int_{[0,1]^s} f(x)e^{-(2\pi\sqrt{-1})g\cdot x}\,\mathrm{d}x$, where $v \cdot y = \sum_{i=1}^s v_i y_i$ is the inner product of $v = (v_1, \ldots, v_s) \in \mathbb{R}^s$ and $y \in (y_1, \ldots, y_s) \in \mathbb{R}^s$. For $\alpha > 1$ and $C > 0$, let $E_\alpha^s(C)$ be the set of such periodic functions $f : \mathbb{R}^s \to \mathbb{R}$ for which $|\hat{f}(g)| \leq Ct(g)^{-\alpha}$ 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 $U \sim \mathcal{U}[0,1]^s$,

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

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 $m \to \infty$, as shown for lattices in L'Ecuyer et al. (2010). 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 $r = 1$), L'Ecuyer et al. (2010) argue that characterizing the error distribution is not practical. Thus, for any fixed $r \geq 1$, the limit distribution as $m \to \infty$ is generally non-normal, motivating our goal of defining rules on (m, r) that ensure a Gaussian CLT.

2.3.2. Scrambled Digital Nets Owen (1995, 1997a) introduces *scrambled digital nets* as another form of RQMC. The method scrambles the digits of special low-discrepancy sequences, namely, digital nets in base b_0 (Niederreiter 1992). The approach 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 Niederreiter 1992 for details), but rather focus on the randomization. For a digital net $\Xi = (\xi_i)_{i\geq 1}$ in base b_0 and dimension s, we express the k-th coordinate $(1 \leq k \leq s)$ of the i-th point $\xi_i = (\xi_i^{(1)}, \ldots, \xi_i^{(s)})$ as $\xi_i^{(k)} = \sum_{l=1}^\infty \xi_i^{(k,\ell)} b_0^{-\ell}$ with each $\xi_i^{(k,l)} \in \{0, \ldots, b_0 - 1\}$.

If we let $U_i' = (U_i^{(1),\prime}, \ldots, U_i^{(s),\prime})$ 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

$$U_i^{(k),\prime} = \sum_{\ell=1}^{\infty} U_i^{(k,\ell),\prime} b_0^{-\ell}, \quad \text{with} \quad U_i^{(k,\ell),\prime} = \begin{cases} \pi_0^k(\xi_i^{(k,1)}) & \text{if } \ell = 1 \\ \pi^k_{0,U_i^{(k,1),\prime},\ldots,U_i^{(k,\ell-1),\prime}}(\xi_i^{(k,\ell)}) & \text{if } \ell > 1 \end{cases},$$

where π_0^k, $\pi_0, \pi^k_{0,U_i^{(k,1),\prime},\ldots,U_i^{(k,\ell-1),\prime}}$, $\forall k, \ell$, are independent and uniformly distributed on the set of $b_0!$ permutations of $\{0, \ldots, b_0 - 1\}$. 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_0 of ξ_i is finite, meaning that the required number of permutations is finite, even if large. Other more computationally efficient forms of scrambling appear in Hong and Hickernell (2003), L'Ecuyer (2018), and Matousěk (1998), including the *linear matrix scramble*.

Scrambled digital nets keep the discrepancy property of the original digital net. Specifically, let Ξ_Π denote the scrambling of a digital net Ξ in base b_0. 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 (L'Ecuyer 2018, Owen 1995)

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

so the estimator X_j in (4) from a single randomization converges at the same speed as (7) when the integrand h has bounded Hardy-Krause variation. For special classes of h (Owen 1997b), the RMSE of the quadrature rule based on nested uniform scrambling can be as small as $O\left(m^{-3/2}(\ln m)^{(s-1)/2}\right)$. For RQMC via digital nets and Owen's full nested scrambling, Loh (2003) establishes a Gaussian CLT for $r = 1$ as $m \to \infty$ (so the number of involved independent permutations also tends to infinity). But the CLT of Loh (2003) is restricted to this specific scrambling, whose large computational cost has limited its use in practice; 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_0 and dimension

s, and recall the notation $\xi_i = (\xi_i^{(1)}, \ldots, \xi_i^{(s)})$ for the i-th point, whose k-th coordinate $(1 \le k \le s)$ is $\xi_i^{(k)} = \sum_{\ell=1}^{\infty} \xi_i^{(k,\ell)} b_0^{-\ell}$ with each $\xi_i^{(k,l)} \in \{0, \ldots, b_0 - 1\}$. The j-th randomization employs a uniform $U_j = (U_j^{(1)}, \ldots, U_j^{(s)}) \sim \mathcal{U}[0,1]^s$, and we write the development in base b_0 of its k-th coordinate as $U_j^{(k)} = \sum_{\ell=1}^{\infty} U_j^{(k,\ell)} b_0^{-\ell}$ (where each $U_j^{(k,\ell)} \in \{0, \ldots, b_0 - 1\}$). For the i-th randomized point $U'_{i,j} = (U_{i,j}^{(1),\prime}, \ldots, U_{i,j}^{(s),\prime})$ from the j-th randomization, its k-th coordinate is defined by

$$U_{i,j}^{(k),\prime} = \sum_{\ell=1}^{\infty} U_{i,j}^{(k,\ell),\prime} b_0^{-\ell}, \qquad \text{where} \qquad U_{i,j}^{(k,\ell),\prime} = \sum_{\ell=1}^{\infty} \left[(\xi_i^{(k,\ell)} + U_j^{(k,\ell)}) \bmod b_0 \right] b_0^{-\ell}.$$

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 (L'Ecuyer 2018). Specifically, for a digital net Ξ, let $\Xi_{U_j}^{\mathrm{Dig}}$ be its digital shift based on $U_j \sim \mathcal{U}[0,1]^s$. Then there exists a constant $C > 0$ such that when $D_m^*(\Xi) \le C m^{-1} (\ln m)^s$ for all m, we also have

$$D_m^*(\Xi_{U_j}^{\mathrm{Dig}}) \le C m^{-1} (\ln m)^s. \tag{9}$$

2.4. Assumptions and Preliminary Results

For a given computation budget of about n integrand evaluations, we define 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 $n \to \infty$, we take $r \equiv r_n \ge 1$ and $m \equiv m_n \ge 1$ as functions of n satisfying the following:

Assumption 1.A. *$m_n r_n \le n$ for each $n \ge 1$, with $m_n r_n / n \to 1$ and $r_n \to \infty$ as $n \to \infty$.*

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

$$\widehat{\mu}_{m_n, r_n}^{\mathrm{RQ}} = \frac{1}{r_n} \sum_{j=1}^{r_n} X_{n,j}, \qquad \text{where} \qquad X_{n,j} = \frac{1}{m_n} \sum_{i=1}^{m_n} h(U'_{i,j}), \tag{10}$$

so $X_{n,j}$ is the estimator from randomization $j = 1, 2, \ldots, r_n$, of m_n points, where $m_n \le n$. Section 3's goal is to determine conditions on the behavior of *allocation* (m_n, r_n) as n grows that ensure $\widehat{\mu}_{m_n, r_n}^{\mathrm{RQ}}$

obeys a CLT (with Gaussian limit) as $n \to \infty$, and Section 4 derives conditions for an asymptotically valid CI for μ. Other papers (e.g., Damerdji 1994, Glynn 1987) adopt a framework similar to Assumption 1.A to study MC methods for analyzing steady-state behavior of a stochastic model.

Assumption 1.A requires $r_n \to \infty$ as $n \to \infty$ because otherwise, a Gaussian CLT may not hold. As noted earlier, L'Ecuyer et al. (2010) show 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 $m \to \infty$ for fixed $r \geq 1$ (see the discussion at the end of Section 2.3.1), and the only known Gaussian CLT being for the (computationally expensive) nested digital scrambling (Loh 2003) (see the end of Section 2.3.2). But while Assumption 1.A specifies that $r_n \to \infty$ as $n \to \infty$, it does not require that $m_n \to \infty$ as $n \to \infty$, 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 $c \in (0,1)$.*

We now give some remarks on Assumption 1.B.

- Under Assumption 1.B, both $r_n, m_n \to \infty$ as $n \to \infty$ because $c \in (0,1)$.
- As m_n and r_n need to be integers, Assumption 1.B 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. Moreover, while our asymptotic analyses also will hold for $(m_n, r_n) = (d_1 n^c, d_2 n^{1-c})$ for positive constants d_1 and d_2, we simplify the discussion by assuming that $(m_n, r_n) = (n^c, n^{1-c})$ is integer-valued under Assumption 1.B.
- Assumption 1.B precludes allocations such as $(m_n, r_n) = (n/\ln n, \ln n)$ allowed by 1.A.

For those $c \in (0,1)$ ensuring that $\widehat{\mu}^{\mathrm{RQ}}_{m_n,r_n}$ satisfies a Gaussian CLT or that AVCI holds, Section 5 determines the optimal c for the best possible convergence rate for $\mathrm{RMSE}[\widehat{\mu}^{\mathrm{RQ}}_{m_n,r_n}]$ as $n \to \infty$.

Let Ξ' be a generic randomized low-discrepancy sequence. It is $\Xi' = \Xi_{U_j}$ for a random shift (Section 2.3.1), $\Xi' = \Xi_{\Pi}$ for digital scrambling (Section 2.3.2), and $\Xi' = \Xi^{\mathrm{Dig}}_{U_j}$ for a digital shift (Section 2.3.2). As we will explain below, each of these randomization methods satisfy the following condition, which will be helpful to analyze RQMC estimators with integrands h having $V_{\mathrm{HK}}(h) < \infty$.

Assumption 2. *For the RQMC method used, there exists a constant* $0<w_0'<\infty$ *such that each randomized sequence* Ξ' *satisfies* $D_m^*(\Xi')\le w_0' m^{-1}(\ln m)^s$ *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).*

We next explain why the RQMC methods in Section 2.3 satisfy Assumption 2, which does not depend on a particular allocation (m_n, r_n), as in Assumption 1.A or 1.B. By (3), there exists a constant $0 < w_0 < \infty$ such that $D_m^*(\Xi) \le w_0 m^{-1}(\ln m)^s$ 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 often consider four alternative conditions on h, presented in order of decreasing strength (see Proposition 1 below).

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

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

Assumption 3.C. *There exists* $b>0$ *such that for* $U\sim\mathcal{U}[0,1]^s$,

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

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 limits the roughness of h. 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. None of 3.A, 3.B, 3.C and 3.D depends on the allocation (m_n, r_n) nor the randomization method.

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.*

We next give two bounds on absolute central moments of the estimator $X_{n,1}$ in (10) based on m_n points from a single randomization, which will be useful when establishing a CLT or AVCI. We omit the proof of the following, which follows by an argument analogous to (Tuffin 1997, Theorem 2).

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

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

By (12), when $V_{\mathrm{HK}}(h) < \infty$ and $m_n \to \infty$ as $n \to \infty$, the order-q absolute central moment of $X_{n,j}$ in (10) shrinks as $O([(\ln m_n)^s / m_n]^q)$ as $n \to \infty$, so $\eta_{n,q} = O(m_n^{-q+\epsilon})$ for each $\epsilon > 0$. But assuming $V_{\mathrm{HK}}(h) < \infty$ is restrictive; e.g., the proof (in Section EC.1) of Proposition 1 notes that $V_{\mathrm{HK}}(h) = \infty$ for $s \geq 2$ if h is an indicator function (so μ is a probability) with discontinuities not aligned with the coordinate axes. If $V_{\mathrm{HK}}(h) < \infty$ is not true or cannot be verified, we can still bound $\eta_{n,q}$ as follows under a moment condition on $h(U)$. (The proof appears in Section EC.1.)

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.*

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 $V_{\mathrm{HK}}(h) < \infty$ and $m_n \to \infty$ as $n \to \infty$, we have that

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

as $n \to \infty$ 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 does not satisfy Assumption 1.B, and then a Gaussian CLT may not be guaranteed (L'Ecuyer et al. 2010), as noted earlier.

The following sections establish various conditions that ensure $\widehat{\mu}^{\mathrm{RQ}}_{m_n,r_n}$ obeys a CLT and when we can obtain an AVCI 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

In analyzing $\widehat{\mu}^{\mathrm{RQ}}_{m_n,r_n}$ in (10) as $n \to \infty$, the distribution of the averaged terms $X_{n,1}, X_{n,2}, \ldots, X_{n,r_n}$ changes with n. A theoretical framework for handling this under Assumption 1.A uses that the $(X_{n,j})_{n=1,2,\ldots;j=1,2,\ldots,r_n}$ in (10) form a *triangular array* (Billingsley 1995, p. 359). In a triangular array, also called a *double array* (e.g., Serfling 1980, Section 1.9.3), 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

$$\sigma^2_{n,j} = \mathrm{Var}[X_{n,j}] \equiv \sigma^2_{m_n},$$

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

$$X_{n,1}, X_{n,2}, \ldots, X_{n,r_n} \text{ are i.i.d., each with some distribution } F_n. \tag{13}$$

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

Although (13) has that the r_n random variables are i.i.d. for each n, the general setting for a triangular array, as in (Billingsley 1995, 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 (Billingsley 1995, Theorem 27.2) for triangular arrays:

For each n, assume that $X_{n,j}$, $j = 1, \ldots, r_n$, are independent (but not necessarily identically distributed), with $E[X_{n,j}] = \mu_{n,j}$ and $\mathrm{Var}[X_{n,j}] = \sigma^2_{n,j} < \infty$, and let $s_n^2 = \sigma^2_{n,1} + \cdots + \sigma^2_{n,r_n}$. Let $G_{n,j}$ denote the distribution of $Y_{n,j} = X_{n,j} - \mu_{n,j}$ and let $\tau^2_{n,j}(t) = \int_{|y|>ts_n} y^2 \,\mathrm{d}G_{n,j}(y)$ 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,

$$\frac{\bar{X}_n - \mu_n}{\sqrt{\mathrm{Var}[\bar{X}_n]}} \Rightarrow \mathcal{N}(0,1), \qquad \text{as } n \to \infty \tag{14}$$

provided that the *Lindeberg condition* holds:

$$\frac{\tau^2_{n,1}(t) + \cdots + \tau^2_{n,r_n}(t)}{\sigma^2_{n,1} + \cdots + \sigma^2_{n,r_n}} \to 0, \qquad \text{as } n \to \infty, \quad \forall t > 0. \tag{15}$$

The Lindeberg condition (15) implies that $\max_{1 \leq j \leq r_n} \sigma^2_{n,j}/s_n^2 \to 0$ as $n \to \infty$ (e.g., Lehmann 1999, 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 precludes the possibility that, e.g., $X_{n,j} \equiv 0$, $2 \le j \le r_n$, so the left side of (14) reduces to $(X_{n,1} - \mu_{n,1})/\sigma_{n,1}$, which can have any distribution with mean 0 and variance 1.

We now adapt the Lindeberg condition (15) to study the RQMC estimator $\widehat{\mu}^{\mathrm{RQ}}_{m_n,r_n}$ in (10). By (13),

$$s_n^2 = \sigma_{n,1}^2 + \cdots + \sigma_{n,r_n}^2 = r_n \sigma_{m_n}^2. \tag{16}$$

Denote the distribution of $Y_{n,j} = X_{n,j} - \mu$ by G_n, which does not depend on j by (13). Note that $\sigma_{m_n}^2 = \int_{y\in\Re} y^2 \,\mathrm{d}G_n(y)$, and let

$$\tau_n^2(t) = \int_{|y|>ts_n} y^2 \,\mathrm{d}G_n(y), \qquad \text{for } t > 0. \tag{17}$$

We will impose another assumption to avoid uninteresting cases. The following precludes the exact result from being eventually always returned by the RQMC estimator.

Assumption 4. *$\sigma_{m_n}^2 > 0$ for all n large enough.*

We omit the proof of the following, which specializes for RQMC the condition (15) as (18) below.

Theorem 1. *Suppose that Assumptions 1.A and 4 hold. If the Lindeberg condition*

$$\frac{\tau_n^2(t)}{\sigma_{m_n}^2} \to 0, \qquad \text{as } n \to \infty, \quad \forall t > 0 \tag{18}$$

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

$$\frac{\widehat{\mu}^{RQ}_{m_n,r_n} - \mu}{\sigma_{m_n}/\sqrt{r_n}} \Rightarrow \mathcal{N}(0,1), \qquad \text{as } n \to \infty. \tag{19}$$

From (Billingsley 1995, p. 361), condition (18) is even necessary and sufficient for the CLT (19) since for all j, $\sigma_{n,j}^2/s_n^2 = 1/r_n \to 0$ as $n \to \infty$ by Assumption 1.A. The Lindeberg condition (18), which imposes restrictions on the tail behavior of G_n through (17), holds under a Lyapounov (moment) condition (Billingsley 1995, Theorem 27.3), which (20) below adapts for our RQMC setting.

Theorem 2. *Suppose that Assumptions 1.A and 4 hold. Further suppose that there exists $b' > 0$ such that $\mathbb{E}\left[|X_{n,1} - \mu|^{2+b'}\right] < \infty$ for each n satisfying Assumption 4 and that*

$$\frac{\mathbb{E}\left[|X_{n,1} - \mu|^{2+b'}\right]}{r_n^{b'/2}\sigma_{m_n}^{2+b'}} \to 0, \qquad \text{as } n \to \infty. \tag{20}$$

Then the Lindeberg condition (18) *and CLT* (19) *hold.*

3.1. Corollaries of Theorems 1 and 2

We now develop various sufficient conditions that secure CLT (19) through Theorem 1 or 2.

Corollary 1. *Suppose that Assumptions 1.A and 4 hold for allocation* (m_n, r_n) *with* $m_n \to \infty$ *as* $n \to \infty$*, and that there are constants* $b' > 0$ *and* $k_1 \equiv k_{1,b'} \in (0, \infty)$ *such that* $\mathbb{E}\left[|A_m - \mu|^{2+b'}\right] < \infty$ *and*

$$\frac{\mathbb{E}\left[|A_m - \mu|^{2+b'}\right]}{\sigma_m^{2+b'}} \le k_1, \quad \text{for all } m \text{ sufficiently large,} \tag{21}$$

where $A_m = \frac{1}{m}\sum_{i=1}^{m} h(U_i')$ *is the estimator based on* m *points from a single randomization* $\Xi' = (U_i')_{i \ge 1}$*, and* $\sigma_m^2 = \mathrm{Var}[A_m]$*. Then the Lyapounov condition* (20) *and CLT* (19) *hold. A sufficient condition for* (21) *is that there exists a constant* $k_2 \in (0, \infty)$ *such that*

$$\mathbb{P}\left(|A_m - \mu| \le k_2 \sigma_m\right) = 1, \quad \text{for all } m \text{ sufficiently large.} \tag{22}$$

Condition (21) (resp., (22)) bounds the moment (resp., almost sure) behavior of the absolute error $|A_m - \mu|$ from a single randomization of m points relative to its standard deviation σ_m for all large m. While Corollary 1 requires $m_n \to \infty$, (21) and (22) do not depend on the allocation (m_n, r_n), so Corollary 1 can fulfill Assumption 1.A with r_n growing slowly to ∞ as $n \to \infty$. Under Assumption 1.B, we may then take $(m_n, r_n) = (n^c, n^{1-c})$ for $c \in (0, 1)$ arbitrarily close to 1, allowing large m_n to exploit QMC's fast convergence rate; Section 5.1 will explore this idea. But Assumption 1.A further permits, e.g., $(m_n, r_n) = (\lfloor n/\ln n \rfloor, \lfloor \ln n \rfloor)$, so m_n can increase even more quickly in n.

Establishing (21) or (22) may be difficult, so we next provide other conditions that can be more readily verifiable to ensure CLT (19). 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* ($V_{\mathrm{HK}}(h) < \infty$)*, and 4. Also, suppose that*

$$r_n^{1-\lambda} \left(\frac{m_n \sigma_{m_n}}{(\ln m_n)^s} \right)^2 \to \infty, \qquad \text{as } n \to \infty, \tag{23}$$

for some constant $\lambda \in (0, 1)$*, which can be chosen arbitrarily small. Then the Lyapounov condition* (20) *and CLT* (19) *hold.*

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 4 hold. If*

$$s_n^2 = r_n \sigma_{m_n}^2 \to \infty, \qquad \text{as } n \to \infty, \tag{24}$$

then the Lindeberg condition (18) *and CLT* (19) *hold.*

As with Corollary 2, the next corollary follows from Theorem 2, but it does not require $V_{\mathrm{HK}}(h) < \infty$, 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 ($h(U)$ has finite absolute central moment of order $2+b$ for some $b>0$), and 4 hold. Also, suppose that*

$$r_n^{b/(2+b)} \sigma_{m_n}^2 \to \infty, \qquad \text{as } n \to \infty. \tag{25}$$

Then the Lyapounov condition (20) *and CLT* (19) *hold.*

While Assumption 1.A specifies that $r_n \to \infty$ as $n \to \infty$, it does not require that $m_n \to \infty$ as $n \to \infty$. The previous corollaries allow $m_n \to \infty$ as $n \to \infty$, although this was not required except for Corollary 1. The following result specializes Theorem 1 to consider the case when m_n is fixed.

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

$$(m_n, r_n) = \left(m_0, \left\lfloor \frac{n}{m_0} \right\rfloor\right) \text{ for all } n, \text{ for some fixed } m_0 \geq 1 \text{ with } \sigma_{m_0}^2 > 0, \tag{26}$$

which implies Assumptions 1.A and 4, then the Lindeberg condition (18) *and CLT* (19) *hold.*

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 (19). 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). (Section 5.8 will provide further comparisons under Assumption 1.B.)

Proposition 2. *If Assumption 1.A holds, then condition* (26) *of Corollary 5 implies condition* (25) *of Corollary 4, and* (25) *in turn implies condition* (24) *of Corollary 3. If in addition*

$$\frac{1}{r_n^{\lambda}}\left(\frac{m_n}{(\ln m_n)^s}\right)^2 \to d_0 \in (0,\infty], \qquad \text{as } n\to\infty, \text{ for } \lambda\in(0,1) \text{ for which (24) holds,} \tag{27}$$

then (24) *implies condition* (23) *of Corollary 2. Condition* (27) *holds, e.g., under Assumption 1.B.*

The condition (27) in Proposition 2 specifies that r_n does not grow too quickly (as $n\to\infty$) compared to m_n. While (27) is always true under Assumption 1.B, it does not hold, e.g., for fixed $m_n \equiv m_0$, as in condition (26) 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 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 $X_{n,j}$, $j=1,2,\ldots,r_n$, are i.i.d. by (13), and their sample variance

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

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

$$I^{\mathrm{RQ}}_{m_n,r_n,\gamma} \equiv \left[\widehat{\mu}^{\mathrm{RQ}}_{m_n,r_n} \pm z_\gamma \widehat{\sigma}_{m_n,r_n}/\sqrt{r_n}\right]. \tag{28}$$

Our goal now is to provide conditions ensuring that $I_{n,\gamma}$ is an AVCI, i.e., (30) below holds.

Theorem 3. *Suppose that Assumptions 1.A and 4 hold. If the CLT* (19) *holds and if*

$$\frac{\widehat{\sigma}^2_{m_n,r_n}}{\sigma^2_{m_n}} \Rightarrow 1, \qquad \text{as } n\to\infty, \tag{29}$$

then

$$\lim_{n\to\infty} \mathbb{P}(\mu \in I^{RQ}_{m_n,r_n,\gamma}) = \gamma. \tag{30}$$

We now want a sufficient condition ensuring that (29) holds to secure AVCI (30).

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

$$\frac{\mathbb{E}\left[(X_{n,1}-\mu)^4\right]}{r_n\sigma^4_{m_n}} \to 0, \qquad \text{as } n\to\infty, \tag{31}$$

then the CLT (19), (29), *and AVCI* (30) *hold.*

4.1. Specific Sufficient Conditions for AVCI

We next establish AVCI (30) (often through Theorem 4) under various conditions.

COROLLARY 6. *Suppose that Assumptions 1.A and 4 hold. If* $\mathbb{E}\left[|X_{n,1}-\mu|^4\right]<\infty$ *and* (21) *holds with* $b'=2$*, then the CLT* (19)*,* (29)*, and AVCI* (30) *hold. A sufficient condition to ensure* (21) *with* $b'=2$ *is that* (22) *holds.*

By Corollaries 1 and 6, the condition (22) ensures both the (19) and AVCI (30). But as we will see in Section 5, securing AVCI often uses stronger conditions than ensuring a CLT.

We next separately consider Assumptions 3.A, 3.C, and 3.D to establish AVCI (30).

COROLLARY 7. *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 4. If*

$$r_n\left(\frac{m_n\sigma_{m_n}}{(\ln m_n)^s}\right)^4\to\infty,\qquad \textit{as } n\to\infty, \tag{32}$$

then the CLT (19)*,* (29)*, and AVCI* (30) *hold.*

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

COROLLARY 8. *Suppose that Assumptions 1.A and 4 hold, along with Assumption 3.C (*$h(U)$ *has finite absolute central moment of order* $2+b$*) for* $b=2$*. If*

$$r_n\sigma_{m_n}^4\to\infty,\qquad \textit{as } n\to\infty, \tag{33}$$

then the CLT (19)*,* (29)*, and AVCI* (30) *hold.*

The next result considers the case when $m_n\equiv m_0$ is fixed, as in condition (26) of Corollary 5.

COROLLARY 9. *The same conditions as in Corollary 5 imply that* (29) *and AVCI* (30) *hold.*

Corollaries 5 and 9 assume the same conditions, where the former establishes the CLT (19), and the latter proves that (28) is AVCI (30). Thus, when $m_n=m_0$ is fixed, AVCI does not require any additional conditions beyond that for a CLT. Corollaries 7 and 8 also allow for $m_n=m_0$, but further permit $m_n\to\infty$. But if $m_n=m_0$ is fixed, Corollary 9 ensures AVCI (30) more broadly than Corollaries 7 and 8, as the latter impose additional restrictions on integrand h.

4.2. Remarks on Comparing Conditions for CLT and AVCI

We now compare Corollaries 7, 8, and 9, each of which ensures AVCI (30). By Proposition 1, Assumption 3.A in Corollary 7 is strictly stronger than Assumption 3.C in Corollary 8, and the latter restriction is strictly stronger than Assumption 3.D (Corollary 9). We next compare the corollaries' constraints on (m_n, r_n), showing that the conditions instead weaken from Corollary 7 to 8 to 9.

Proposition 3. *Under Assumption 1.A, condition* (26) *in Corollary 9 implies condition* (33) *of Corollary 8. If also $m_n > 1$ for all n sufficiently large, then* (33) *implies condition* (32) *of Corollary 7.*

Comparing the conditions securing CLT (19) and AVCI (30) provides further insights. Corollary 1 (resp., 6) ensures the CLT (resp., AVCI) when condition (21) holds for $b' > 0$ (resp., $b' = 2$), so the corollaries impose a more stringent condition to ensure AVCI beyond a CLT. But both Corollaries 1 and 6 also guarantee CLT and AVCI under the more restrictive requirement (22). Corollaries 5 and 9 assume the same conditions, so when $m_n = m_0$ is fixed, AVCI (30) does not require any additional conditions beyond that for the CLT (19). Theorem 4 and Corollaries 7 and 8 provide sufficient conditions for AVCI, which also imply CLT (19). 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 7, and compare Corollary 4 to Corollary 8. However, as such a comparison is long, we instead will carry out the analysis only under Assumption 1.B in Section 5.8, 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 (19) or AVCI (30), and which of those c lead to the best rates of convergence for RMSE$[\widehat{\mu}^{\mathrm{RQ}}_{m_n,r_n}]$ as $n \to \infty$. Proposition 4 at the end of this section will summarize the results under Assumption 1.B.

With RQMC, we typically have $\sigma_m \equiv \left(\mathrm{Var}[\frac{1}{m}\sum_{i=1}^{m} h(U'_{i,1})]\right)^{1/2} = O(m^{-\alpha})$ as $m \to \infty$ with $\alpha > 1/2$ (e.g., see (12) when $V_{\mathrm{HK}}(h) < \infty$). 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 $\mathrm{RMSE}[\widehat{\mu}_m^{\mathrm{MC}}] = \sqrt{\mathrm{Var}[\widehat{\mu}_m^{\mathrm{MC}}]} = \psi/\sqrt{m}$ by (1) with $n = m$. Let

$$\alpha_* = -\lim_{m\to\infty} \frac{\ln(\sigma_m)}{\ln(m)}, \tag{34}$$

where we assume the limit exists, so α_* is the only constant such that the rate (as $m\to\infty$) 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.,

$$\sigma_m = o(m^{-\alpha_*+\epsilon}) \quad \text{and} \quad \sigma_m = \omega(m^{-\alpha_*-\epsilon}) \quad \text{as} \quad m\to\infty \quad \text{for any} \quad \epsilon > 0, \tag{35}$$

where $f(m) = o(g(m))$ as $m\to\infty$ means that $f(m)/g(m)\to 0$ as $m\to\infty$, and $f(m) = \omega(g(m))$ as $m\to\infty$ means that $f(m)/g(m)\to\infty$ as $n\to\infty$ (i.e., $g(n) = o(f(n))$). By (12), we see that

$$\alpha_* \geq 1 \qquad \text{for RQMC when } V_{\mathrm{HK}}(h) < \infty \tag{36}$$

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

$$\alpha_* > \frac{1}{2}. \tag{37}$$

The value of α_* depends on the integrand h and the method to construct the randomized 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. (Section EC.4.1 of the e-companion presents results from numerical experiments to estimate α_* for different RQMC methods and various integrands corresponding to Assumptions 3.A, 3.B, and 3.C. In particular, Table EC.1 shows that for all randomization methods considered, (36) seems to hold for the integrands satisfying Assumption 3.A, and the other integrands appear to be consistent with (37).)

For $m_n = n^c$, we get from (35) that (as c is bounded)

$$\sigma_{m_n} = \omega(n^{-c\alpha_*-\epsilon}), \quad \text{as } n\to\infty, \text{ for any } \epsilon > 0. \tag{38}$$

Therefore, using $r_n = n^{1-c}$ and (10), (13), and (38) lead to

$$\mathrm{RMSE}\left[\widehat{\mu}_{m_n,r_n}^{\mathrm{RQ}}\right] = \frac{\sigma_{m_n}}{\sqrt{r_n}} = \omega\left(n^{-v(\alpha_*,c)-\epsilon}\right), \qquad \text{as } n\to\infty, \text{ for any } \epsilon > 0, \tag{39}$$

where $v(\alpha_*, c) \equiv c\left[\alpha_* - \frac{1}{2}\right] + \frac{1}{2}$. Similarly, (35) also yields RMSE $\left[\widehat{\mu}^{\mathrm{RQ}}_{m_n,r_n}\right] = o\left(n^{-v(\alpha_*,c)+\epsilon}\right)$ as $n \to \infty$ for all $\epsilon > 0$. As ϵ can be arbitrarily small, RMSE $\left[\widehat{\mu}^{\mathrm{RQ}}_{m_n,r_n}\right]$ then decreases at about rate (ignoring the leading coefficient and lower-order terms) $n^{-v(\alpha_*,c)}$ as $n \to \infty$. Our assumption (37) 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 $n^{-v_{\mathrm{MC}}}$ as $n \to \infty$, where

$$v_{\mathrm{MC}} \equiv \frac{1}{2} \tag{40}$$

by (1). Moreover, for any fixed α_* satisfying (37), $v(\alpha_*, c)$ 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 (19) or AVCI (30).

As will be shown in Sections 5.1–5.6, the corollaries guaranteeing CLT (19) or AVCI (30) in Sections 3 and 4 will typically lead to imposing restrictions on c of the form

$$c < c_k(\alpha_*) \tag{41}$$

for some $0 < c_k(\alpha_*) \le 1$ depending on the particular Corollary k considered. (The only exceptions to constraints as in (41) are Corollaries 5 and 9, which essentially have $c = 0$ because they assume that $n^c = m_n = m_0$ is fixed in (26); see Section 5.7 for more details.) We will see that except for $k = 1$ and 6, each upper bound $c_k(\alpha_*)$ in (41) 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 (19) or AVCI (30) 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 (41) 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 (39) then becomes $-[(c_k(\alpha_*) - \delta)(\alpha_* - 1/2) + 1/2] - \epsilon = -v_k(\alpha_*) - t_1(\epsilon, \delta)$, where

$$v_k(\alpha_*) \equiv c_k(\alpha_*)\left(\alpha_* - \frac{1}{2}\right) + \frac{1}{2} \tag{42}$$

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 (37). We then get by (39) that

$$\mathrm{RMSE}\left[\widehat{\mu}^{\mathrm{RQ}}_{m_n,r_n}\right] = \omega\left(n^{-v_k(\alpha_*) - t_1(\epsilon,\delta)}\right) \quad \text{as } n \to \infty, \text{ for any } \epsilon > 0 \text{ and any } \delta \in (0, \delta_0(\epsilon)). \tag{43}$$

By (35), we also have that $\sigma_m = o(m^{-\alpha_* + \epsilon'})$ as $m \to \infty$ for all arbitrarily small $\epsilon' > 0$, so under Assumption 1.B, we get $\sigma_{m_n} = o(n^{-c\alpha_* + \epsilon})$ as $n \to \infty$ for all $\epsilon > 0$. This then leads to $\mathrm{RMSE}\left[\widehat{\mu}^{\mathrm{RQ}}_{m_n,r_n}\right] = \sigma_{m_n}/\sqrt{r_n} = o(n^{-[c(\alpha_* - 1/2) + 1/2] + \epsilon})$ as $n \to \infty$, and using $c = c_k(\alpha_*) - \delta$ results in

$$\mathrm{RMSE}\left[\widehat{\mu}^{\mathrm{RQ}}_{m_n,r_n}\right] = o\left(n^{-v_k(\alpha_*) + t_2(\epsilon,\delta)}\right) \quad \text{as } n \to \infty, \text{ for any } \epsilon > 0 \text{ and any } \delta \in (0, \delta_0(\epsilon)), \tag{44}$$

where $t_2(\epsilon,\delta) \equiv \epsilon + \delta[\alpha_* - 1/2] > 0$ under our assumption (37). Therefore, because $t_1(\epsilon,\delta) > 0$ and $t_2(\epsilon,\delta) > 0$ in (43) and (44) can be made arbitrarily small by taking $\epsilon \to 0$ (which also ensures $\delta \to 0$), the optimal rate at which $\mathrm{RMSE}\left[\widehat{\mu}^{\mathrm{RQ}}_{m_n,r_n}\right]$ decreases under (41) is about

$$\mathrm{RMSE}\left[\widehat{\mu}^{\mathrm{RQ}}_{m_n,r_n}\right] \approx n^{-v_k(\alpha_*)} \quad \text{as } n \to \infty, \tag{45}$$

for $v_k(\alpha_*)$ from (42).

We see by (42) that $v_k(\alpha_*) > 1/2$ under our assumption that $\alpha_* > 1/2$ in (37) because $c_k(\alpha_*) > 0$ always holds in (41). Hence, the optimal rate at which the RQMC estimator's RMSE shrinks is faster than MC's rate, which has exponent $v_{\mathrm{MC}} = 1/2$ from (40). For any Corollaries k and k', (42) also implies that under our assumption (37),

$$v_k(\alpha_*) > v_{k'}(\alpha_*) \quad \text{if and only if} \quad c_k(\alpha_*) > c_{k'}(\alpha_*), \tag{46}$$

so expanding the range of valid values for c in (41) corresponds to better optimal approximate RMSE convergence rate by (45). If the constraint (41) has $c_k(\alpha_*) = 1$, which is the largest possible upper bound for c in Assumption 1.B, then (42) shows that $v_k(\alpha_*) = \alpha_*$; thus in this case, even with the number of randomizations slowly growing large (i.e., $c < 1$ with $c \approx 1$), $\mathrm{RMSE}\left[\widehat{\mu}^{\mathrm{RQ}}_{m_n,r_n}\right]$ decreases at about the rate for a single randomization of a low-discrepancy sequence of full length $m_n = n$ (or for an RQMC estimator with a *fixed* number $r_0 \geq 1$ of randomizations, each of length $m_n = \lfloor n/r_0 \rfloor$).

The next few subsections will specialize $c_k(\alpha_*)$ in (41) and $v_k(\alpha_*)$ in (42) for Corollaries $k = 1$, 2, 3, 4, 6, 7, and 8. Sections EC.2 and EC.3 of the e-companion compare the resulting values analytically and graphically, respectively.

5.1. CLT and AVCI Conditions Under Corollaries 1 and 6

For $m_n \to \infty$ as $n \to \infty$, Corollary 1 ensures CLT (19) under condition (21) for some $b' > 0$ (used for the order-$(2+b')$ absolute central moment of A_m); Corollary 6 secures AVCI (30) using condition (21) with $b' = 2$. As (21) does not depend on the allocation (m_n, r_n), Corollaries 1 and 6 allow any $c < 1$ in Assumption 1.B, so we define $c_1(\alpha_*)$ and $c_6(\alpha_*)$ in constraint (41) and $v_1(\alpha_*)$ and $v_6(\alpha_*)$ in (42) as

$$c_1(\alpha_*) = c_6(\alpha_*) \equiv 1 \quad \text{and} \quad v_1(\alpha_*) \equiv c_1(\alpha_*)\left(\alpha_* - \frac{1}{2}\right) + \frac{1}{2} = \alpha_* \equiv v_6(\alpha_*). \tag{47}$$

Thus, as noted in the discussion after (46), because $v_1(\alpha_*) = v_6(\alpha_*) = \alpha_*$, the RMSE of $\widehat{\mu}^{\mathrm{RQ}}_{m_n,r_n}$ as $n \to \infty$ with the number r_n of randomizations slowly growing large (i.e., $c < 1$ with $c \approx 1$) decreases at about the same rate as for an RQMC estimator with a *fixed* number $r_0 \geq 1$ of randomizations.

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

We now derive a constraint on c as in (41) to ensure CLT (19) through Corollary 2, which requires $V_{\mathrm{HK}}(h) < \infty$ (Assumption 3.A) and that $(m_n, r_n) = (n^c, n^{1-c})$ satisfies (23) for some (small) $\lambda \in (0,1)$. Under Assumption 1.B, the left side of (23) becomes $n^{(1-c)(1-\lambda)}\left(\frac{n^c \sigma_{m_n}}{(c \ln n)^s}\right)^2$. Thus, squaring (38) implies that (23) holds when $(1-c)(1-\lambda) + 2c - 2c\alpha_* > 0$ for some $\lambda \in (0,1)$. This is equivalent to

$$c < \frac{1-\lambda}{2\alpha_* - 1 - \lambda} \equiv c_2'(\alpha_*, \lambda). \tag{48}$$

Because Corollary 2 assumes that $V_{\mathrm{HK}}(h) < \infty$, (36) implies $\alpha_* \geq 1$, and $c_2'(\alpha_*, \lambda) = 1$ when $\alpha_* = 1$. Now consider any fixed $\alpha_* > 1$. As $c_2'(\alpha_*, \lambda)$ is strictly decreasing in $\lambda \in (0,1)$, choosing λ smaller leads to a looser constraint (i.e., more possible choices for c satisfying (48)), and Corollary 2 allows taking

$\lambda \in (0,1)$ in (23) to be arbitrarily small. But as $c_2'(\alpha_*, \lambda)$ is continuous in $\lambda \in (0,1)$ and because (48) has a strict inequality, we can replace (48) with the constraint

$$c < \frac{1}{2\alpha_* - 1} \equiv c_2(\alpha_*), \quad \text{which satisfies} \quad 0 < c_2(\alpha_*) \le 1 \tag{49}$$

since $\alpha_* \ge 1$ by (36). (To see why (48) can be replaced by the constraint in (49), note that if $c < c_2(\alpha_*)$, then (48) also holds for any $\lambda \in (0, \lambda_0)$, where $\lambda_0 \equiv (1 - 2c\alpha_* + c)/(1-c)$ is strictly positive because $1 - 2c\alpha_* + c > 0$ by (49). Also, (36) ensures that $\lambda_0 \le 1$.) If $\alpha_* = 1$, then $c_2(\alpha_*) = 1$, making (49) the weakest possible constraint on c under Assumption 1.B. For $\alpha_* > 1$, we get $c_2(\alpha_*) < 1$.

Under Corollary 2, the optimal rate that $\mathrm{RMSE}[\widehat{\mu}^{\mathrm{RQ}}_{m_n,r_n}]$ decreases is about $n^{-v_2(\alpha_*)}$ as $n \to \infty$, with

$$v_2(\alpha_*) \equiv c_2(\alpha_*) \left(\alpha_* - \frac{1}{2}\right) + \frac{1}{2} = 1 \tag{50}$$

by (42), (49) and (36), so $v_2(\alpha_*) > v_{\mathrm{MC}} = 1/2$ by (40).

5.3. CLT Conditions for Bounded h and Corollary 3

We now derive a constraint on c as in (41) to ensure CLT (19) through Corollary 3, which requires that the integrand h is bounded (Assumption 3.B) and that condition (24) holds. Under Assumption 1.B, (38) implies that $(m_n, r_n) = (n^c, n^{1-c})$ satisfies (24) if $r_n \sigma^2_{m_n} = \omega\left(n^{1-c} n^{-2c\alpha_* - 2\epsilon}\right) \to \infty$ for all sufficiently small $\epsilon > 0$ as $n \to \infty$, which is true when $1 - c - 2c\alpha_* > 0$, or equivalently,

$$c < \frac{1}{2\alpha_* + 1} \equiv c_3(\alpha_*), \quad \text{which satisfies} \quad 0 < c_3(\alpha_*) < \frac{1}{2} \tag{51}$$

under the assumption that $\alpha_* > 1/2$ in (37). If $\alpha_* \ge 1$, as (36) ensures when $V_{\mathrm{HK}}(h) < \infty$, which is not required by Corollary 3, then $0 < c_3(\alpha_*) \le 1/3$.

By (42) and (51), the optimal rate at which $\mathrm{RMSE}\left[\widehat{\mu}^{\mathrm{RQ}}_{m_n,r_n}\right]$ decreases (as $n \to \infty$) under Corollary 3 is about $n^{-v_3(\alpha_*)}$ with

$$v_3(\alpha_*) \equiv c_3(\alpha_*) \left(\alpha_* - \frac{1}{2}\right) + \frac{1}{2} = \frac{2\alpha_*}{2\alpha_* + 1}, \quad \text{which satisfies} \quad \frac{1}{2} < v_3(\alpha_*) < 1 \tag{52}$$

when $\alpha_* > 1/2$, as assumed in (37). (If $\alpha_* \ge 1$, as in (36), then $2/3 \le v_3(\alpha_*) < 1$.) Hence, under Corollary 3, the RQMC estimator's RMSE shrinks faster than the MC estimator's RMSE by (40).

5.4. CLT Conditions Under Moment Conditions of Corollary 4

We now derive a constraint on c as in (41) to ensure CLT (19) through Corollary 4, which assumes that (11) (from Assumption 3.C) and (25) both hold for some $b > 0$ (used for the order-$(2+b)$ absolute central moment of $h(U)$). In (25) 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 (38). Thus, $(m_n, r_n) = (n^c, n^{1-c})$ satisfies (25) if $r_n^{b/(2+b)}\sigma_{m_n}^2 = \omega\left(n^{(1-c)\frac{b}{2+b} - 2c\alpha_* - 2\epsilon}\right) \to \infty$ as $n \to \infty$, which holds when $(1-c)\frac{b}{2+b} - 2c\alpha_* > 0$, or equivalently,

$$c < \frac{1}{2\alpha_*(1+\frac{2}{b})+1} \equiv c_4(\alpha_*, b), \quad \text{which satisfies} \quad 0 < c_4(\alpha_*, b) < \frac{1}{2} \tag{53}$$

for each $b > 0$ when $\alpha_* > 1/2$, as assumed in (37). If $\alpha_* \geq 1$, as (36) ensures when $V_{\mathrm{HK}}(h) < \infty$ (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 (53).

By (42), under Corollary 4, the optimal rate at which $\mathrm{RMSE}[\widehat{\mu}^{\mathrm{RQ}}_{m_n,r_n}]$ decreases (as $n \to \infty$) is about $n^{-v_4(\alpha_*,b)}$, where for any $b > 0$,

$$v_4(\alpha_*, b) \equiv c_4(\alpha_*, b)\left(\alpha_* - \frac{1}{2}\right) + \frac{1}{2} = \frac{2\alpha_*(1+\frac{1}{b})}{2\alpha_*(1+\frac{2}{b})+1}, \quad \text{which satisfies} \quad \frac{1}{2} < v_4(\alpha_*, b) < 1 \tag{54}$$

when $\alpha_* > 1/2$, as in (37). Hence, we get $v_4(\alpha_*, b) > v_{\mathrm{MC}} = 1/2$ by (40) 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 $\mathrm{RMSE}[\widehat{\mu}^{\mathrm{RQ}}_{m_n,r_n}]$ under Corollary 4 improves as $h(U)$ has more finite absolute central moments.

5.5. AVCI Conditions When $V_{\mathrm{HK}}(h) < \infty$ Under Corollary 7

For $I^{\mathrm{RQ}}_{m_n,r_n,\gamma}$ in (28) to be AVCI (30), we assumed, in addition to the CLT in (19), that (29) holds. Corollary 7 ensures (29) is satisfied when $V_{\mathrm{HK}}(h) < \infty$ (Assumption 3.A) and condition (32) holds, which is the same as the square of CLT condition (23) for $\lambda = 1/2$. In the setting of Assumption 1.B, (23) holds for $\lambda = 1/2$ if c satisfies (48) with $\lambda = 1/2$, so AVCI condition (32) is true when

$$c < \frac{1}{4\alpha_* - 3} \equiv c_7(\alpha_*), \quad \text{which satisfies} \quad 0 < c_7(\alpha_*) \leq 1 \tag{55}$$

because $\alpha_* \geq 1$ by (36). Note that $c_7(\alpha_*) = 1$ when $\alpha_* = 1$, and $c_7(\alpha_*) < 1$ for $\alpha_* > 1$.

By (42), under Corollary 7, the optimal rate at which RMSE$[\widehat{\mu}^{\mathrm{RQ}}_{m_n,r_n}]$ decreases (as $n \to \infty$) is about $n^{-v_7(\alpha_*)}$, with

$$v_7(\alpha_*) \equiv c_7(\alpha_*) \left(\alpha_* - \frac{1}{2}\right) + \frac{1}{2} = \frac{3\alpha_* - 2}{4\alpha_* - 3}, \quad \text{which satisfies} \quad \frac{3}{4} < v_7(\alpha_*) \leq 1 \tag{56}$$

since $\alpha_* \geq 1$ by (36). Note that $v_7(\alpha_*) = 1$ when $\alpha_* = 1$, and $v_7(\alpha_*) < 1$ for $\alpha_* > 1$. Also, (56) implies that $v_7(\alpha_*) > v_{\mathrm{MC}} = 1/2$ by (40).

5.6. AVCI Conditions Under Moment Conditions of Corollary 8

By Corollary 8, which does not require $V_{\mathrm{HK}}(h) < \infty$, the combination of (11) holding for $b = 2$ (Assumption 3.C) and condition (33) implies (29), then ensuring AVCI (30). Under Assumption 1.B, we have in (33) that $\sigma^4_{m_n} = \omega(n^{-4c\alpha_* - 4\epsilon})$ as $n \to \infty$ for any $\epsilon > 0$ by (38) and $r_n = n^{1-c}$. Thus, (33) holds if $r_n \sigma^4_{m_n} = \omega(n^{1-c} n^{-4c\alpha_* - 4\epsilon}) \to \infty$ as $n \to \infty$, which is true if $1 - c - 4c\alpha_* > 0$, or equivalently,

$$c < \frac{1}{4\alpha_* + 1} \equiv c_8(\alpha_*), \quad \text{which satisfies} \quad 0 < c_8(\alpha_*) < \frac{1}{3} \tag{57}$$

when $\alpha_* > 1/2$, as assumed in (37). If $\alpha_* \geq 1$, as (36) ensures when $V_{\mathrm{HK}}(h) < \infty$ (Assumption 3.A) which is not required by Corollary 8, then $0 < c_8(\alpha_*) \leq 1/5$.

By (42), under Corollary 8, the optimal rate at which RMSE$[\widehat{\mu}^{\mathrm{RQ}}_{m_n,r_n}]$ decreases (as $n \to \infty$) is about $n^{-v_8(\alpha_*)}$, where

$$v_8(\alpha_*) \equiv c_8(\alpha_*) \left(\alpha_* - \frac{1}{2}\right) + \frac{1}{2} = \frac{3\alpha_*}{4\alpha_* + 1}, \quad \text{which satisfies} \quad \frac{1}{2} < v_8(\alpha_*) < \frac{3}{4} \tag{58}$$

when $\alpha_* > 1/2$, as assumed in (37). Thus, we have $v_8(\alpha_*) > v_{\mathrm{MC}} = 1/2$ by (40).

5.7. Remarks on the Case $m_n \equiv m_0$ in Corollaries 5 and 9

As noted before after (41), Corollaries 5 and 9 essentially have $c = 0$ as they assume that $n^c = m_n = m_0$ is fixed in (26). Thus, under Corollaries 5 and 9, the exact rate at which RMSE$[\widehat{\mu}^{\mathrm{RQ}}_{m_n,r_n}]$ decreases (as $n \to \infty$) is $n^{-1/2}$, which can be seen through (42) and (45) by setting $c_k(\alpha_*) = 0$. Hence, the RMSE for this case of RQMC shrinks at the same rate as for MC by (40), but RQMC, which can be viewed as a variance-reduction technique (e.g., see L'Ecuyer 2018, L'Ecuyer and Lemieux 2000), typically has a smaller leading coefficient for its rate.

5.8. Summary of Results Under Assumption 1.B

Section EC.2 (resp., EC.3) of the e-companion provides analytical (resp., graphical) comparisons of some of the $c_k(\alpha_*)$ and $v_k(\alpha_*)$ values for the various Corollaries k. We next summarize those asymptotic (as $n \to \infty$) results and the others from this section. (Section EC.4 of the e-companion presents results from numerical experiments studying the behavior of the asymptotic results for finite n.)

PROPOSITION 4. *Under Assumption 1.B that* $(m_n, r_n) = (n^c, n^{1-c})$ *for some* $c \in (0,1)$*, define* α_* *as in* (34)*, which is the exponential decreasing rate of the standard deviation of an estimator from a single randomization, and which satisfies* $\alpha_* > 1/2$ *under the assumption* (37)*. Also, define* $c_k(\alpha_*)$ *as in* (41)*, which is the supremum of values of c (depending on the particular Corollary k considered) to guarantee CLT* (19) *or AVCI* (30)*. Also, consider* $v_k(\alpha_*)$ *in* (45)*, which is the corresponding maximum RMSE convergence rate from* r_n *i.i.d. randomizations. Then the following properties hold:*

- *Under absolute central moment condition* (21) *on* $A_m = (1/m)\sum_{i=1}^{m} h(U_i')$ *for some* $b' > 0$*, Corollary 1 ensures the CLT when* $c < c_1(\alpha_*) = 1$*, which leads to* $v_1(\alpha_*) = \alpha_*$*; see* (47)*. If* (21) *holds for* $b' = 2$*, Corollary 6 secures AVCI when* $c < c_6(\alpha_*) = 1$*, which leads to* $v_6(\alpha_*) = \alpha_*$*; see* (47)*.*
- *When* $V_{\mathrm{HK}}(h) < \infty$ *(Assumption 3.A), which implies that* $\alpha_* \geq 1$ *by* (36)*, Corollary 2 ensures the CLT holds for* $c < c_2(\alpha_*) = \frac{1}{2\alpha_*-1}$*, leading to* $v_2(\alpha_*) = 1$*; see* (49) *and* (50)*. A condition for AVCI under Corollary 7 is* $c < c_7(\alpha_*) = \frac{1}{4\alpha_*-3}$*, which leads to* $v_7(\alpha_*) = \frac{3\alpha_*-2}{4\alpha_*-3}$*; see* (55) *and* (56)*.*
- *When h is bounded (Assumption 3.B), Corollary 3 secures a CLT for supremum values* $c_3(\alpha_*) = \frac{1}{2\alpha_*+1}$ *and* $v_3(\alpha_*) = \frac{2\alpha_*}{2\alpha_*+1}$*; see* (51) *and* (52)*.*
- *When* $h(U)$ *has finite absolute central moment of order* $2+b$ *(Assumption 3.C), Corollary 4 secures a CLT for* $c < c_4(\alpha_*, b) = \frac{1}{2\alpha_*(1+\frac{2}{b})+1}$*, for which* $v_4(\alpha_*, b) = \frac{2\alpha_*(1+\frac{1}{b})}{2\alpha_*(1+\frac{2}{b})+1}$*; see* (53) *and* (54)*. AVCI holds for* $b = 2$ *under Corollary 8 for* $c < c_8(\alpha_*) = \frac{1}{4\alpha_*+1}$*, for which* $v_8(\alpha_*) = \frac{3\alpha_*}{4\alpha_*+1}$*; see* (57) *and* (58)*.*

- *For each fixed* $\alpha_* > 1/2$, $c_4(\alpha_*, b)$ *(resp.,* $v_4(\alpha_*, b)$*) strictly increases in* $b > 0$ *and converges to* $c_3(\alpha_*)$ *(resp.,* $v_3(\alpha_*)$*) as* $b \to \infty$.

- $v_k(\alpha_*) > 1/2$ *for each* $k \in \{1, 2, 3, 4, 6, 7, 8\}$*, all of which have* $c_k(\alpha_*) > 0$*; i.e., the RMSE's optimal approximate convergence rate for RQMC is better than for MC. Under Corollaries 5 and 9 (fixed* $m_n \equiv m_0$*),* $\mathrm{RMSE}[\widehat{\mu}^{RQ}_{m_n,r_n}]$ *decreases (as* $n \to \infty$*) at the MC rate* $n^{-1/2}$.

- *Stricter conditions on* h *(see Proposition 1) in a Corollary* k *lead to a larger corresponding* $c_k(\alpha_*)$*; e.g.,* $c_2(\alpha_*) > c_3(\alpha_*) > c_4(\alpha_*, b)$ *and* $c_7(\alpha_*) > c_8(\alpha_*)$ *for each* $b > 0$ *and* $\alpha_* > 1/2$ *(*$c_2(\alpha_*)$ *being valid only when* $\alpha_* \geq 1$*), as shown in Sections EC.2.1 and EC.2.2 of the e-companion.*

- *For a particular condition on* h*,* $c_k(\alpha_*)$ *ensuring AVCI* (30) *is no larger (and often strictly smaller) than the corresponding* $c_{k'}(\alpha_*)$ *securing CLT* (19)*:* $c_7(\alpha_*) \leq c_2(\alpha_*)$ *and* $c_8(\alpha_*) < c_3(\alpha_*)$*; see Sections EC.2.3 and EC.2.4 of the e-companion.*

- $v_k(\alpha_*) > v_{k'}(\alpha_*)$ *if and only if* $c_k(\alpha_*) > c_{k'}(\alpha_*)$*: the stronger the requirement on the integrand, the faster the convergence rate; see* (46).

- $v_1(\alpha_*) = v_6(\alpha_*) \geq v_k(\alpha_*)$ *for all* $k \notin \{1, 6\}$ *because* $c_1(\alpha_*) = c_6(\alpha_*) = 1 \geq c_k(\alpha_*)$.

6. 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 i.i.d. randomizations, but typically limiting the size of r 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 (Loh 2003) is for scrambled digital nets with nested uniform scrambling, but this RQMC technique is computationally expensive. L'Ecuyer et al. (2010) proved that otherwise, increasing m for any fixed number of randomizations can lead to a non-normal limiting 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 the framework of Lindeberg's condition 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 of the conditions and the convergence speed of the resulting estimators.

While many of our corollaries specify that an allocation (m_n, r_n) has that m_n does not grow too quickly relative to r_n as $n \to \infty$, Corollary 1 (resp., 6) secures a CLT (resp., AVCI) for $r_n \to \infty$ at *any* arbitrarily slow rate as $n \to \infty$ when condition (21) holds for some $b' > 0$ (resp., $b' = 2$). We are currently investigating alternatives to condition (21) that similarly allow $r_n \to \infty$ at any rate (possibly sub-polynomial, e.g., $r_n = \lfloor \ln n \rfloor$). 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 (34), 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. Moreover, rather than build a CI for μ based on a CLT, we are additionally investigating instead employing resampling methods, such as the bootstrap t (Owen 2019, Chapter 17 end notes).

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

Asmussen S, Glynn P (2007) *Stochastic Simulation: Algorithms and Analysis* (New York: Springer).

Billingsley P (1995) *Probability and Measure* (New York: John Wiley and Sons), 3rd edition.

Breneis S (2020) Functions of bounded variation in one and multiple dimensions. mathesis, Johannes Kepler Universtat Linz, Austria.

Casella G, Berger RL (2002) *Statistical Inference* (Pacific Grove, Calif.: Duxbury), 2nd edition.

Choi SCT, Hickernell FJ, McCourt M, Sorokin A (2020) QMCPy: A quasi-Monte Carlo Python library. URL https://github.com/QMCSoftware/QMCSoftware.

Cranley R, Patterson TNL (1976) Randomization of number theoretic methods for multiple integration. *SIAM Journal on Numerical Analysis* 13(6):904–914.

Damerdji H (1994) Strong consistency of the variance estimator in steady-state simulation output analysis. *Mathematics of Operations Research* 19:494–512.

Durrett R (2019) *Probability: Theory and Examples* (Cambridge, UK: Cambridge University Press), 5th edition, URL https://services.math.duke.edu/~rtd/PTE/PTE5_011119.pdf.

Glynn PW (1987) Limit theorems for the method of replication. *Stochastic Models* 3(3):343–355.

Hickernell FJ (1998) A generalized discrepancy and quadrature error bound. *Mathematics of Computation* 67(221):299–322.

Hickernell FJ (2018) Error analysis for quasi-Monte Carlo methods. Glynn PW, Owen AB, eds., *Monte Carlo and Quasi-Monte Carlo Methods: MCQMC 2016*, 3–27 (Berlin: Springer).

Hong HS, Hickernell FH (2003) Algorithm 823: Implementing scrambled digital sequences. *ACM Transactions on Mathematical Software* 29:95–109.

Kroese DP, Taimre T, Botev ZI (2011) *Handbook of Monte Carlo Methods* (New York: John Wiley and Sons).

Kuipers L, Niederreiter H (1974) *Uniform Distribution of Sequences* (John Wiley, New York).

L'Ecuyer P (2018) Randomized quasi-Monte Carlo: An introduction for practitioners. Glynn PW, Owen AB, eds., *Monte Carlo and Quasi-Monte Carlo Methods: MCQMC 2016*, 29–52 (Berlin: Springer).

L'Ecuyer P, Lemieux C (2000) Variance reduction via lattice rules. *Management Science* 46(9):1214–1235.

L'Ecuyer P, Lemieux C (2002) Recent advances in randomized quasi-Monte Carlo methods. Dror M, L'Ecuyer P, Szidarovszky F, eds., *Modeling Uncertainty: An Examination of Stochastic Theory, Methods, and Applications*, 419–474 (Boston: Kluwer Academic).

L'Ecuyer P, Munger D, Tuffin B (2010) On the distribution of integration error by randomly-shifted lattice rules. *Electronic Journal of Statistics* 4:950–993.

Lehmann EL (1999) *Elements of Large-Sample Theory* (New York: Springer).

Lemieux C (2006) Quasi-random number techniques. Henderson SG, Nelson BL, eds., *Simulation*, 351–380, Handbooks in Operations Research and Management Science (Amsterdam, The Netherlands: Elsevier), chapter 12.

Lemieux C (2009) *Monte Carlo and Quasi-Monte Carlo Sampling*. Series in Statistics (New York: Springer).

Loh WL (2003) On the asymptotic distribution of scrambled net quadrature. *Annals of Statistics* 31(4):1282–1324.

Matousěk J (1998) On the L_2-discrepancy for anchored boxes. *J. of Complexity* 14:527–556.

Nakayama MK, Tuffin B (2021) Sufficient conditions for a central limit theorem to assess the error of randomized quasi-Monte Carlo methods. *Proceedings of the 2021 Winter Simulation Conference* (IEEE), accepted.

Niederreiter H (1992) *Random Number Generation and Quasi-Monte Carlo Methods*, volume 63 (Philadelphia: SIAM).

Owen A (2005) Multidimensional variation for quasi-Monte Carlo. *Contemporary Multivariate Analysis and Design of Experiments*, 49–74 (World Scientific), URL http://dx.doi.org/10.1142/9789812567765_0004.

Owen A, Rudolf D (2020) A strong law of large numbers for scrambled net integration, https://arxiv.org/abs/2002.07859.

Owen AB (1995) Randomly permuted (t,m,s)-nets and (t,s)-sequences. *Monte Carlo and Quasi-Monte Carlo Methods in Scientific Computing: Lecture Notes in Statistics*, volume 106, 299–317 (Springer).

Owen AB (1997a) Monte Carlo variance of scrambled net quadrature. *SIAM Journal of Numerical Analysis* 34:1884–1910.

Owen AB (1997b) Scrambled net variance for integrals of smooth functions. *Annals of Statistics* 25(4):1541–1562.

Owen AB (2019) *Monte Carlo Theory, Methods and Examples* (Draft available at http://statweb.stanford.edu/~owen/mc/), in preparation.

Serfling RJ (1980) *Approximation Theorems of Mathematical Statistics* (New York: John Wiley and Sons).

Shapiro SS, Wilk MB (1965) An analysis of variance test for normality (complete samples). *Biometrika* 52(3/4):591–611.

Shapiro SS, Wilk MB, Chen HJ (1968) A comparative study of various tests of normality. *Journal of the American Statistical Association* 63:1343–1372.

Tuffin B (1997) Variance reduction applied to product-form multi-class queuing networks. *ACM Transactions on Modeling and Computer Simulation* 7(4):478–500.

Tuffin B (1998) Variance reduction order using good lattice points in Monte Carlo methods. *Computing* 61(4):371–378.

Tuffin B (2004) Randomization of quasi-monte carlo methods for error estimation: Survey and normal approximation. *Monte Carlo Methods and Applications* 10(3-4):617–628.

Proofs and Other Results

EC.1. Proofs

Proof of Proposition 1. As shown in Owen and Rudolf (2020), functions of bounded Hardy-Krause variation are Riemann integrable, and therefore bounded, which Owen (2005) also uses. Thus, Assumption 3.A is stronger than Assumption 3.B. To show that the relation is strict, for any dimension $s \ge 2$, consider h to be an indicator function (so μ is a probability) with discontinuities not lining up with the axes: it is bounded but has $V_{\mathrm{HK}}(h) = \infty$. An explicit example (Owen and Rudolf 2020) is $h(u) = \mathcal{I}(\sum_{i=1}^{s} u_i \le 1)$, for $u = (u_1, \ldots, u_s) \in [0,1]^s$, 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 $|h(u) - \mu| \le 2t_0$ for all $u \in [0,1]^s$. It follows that $\mathbb{E}[|h(U) - \mu|^q] \le (2t_0)^q$ 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 $F^{\leftarrow}(u) = \inf\{x : F(x) \ge u\}$ 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 (Billingsley 1995, pp. 81 and 277). To show that the relation is strict, we modify an example of (Durrett 2019, p. 366). For $s = 1$, let $h(u) = F^{\leftarrow}(u)$, where F is the CDF of the density function $f(x) = k_0 \mathcal{I}(x \ge e)/[x^3 (\ln x)^2]$ with $k_0 = (\int_e^\infty \mathrm{d}x/[x^3(\ln x)^2])^{-1} \doteq 26.64$. As $h(U) \sim F$, we have that $\mathbb{E}[h(U)^2] = \int_e^\infty \frac{k_0 x^2}{x^3 (\ln x)^2}\,\mathrm{d}x = \int_e^\infty \frac{k_0}{x (\ln x)^2}\,\mathrm{d}x = \left[-\frac{k_0}{\ln x}\right]_{x=e}^{\infty} = k_0$, so $\mathrm{Var}[h(U)] < \infty$. But for any fixed $b > 0$, we get $\mathbb{E}[h(U)^{2+b}] = \int_e^\infty \frac{k_0 x^{2+b}}{x^3 (\ln x)^2}\,\mathrm{d}x = k_0 \int_e^\infty \frac{1}{x^{1-b}(\ln x)^2}\,\mathrm{d}x \ge (k_0/k_b) \int_e^\infty \frac{1}{x^{1-b/2}}\,\mathrm{d}x = \infty$ for $k_b = [4/(be)]^2$ because $(\ln x)^2 \le k_b x^{b/2}$ for all $x \ge e$. $\square$

Proof of Lemma 2. As seen by (10), $X_{n,1}$ averages m_n *dependent* terms, which we will handle through Minkowski's inequality (Billingsley 1995, eq. (5.40)):

$$\eta_{n,q} = \mathbb{E}\left[\left|\sum_{i=1}^{m_n} \frac{h(U'_{i,1}) - \mu}{m_n}\right|^q\right] \le \left[\sum_{i=1}^{m_n} \left(\mathbb{E}\left[\left|\frac{h(U'_{i,1}) - \mu}{m_n}\right|^q\right]\right)^{1/q}\right]^q$$
$$= \left[m_n \left(\mathbb{E}\left[\left|\frac{h(U) - \mu}{m_n}\right|^q\right]\right)^{1/q}\right]^q = \mathbb{E}\left[|h(U) - \mu|^q\right],$$

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

Proof of Theorem 2. Fix any $t>0$ for $\tau_n^2(t)$ in (17), and we will bound $\tau_n^2(t)$ as in the proof of the Lyapounov CLT in (Billingsley 1995, Theorem 27.3). Specifically, the condition $|y|>ts_n$ in (17) implies that $|y|^{b'}/(t^{b'}s_n^{b'})>1$, so

$$\tau_n^2(t) \le \frac{1}{t^{b'}s_n^{b'}}\int_{|y|>ts_n} y^2|y|^{b'}\,\mathrm{d}G_n(y) \le \frac{1}{t^{b'}s_n^{b'}}\mathbb{E}\left[|Y_{n,1}|^{2+b'}\right] = \frac{1}{t^{b'}r_n^{b'/2}\sigma_{m_n}^{b'}}\mathbb{E}\left[|X_{n,1}-\mu|^{2+b'}\right] \tag{EC.1}$$

by (16), with $\mathbb{E}\left[|X_{n,1}-\mu|^{2+b'}\right]<\infty$ by assumption, which ensures $\sigma_{m_n}^2<\infty$. Thus, as in (18), dividing throughout (EC.1) by $\sigma_{m_n}^2$, which is strictly positive for all n large enough (Assumption 4), shows that (20) guarantees (18) because $t>0$ is fixed and arbitrary, so CLT (19) holds by Theorem 1. □

Proof of Corollary 1. The moments of $X_{n,1}$ in (20) are the same as the moments of A_{m_n} in (21). As a consequence, because $m_n\to\infty$ as $n\to\infty$, (21) yields

$$\frac{\mathbb{E}\left[|X_{n,1}-\mu|^{2+b'}\right]}{\sigma_{m_n}^{2+b'}} \le k_1, \quad \text{for all } n \text{ sufficiently large.} \tag{EC.2}$$

For all n satisfying (EC.2) and Assumption 4, (EC.2) implies that the left side of (20) is bounded above by $k_1/r_n^{b'/2}$, which vanishes as $n\to\infty$ because $b'>0$ and $r_n\to\infty$ as $n\to\infty$ under Assumption 1.A. Thus, (20) holds, so CLT (19) follows from Theorem 2. Under (22), we have $\mathbb{E}\left[|X_{n,1}-\mu|^{2+b'}\right]\le k_2^{2+b'}k_2^{2+b'}$ for all n sufficiently large because $m_n\to\infty$, securing (21) with $k_1=k_2^{2+b'}$. □

Proof of Corollary 2. 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 4 also holds. Because $V_{\mathrm{HK}}(h)<\infty$, 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 (20) satisfies

$$\frac{\mathbb{E}\left[|X_{n,1}-\mu|^{2+b'}\right]}{r_n^{b'/2}\sigma_{m_n}^{2+b'}} \le \frac{1}{r_n^{b'/2}}\left(\frac{w_0'V_{\mathrm{HK}}(h)(\ln m_n)^s}{m_n\sigma_{m_n}}\right)^{2+b'}. \tag{EC.3}$$

Raising the right side of (EC.3) to the $2/(2+b')$ power shows that it vanishes as $n\to\infty$ if and only if

$$r_n^{b'/(2+b')}\left(\frac{m_n\sigma_{m_n}}{(\ln m_n)^s}\right)^2 \to\infty, \qquad \text{as } n\to\infty. \tag{EC.4}$$

But $b'/(2+b')=1-\lambda$, so (23) and (EC.4) imply (20), and (18) and (19) follow from Theorem 2. □

Proof of Corollary 3. As h is bounded, Assumption 3.B implies that $|Y_{n,j}| \le 2t_0$ for all n and j. Hence, (17) ensures $\tau_n^2(t) = 0$ whenever $ts_n > 2t_0$, which holds, given any $t > 0$, for all $n > n_0$ for some n_0 by (24). Thus, the numerator of (18) is zero for all $n > n_0$, so (19) follows from Theorem 1. □

Proof of Corollary 4. Because $b > 0$, raising (25), which has nonnegative left side, to the $(2+b)/2$ power is equivalent to

$$r_n^{b/2}\sigma_{m_n}^{2+b} \to \infty, \qquad \text{as } n \to \infty. \tag{EC.5}$$

By (11), Lemma 2 with $q = 2+b$ then implies that the left side of (20) satisfies

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

by (EC.5), so (20) holds. Hence, (18) and (19) follow from Theorem 2. □

Proof of Corollary 5. For $n \ge m_0$, we have by (26) that $X_{n,j} = X_j$ for $j = 1, 2, \ldots, r_n$, 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, $X_1, X_2, \ldots$ form a single i.i.d. sequence with mean μ and variance $\sigma_{m_0}^2 = \mathrm{Var}[X_1] = \mathrm{Var}[X_{n,1}]$, which Lemma 2 with $q = 2$ ensures is finite by Assumption 3.D. For G as the distribution of $X_1 - \mu$, (18) reduces to $\lim_{n\to\infty} \frac{1}{\sigma_{m_0}^2}\int_{|y|>t\sigma_{m_0}\sqrt{r_n}} y^2 \, \mathrm{d}G(y) \to 0$ as $n \to \infty$ for each $t > 0$, which holds because $\{|y| > t\sigma_{m_0}\sqrt{r_n}\} \to \emptyset$ as $n \to \infty$ since $0 < \sigma_{m_0} < \infty$ and $r_n = \lfloor n/m_0 \rfloor \to \infty$ as $n \to \infty$. Hence, Theorem 1 yields (19). □

Proof of Proposition 2. We first compare the restrictions on (m_n, r_n) in Corollaries 4 and 5. If (26) holds, then the left side of (25) becomes $r_n^{b/(2+b)}\sigma_{m_n}^2 = (\lfloor n/m_0 \rfloor)^{b/(2+b)}\sigma_{m_0}^2 \to \infty$ as $n \to \infty$ because $\sigma_{m_0}^2 > 0$, so (26) implies (25).

We next analyze the requirements on (m_n, r_n) in Corollaries 3 and 4. Expressing the left side of (24) as $r_n\sigma_{m_n}^2 = r_n^{2/(2+b)} r_n^{b/(2+b)}\sigma_{m_n}^2$ shows that (25) implies (24) since $r_n^{2/(2+b)} \to \infty$ as $n \to \infty$ by Assumption 1.A.

We finally compare condition (23) of Corollary 2 with condition (24) of Corollary 3. Writing the left side of (23) as

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

makes clear that choosing (m_n, r_n) to satisfy (24) ensures that (23) is satisfied when (27) holds. To verify (27) 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 (27) 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 $n \to \infty$ for any $\lambda \in (0, \lambda_0)$, where $\lambda_0 = \min(2c/(1-c), 1) > 0$. (Corollary 2 allows taking $\lambda > 0$ arbitrarily small.) $\square$

Proof of Theorem 3. Note that (29) implies that $\sigma_{m_n}/\widehat{\sigma}_{m_n,r_n} \Rightarrow 1$ as $n \to \infty$ by the continuous-mapping theorem (e.g., Billingsley 1995, Theorem 25.7). Thus, (19) ensures that

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

by Slutsky's theorem (Serfling 1980, Theorem 1.5.4). Hence,

$$\mathbb{P}(\mu \in I^{\mathrm{RQ}}_{m_n,r_n,\gamma}) = \mathbb{P}\left(-z_\gamma \le \frac{\widehat{\mu}^{\mathrm{RQ}}_{m_n,r_n} - \mu}{\widehat{\sigma}_{m_n,r_n}/\sqrt{r_n}} \le z_\gamma\right) \to \gamma, \qquad \text{as } n \to \infty,$$

by the portmanteau theorem (Billingsley 1995, Theorem 25.8), establishing (30). $\square$

Proof of Theorem 4. As (31) is equivalent to (20) for $b' = 2$, Theorem 2 guarantees CLT (19) because we assumed that $\mathbb{E}\left[(X_{n,1} - \mu)^4\right] < \infty$. Thus, if (29) holds, then Theorem 3 will imply (30).

We will prove (29) by establishing that $p_n \equiv \mathbb{P}\left(\left|\frac{\widehat{\sigma}^2_{m_n,r_n}}{\sigma^2_{m_n}} - 1\right| > v\right) \to 0$ as $n \to \infty$ for each fixed $v > 0$. Assume that n is large enough so that $r_n \ge 2$ (see Assumption 1.A). By Chebyshev's inequality (Billingsley 1995, eq. (21.13)) and because $\mathbb{E}[\widehat{\sigma}^2_{m_n,r_n}] = \sigma^2_{m_n}$ (Serfling 1980, p. 173), we get

$$\begin{aligned} p_n &\le \frac{1}{v^2}\mathbb{E}\left[\left(\frac{\widehat{\sigma}^2_{m_n,r_n}}{\sigma^2_{m_n}} - 1\right)^2\right] = \frac{1}{v^2\sigma^4_{m_n}}\mathbb{E}\left[\left(\widehat{\sigma}^2_{m_n,r_n} - \sigma^2_{m_n}\right)^2\right] = \frac{1}{v^2\sigma^4_{m_n}}\mathrm{Var}\left[\widehat{\sigma}^2_{m_n,r_n}\right] \\ &= \frac{1}{v^2 r_n \sigma^4_{m_n}}\left(\mathbb{E}\left[(X_{n,1} - \mu)^4\right] - \frac{r_n - 3}{r_n - 1}\sigma^4_{m_n}\right), \end{aligned} \tag{EC.6}$$

where (EC.6) follows from (Serfling 1980, p. 184), with $\mathbb{E}\left[(X_{n,1} - \mu)^4\right] < \infty$ by assumption. Because $r_n \to \infty$ as $n \to \infty$ by Assumption 1.A, (EC.6) vanishes as $n \to \infty$ by (31) since $v > 0$ is fixed, thus verifying (29). $\square$

Proof of Corollary 6. As shown in (EC.2), because $m_n \to \infty$, (21) with $b' = 2$ implies that for all n sufficiently large, the left side of (31) is bounded above by k_1/r_n, which vanishes as $n \to \infty$ because $r_n \to \infty$ as $n \to \infty$ under Assumption 1.A. Thus, (31) holds, so CLT (19), (29), and AVCI

(30) follow from Theorem 4. Under (22), we have $\mathbb{E}[|X_{n,1}-\mu|^4] \le k_2^4 \sigma_{m_n}^4$ for all n sufficiently large because $m_n \to \infty$, securing (EC.2) for $b'=2$ with $k_1 = k_2^4$. $\square$

Proof of Corollary 7. As the left side of (32) is nonnegative, taking the square-root of (32) shows that it is equivalent to (23) with $\lambda = 1/2$, so (32) guarantees CLT (19) by Corollary 2. We will next establish (29) and (30) 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 4 also holds. Because $V_{\mathrm{HK}}(h) < \infty$ from Assumption 3.A, Lemma 1 implies that $\mathbb{E}\left[(X_{n,1}-\mu)^4\right] < \infty$ by (12) with $q=4$. Also, (12) ensures that the left side of (31) satisfies

$$\frac{\mathbb{E}\left[(X_{n,1}-\mu)^4\right]}{r_n \sigma_{m_n}^4} \le \frac{1}{r_n}\left(\frac{w_0' V_{\mathrm{HK}}(h)(\ln m_n)^s}{m_n \sigma_{m_n}}\right)^4. \tag{EC.7}$$

Using (32) in (EC.7) verifies (31), so (29) and (30) follow from Theorem 4. $\square$

Proof of Corollary 8. As the left side of (33) is nonnegative, taking the square-root of (33) shows it is equivalent to (25) with $b=2$, so (33) ensures CLT (19) by Corollary 4 because we further assumed that (11) holds for $b=2$. We will next establish (29) and (30) by verifying the conditions of Theorem 4. By Lemma 2 with $q=4$, we see that the numerator in (31) 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, (33) implies (31), so Theorem 4 yields (29) and (30). $\square$

Proof of Corollary 9. Corollary 5 ensures the CLT (19) is true. Also, as each $X_{n,j} = X_j$ in (13) does not depend on n, the triangular array reduces to a single i.i.d. sequence. Consequently, standard arguments (e.g., Serfling 1980, Theorem 2.2.3A) show that (29) holds, so Theorem 3 implies (30). $\square$

Proof of Proposition 3. If (26) holds, then $r_n \sigma_{m_n}^4 = \lfloor n/m_0 \rfloor \sigma_{m_0}^4 \to \infty$ as $n \to \infty$ because $\sigma_{m_0}^4 > 0$, so (26) implies (33). To compare (32) and (33), 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 (32) is weaker than (33). $\square$

EC.2. Analytical Comparisons of the $c_k(\alpha_*)$ and the $v_k(\alpha_*)$ from Section 5

For the various Corollaries k from Sections 3.1 and 4.1, we now compare their corresponding values of the upper bounds $c_k(\alpha_*)$ in (41) for $c \in (0,1)$ in Assumption 1.B and the optimal approximate rates

$v_k(\alpha_*)$ in (42) from Sections 5.1–5.6. The cases $k=1$ and 6 are omitted from the discussion since their condition (21) or (22) is different from the assumptions used in the other corollaries, combining properties of the RQMC method and the integrand. Conditions (21) and (22) may also be difficult to verify in practice. Note nevertheless that $c_1(\alpha_*) = c_6(\alpha_*) = 1 \geq c_k(\alpha_*)$ and $v_1(\alpha_*) = v_6(\alpha_*) = \alpha_* \geq v_k(\alpha_*)$ for all $k \notin \{1, 6\}$. As shown before in Propositions 1 and 2 under Assumption 1.A, we will see trade-offs in the conditions that ensure CLT (19) 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 (41). A similar situation will also hold for guaranteeing AVCI (30), 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.

EC.2.1. $k=2$ vs. $k=3$ and $k=4$

Comparing $c_2(\alpha_*)$ in (49) of Corollary 2 with $c_3(\alpha_*)$ from (51) for Corollary 3 shows that $c_2(\alpha_*) > c_3(\alpha_*)$; also see Proposition 2. This then implies that $v_2(\alpha_*) > v_3(\alpha_*)$ by (46). But recall that Corollary 2 required that $V_{\mathrm{HK}}(h) < \infty$ (Assumption 3.A), which is stronger (Proposition 1) than restricting h to be bounded (Assumption 3.B), as Corollary 3 imposed.

We now compare $c_2(\alpha_*)$ in (49) from Corollary 2 to $c_4(\alpha_*, b)$ from (53) of Corollary 4 when $V_{\mathrm{HK}}(h) < \infty$ (Assumption 3.A), as required by Corollary 2 but not by Corollary 4. We have that $c_2(\alpha_*) > c_4(\alpha_*, b)$ for all $b > 0$. Thus, condition (49) is (substantially) less restrictive on our choices for c than (53) for each $b > 0$. This further implies that $v_2(\alpha_*) > v_4(\alpha_*, b)$ by (46).

EC.2.2. $k=3$ vs. $k=4$

Observe that $c_3(\alpha_*) > c_4(\alpha_*, b)$ for each $b > 0$ by (51) and (53), so the condition (53) for Corollary 4 restricts our choices for c more than condition (51) from Corollary 3, but (51) 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 (53) for Corollary 4. Because $c_4(\alpha_*, b)$ is strictly increasing in b, the constraint on c from condition (53) 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 (53). Similarly, by (46), the exponent $v_4(\alpha_*, b)$ from (54) governing the optimal rate at which $\mathrm{RMSE}[\widehat{\mu}^{\mathrm{RQ}}_{m_n,r_n}]$ decreases as $n \to \infty$ under Corollary 4 is strictly worse than the exponent $v_3(\alpha_*)$ from (52) under Corollary 3.

Moreover, note that $\lim_{b\to\infty} c_4(\alpha_*, b) = c_3(\alpha_*)$, the latter being the upper bound for c in (51) when the integrand h is bounded. Similarly, we have that $\lim_{b\to\infty} v_4(\alpha_*, b) = v_3(\alpha_*)$. Thus, the tradeoffs of the conditions of Corollaries 3 and 4 disappear as $b \to \infty$. While this did not necessarily have to happen (e.g., as in the proof of Proposition 1, consider $h(u) = \Phi^{\leftarrow}(u)$, which is unbounded but $h(U)$ has finite moments of all orders), it is reasonable: we can think of bounded h as being an extreme special case as $b \to \infty$ of the order-$(2+b)$ absolute central moment of $h(U)$ being finite.

EC.2.3. $k = 7$ vs. $k = 2$ and $k = 8$

We first compare $c_7(\alpha_*)$ from (55) and $c_8(\alpha_*)$ in (57), each of which is an upper bound for c to ensure AVCI (30). It is clear that $c_7(\alpha_*) > c_8(\alpha_*)$, so condition (55), obtained under $V_{\mathrm{HK}}(h) < \infty$ (Assumption 3.A), is a strictly weaker restriction on the choice of c than the condition (57), derived under moment conditions (Assumption 3.C) but without requiring $V_{\mathrm{HK}}(h) < \infty$; also see Proposition 3. Thus, assuming the stricter condition $V_{\mathrm{HK}}(h) < \infty$ (see Proposition 1) allows us to expand the values of c that ensure AVCI. Moreover, we have that $v_7(\alpha_*) > v_8(\alpha_*)$ by (46), so when $V_{\mathrm{HK}}(h) < \infty$, the rate at which $\mathrm{Var}[\widehat{\mu}^{\mathrm{RQ}}_{m_n,r_n}]$ converges is faster by choosing c to optimally satisfy Corollary 7 rather than Corollary 8.

We now want to see if the upper bound $c_7(\alpha_*)$ in AVCI condition (55) is more restrictive than the upper bound $c_2(\alpha_*)$ from CLT condition (49), where both were obtained under the assumption that $V_{\mathrm{HK}}(h) < \infty$. When $\alpha_* = 1$, we have $c_7(\alpha_*) = c_2(\alpha_*) = 1$. For $\alpha_* > 1$, we get $c_7(\alpha_*) < c_2(\alpha_*)$, Thus, when $\alpha_* = 1$, the same values of c guarantee both CLT (19) and AVCI (30). But for $\alpha_* > 1$, ensuring AVCI (30) requires restricting c more than what is needed to make sure CLT (19) holds.

EC.2.4. $k = 8$ vs. $k = 3$ and $k = 4$

Note that $c_8(\alpha_*) < c_3(\alpha_*)$ for all $\alpha_* > 0$ by (51) and (57), so AVCI condition (57) obtained from Corollary 8 restricts our choices for c more than the CLT condition (51) under Corollary 3. Hence, ensuring AVCI (30) under Corollary 8 (finite 4th absolute central moment) requires further constraining c compared to obtaining CLT (19) under Corollary 3 (bounded h). Moreover, by (46), it follows that $v_8(\alpha_*) < v_3(\alpha_*)$.

By Corollary 4, CLT (19) also is secured when (53) and (11) hold for some $b > 0$, as discussed in Section 5.4. To compare the AVCI upper bound $c_8(\alpha_*)$ in (57) to CLT bound $c_4(\alpha_*)$ in (53), which has $c_4(\alpha_*, b) = 1/[1 + 2\alpha_*(1 + \frac{2}{b})]$, note that $c_4(\alpha_*, b) \leq c_8(\alpha_*)$ 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 (19) under Corollary 4, then the same c also yields AVCI (30).

EC.3. Graphical Comparisons of the $c_k(\alpha_*)$ and the $v_k(\alpha_*)$ from Section 5

Figure EC.1 shows the upper bounds $c_k(\alpha_*)$ in (41) for c in Assumption 1.B, where $c_k(\alpha_*)$ is given by $c_2(\alpha_*)$ in (49) for Corollary 2, $c_3(\alpha_*)$ in (51) for Corollary 3, $c_4(\alpha_*, b)$ in (53) for Corollary 4, $c_7(\alpha_*)$ in (55) for Corollary 7, and $c_8(\alpha_*)$ in (57) for Corollary 8. 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 (49), (51), (53), (55), and (57). Note that $c_* = 1$ also corresponds to $c_1(\alpha_*) = c_6(\alpha_*) = 1$ in (47). We do not include $v_1(\alpha_*) = v_6(\alpha_*) = \alpha_*$ from (47) in the graphs to better see the differences among the other corollaries.

Figure EC.1 corroborates the results obtained in Section EC.2 and summarized in Proposition 4. We can readily check the following:

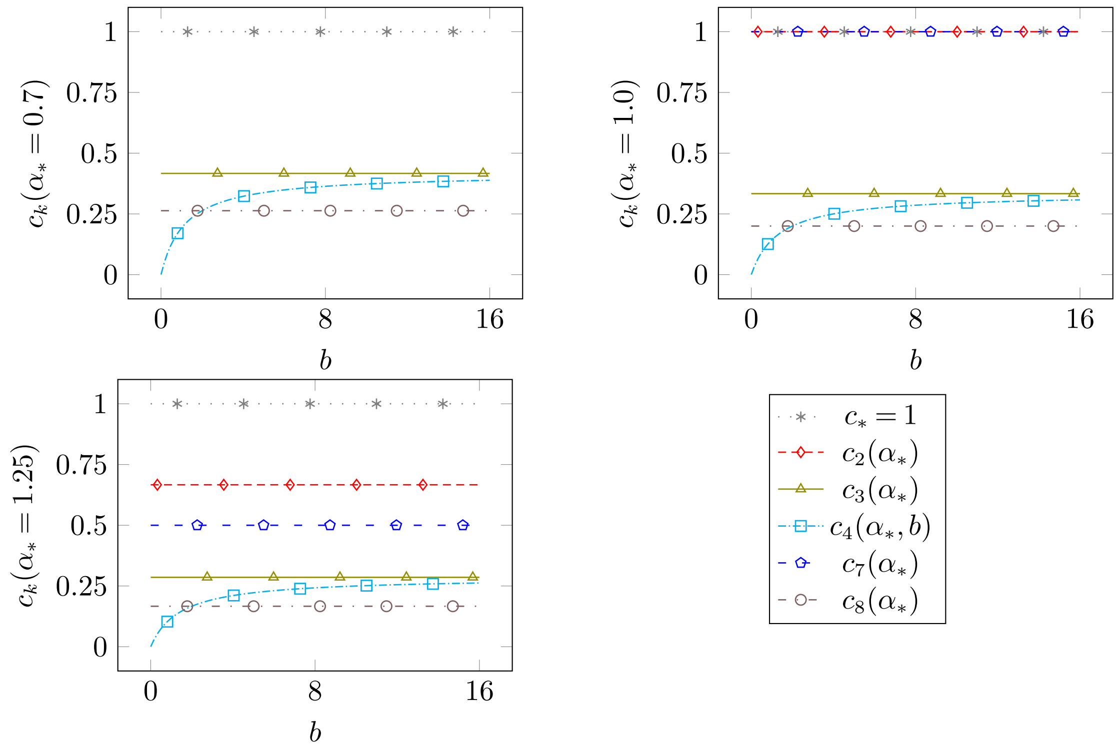


Figure EC.1 Plots of the upper bounds $c_k(\alpha_*)$ in (41) of c in Assumption 1.B for different Corollaries k. The plots display the $c_k(\alpha_*)$ as functions of b from Assumption 3.C for different fixed values of α_*. The upper left panel does not include $c_2(\alpha_*)$ and $c_7(\alpha_*)$ because these require $V_{\mathrm{HK}}(h) < \infty$, which then implies $\alpha_* \geq 1$ by (36). The plots show that stronger restrictions on the integrand h lead to larger $c_k(\alpha_*)$.

- $c_4(\alpha_*, b) < c_3(\alpha_*) < c_2(\alpha_*)$ ($c_2(\alpha_*)$ 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 (19) (Proposition 2).
- $c_8(\alpha_*) < c_7(\alpha_*)$, so we similarly see that a stronger condition on h leads to larger range of values of c that ensure AVCI (30); also see Proposition 3.
- $c_4(\alpha_*, b)$ approaches $c_3(\alpha_*)$ as b grows large in Assumption 3.C, which agrees with the principle that finite absolute central moments of order $2+b$ as $b \to \infty$ is "close" to meaning a bounded integrand; see the related discussion at the end of Section EC.2.2.

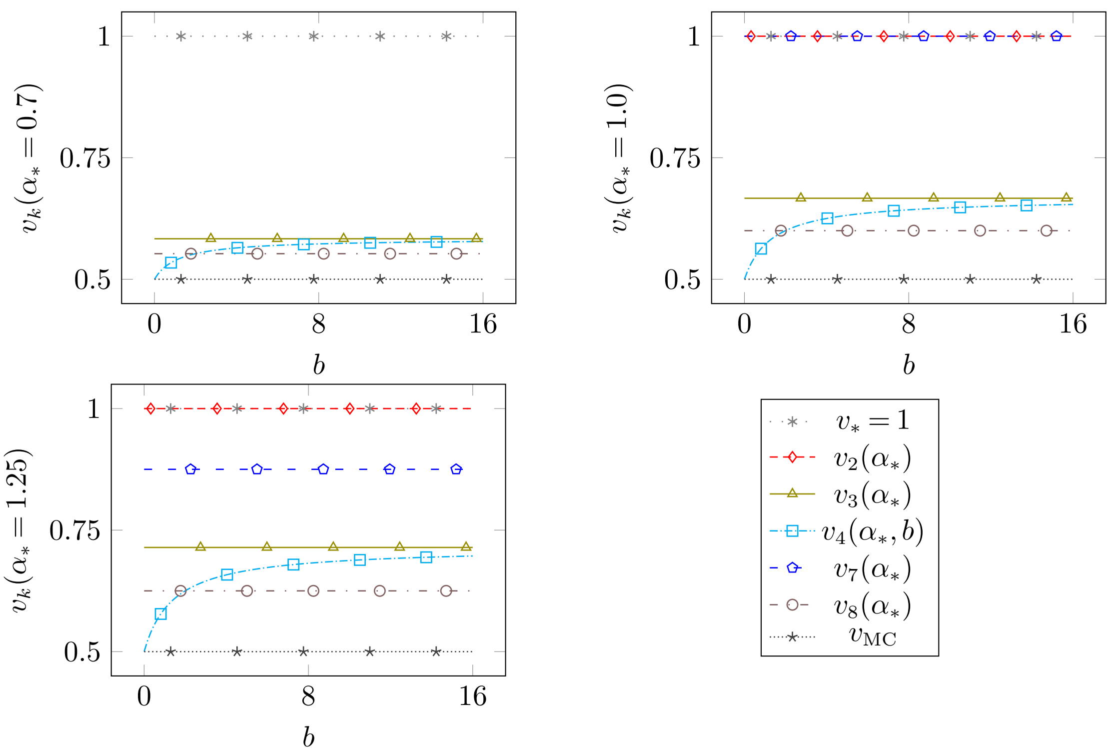


Figure EC.2 Plots of the negative exponent $v_k(\alpha_*)$ of the optimal rate at which the estimator RMSE decreases as functions of b for different values of α_*. The upper left panel does not include $v_2(\alpha_*)$ and $v_7(\alpha_*)$ because these require $V_{\mathrm{HK}}(h) < \infty$, which then implies $\alpha_* \geq 1$ by (36). The plots show that stronger restrictions on the integrand h lead to larger $v_k(\alpha_*)$.

- As b increases in Assumption 3.C (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 (30) often (but not always) entails restricting c more than what guarantees a CLT, which can be seen from $c_8(\alpha_*) < c_3(\alpha_*)$ and $c_7(\alpha_*) \leq c_2(\alpha_*)$.

Figure EC.2 plots, as functions of $b > 0$ from Assumption 3.C, the (negative) exponent $v_k(\alpha_*)$ from (42) of the optimal rate at which the RQMC estimator's RMSE decreases, given by $v_2(\alpha_*)$ in (50) for Corollary 2, $v_3(\alpha_*)$ in (52) for Corollary 3, $v_4(\alpha_*, b)$ in (54) for Corollary 4, $v_7(\alpha_*)$ in (56) for Corollary 7, and $v_8(\alpha_*)$ in (58) for Corollary 8. By (50), (52), (54), (56), and (58), each of these exponents is at most $v_* = 1$, which is also plotted for reference. The figure further includes $v_{\mathrm{MC}} = 1/2$

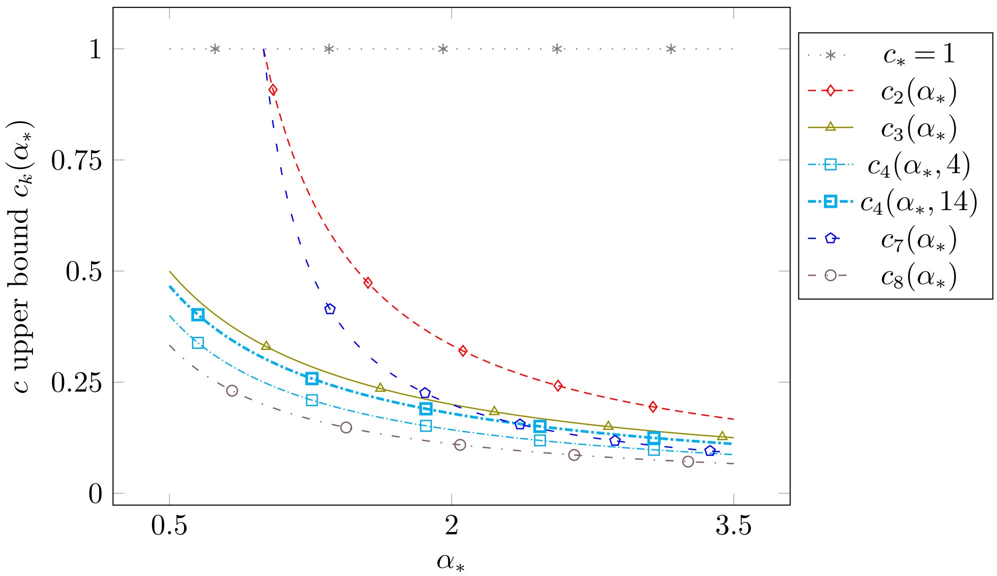


Figure EC.3 Plots of the upper bounds $c_k(\alpha_*)$ of c as functions of α_*. Functions $c_2(\alpha_*)$ and $c_7(\alpha_*)$ require $V_{\mathrm{HK}}(h) < \infty$, so they are shown for only $\alpha_* \geq 1$ because of (36).

from (40) for comparison, and all of these $v_k(\alpha_*)$ are strictly greater than v_{MC}, so RQMC with $c > 0$ always has that the optimal rate at which its RMSE decreases is better than for MC. By (46), 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 EC.1, we see that each upper bound $c_k(\alpha_*)$ in (41) on c decreases as α_* increases. To investigate this further, Figures EC.3 and EC.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 EC.3 more clearly illustrates that as α_* grows, each upper bound $c_k(\alpha_*)$ in (41) on c decreases, so guaranteeing a CLT (19) or AVCI (30) requires putting more effort on the MC part (i.e., for fixed $n > 0$, $r_n = n^{1-c}$ 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 EC.4 shows that this is not the case for all CLT and most AVCI conditions: the

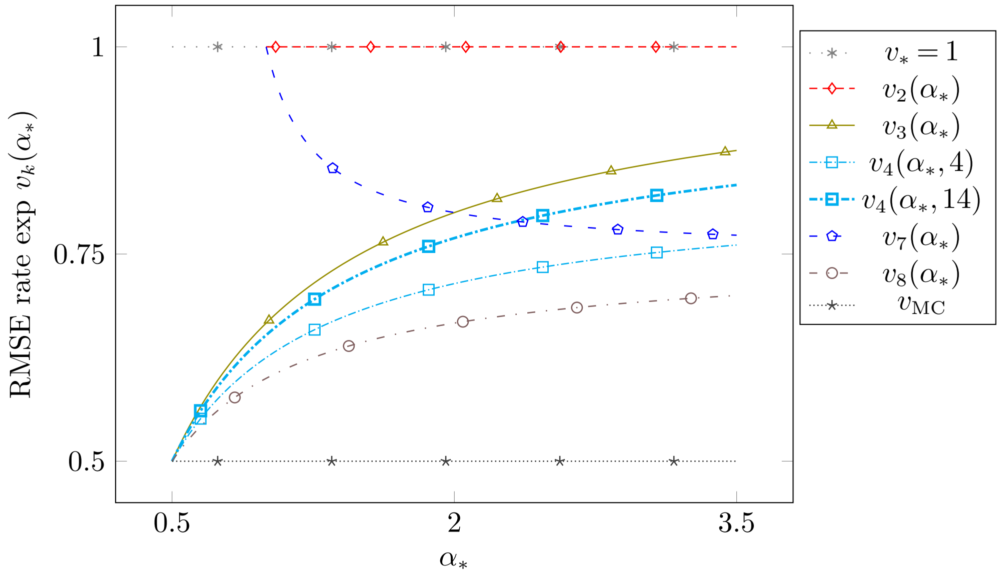


Figure EC.4 Plots of the negative exponent $v_k(\alpha_*)$ of the optimal rate at which the estimator RMSE decreases as functions of α_*. Functions $v_2(\alpha_*)$ and $v_7(\alpha_*)$ require $V_{\mathrm{HK}}(h) < \infty$, so they are shown for only $\alpha_* \geq 1$ because of (36).

RMSE optimal convergence speed determined by $v_k(\alpha_*)$, as in (42), is generally increasing in α_*. The one exception is $v_7(\alpha_*)$, which determines the optimal RMSE convergence rate to ensure AVCI when $V_{\mathrm{HK}}(h) < \infty$. In this case, Figure EC.3 shows that as α_* increases, $c_7(\alpha_*)$ drops off quickly, so the number of randomizations must grow rapidly as α_* increases to ensure AVCI when $V_{\mathrm{HK}}(h) < \infty$. But even so, we still have $v_7(\alpha_*) > v_8(\alpha_*)$ for all α_*, where $v_8(\alpha_*)$ in (58) is the optimal RMSE rate exponent for AVCI obtained under the moment conditions of Corollary 8.

EC.4. Numerical results

The goal of this section is to study numerically the asymptotic results in our paper. We aim to see for various values of $c \in (0,1)$ from Assumption 1.B if a CLT seems to actually hold as $n \to \infty$, 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 (Choi et al. 2020). We consider three different randomized sequences: Sobol' sequences with digital shift (DS) (L'Ecuyer 2018), scrambled Sobol' sequence with linear matrix scrambling (LMS) (Hong and Hickernell 2003, Matousěk 1998), and Lattice rules with random shift (RS) (Cranley and Patterson 1976, Tuffin 1998, L'Ecuyer and Lemieux 2000). 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 three of the cases of Assumption 3.

1. For $u=(u_1,\ldots,u_s)\in[0,1]^s$, function $h_{\mathrm{v}}(u)=\prod_{l=1}^s\left(u_l-\frac{1}{2}\right)^2$ is of bounded Hardy-Krause variation (as h_{v} is a product of functions of bounded variation (Breneis 2020, Section 9.3)), so Assumption 3.A holds.

2. Consider the unidimensional function $g_{\mathrm{b}}(u_1)=\sin(1/u_1)$ for $u_1\in(0,1]$ and $g_{\mathrm{b}}(0)=0$. This function is known to have $V_{\mathrm{HK}}(g_{\mathrm{b}})=\infty$ (e.g., see (Breneis 2020, Example 2.1.9)) and to be bounded. The product function for $u=(u_1,\ldots,u_s)\in[0,1]^s$ is

$$h_{\mathrm{b}}(u)=\prod_{l=1}^s g_{\mathrm{b}}(u_l),$$

which also has $V_{\mathrm{HK}}(h_{\mathrm{b}})=\infty$ but is bounded, so Assumption 3.B holds but not Assumption 3.A.

3. For any fixed $\theta>0$, consider now $g_{\mathrm{ub},\theta}(u_1)=u_1^{-\theta}$ for $u_1\in(0,1]$ and $g_{\mathrm{ub},\theta}(0)=0$. Define

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

for $u\in[0,1]^s$. Function $h_{\mathrm{ub},\theta}$ is unbounded and the moment of order $2+b$ of $h_{\mathrm{ub},\theta}(U)$ for $U\sim U[0,1]^s$ is $\mathbb{E}[(h_{\mathrm{ub},\theta}(U))^{2+b}]=\int_{[0,1]^s}(h_{\mathrm{ub},\theta}(u))^{2+b}\,\mathrm{d}u=\left[\int_0^1(u_1)^{-\theta(2+b)}\,\mathrm{d}u_1\right]^s$, which is finite if and only if $\theta(2+b)<1$. (Our experiments used $\theta=0.35$, so the $2+b$ moment does not exist for $b>0.858$.) In this case, Assumption 3.C holds but not Assumption 3.B.

To test whether a CLT roughly holds for a given (m_n, r_n), we generated 100 independent values of $\widehat{\mu}^{\mathrm{RQ}}_{m_n,r_n}$. We then applied a statistical test on the resulting sample to test for normality. Among existing tests, we chose to use the Shapiro and Wilk (1965) test, 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 (Shapiro et al. 1968). 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 gives strong evidence that the normal property (and the CLT) can be rejected.

EC.4.1. 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 (34). 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 $m^{(i)}$, $1 \le i \le K$, 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 $\nu_i = \ln(\sigma_{m^{(i)}})$ for each $1 \le i \le K$, standard linear regression yields an estimator of α_* as

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

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 $i \in \{1, 2, \ldots, 17\}$. Table EC.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 $h_{\mathrm{ub},0.35}$) in dimensions $s = 3$

Integrand	Dimension s	Estimated α_* Sobol' DS	Sobol' LMS	Lattice RS
h_{v}	3	1.3420	1.3040	1.4218
h_{b}	3	0.6843	0.6735	0.6173
$h_{\mathrm{ub},0.35}$	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
$h_{\mathrm{ub},0.35}$	4	0.6115	0.5624	0.5437

Table EC.1 Estimated values of α_* in (34) from log-log regressions for different integrands, dimensions, and RQMC methods.

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 $h_{\mathrm{ub},0.35}$, which have infinite Hardy-Krause variation, estimated values of α^* are smaller then 1 but larger than 1/2 (see (37)), the exponential rate when using MC methods.

EC.4.2. 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 EC.5 displays, as c increases, the p-values of the Shapiro-Wilk test for h_{v}, h_{b} and $h_{\mathrm{ub},0.35}$ 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 7 apply; thus, taking

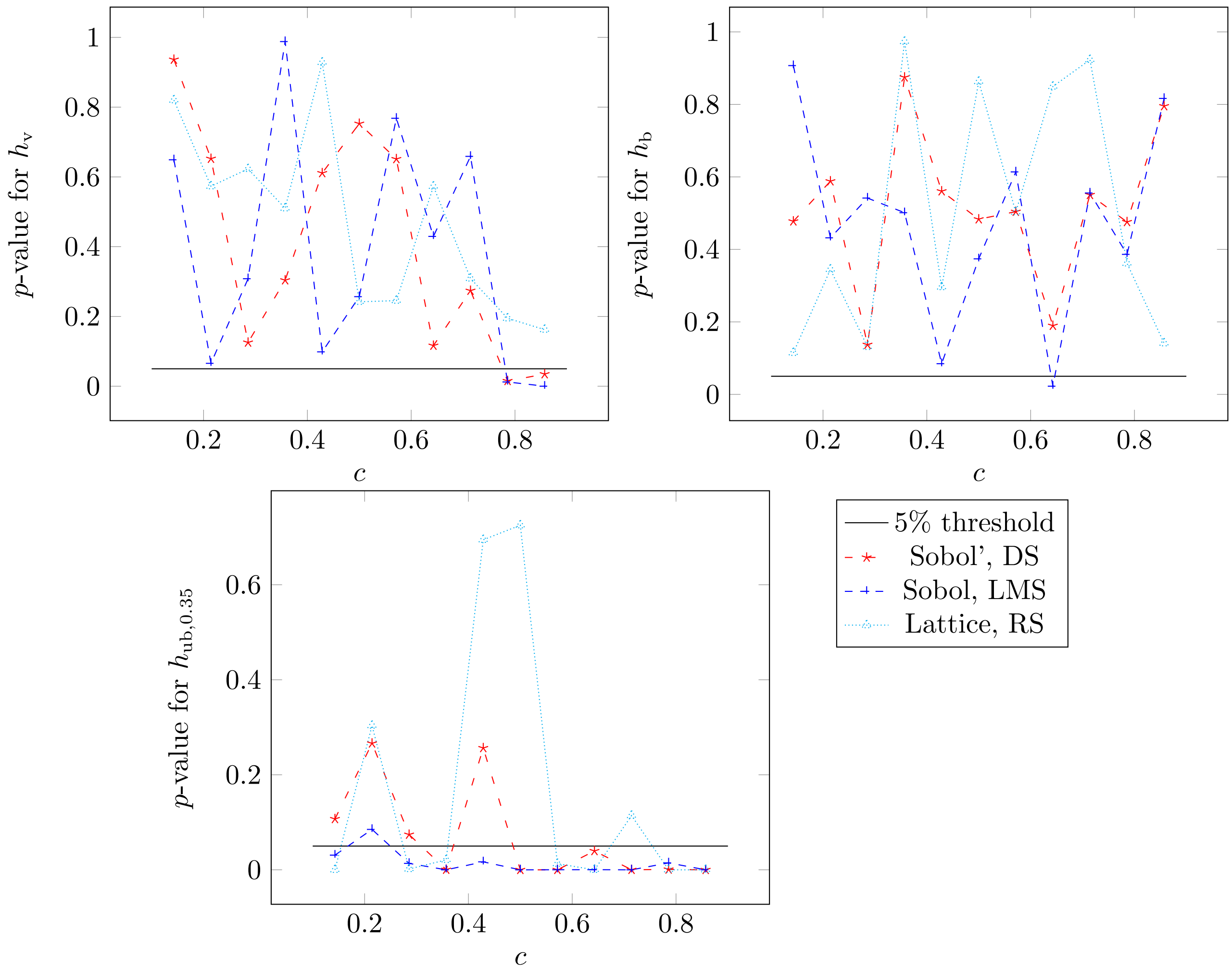


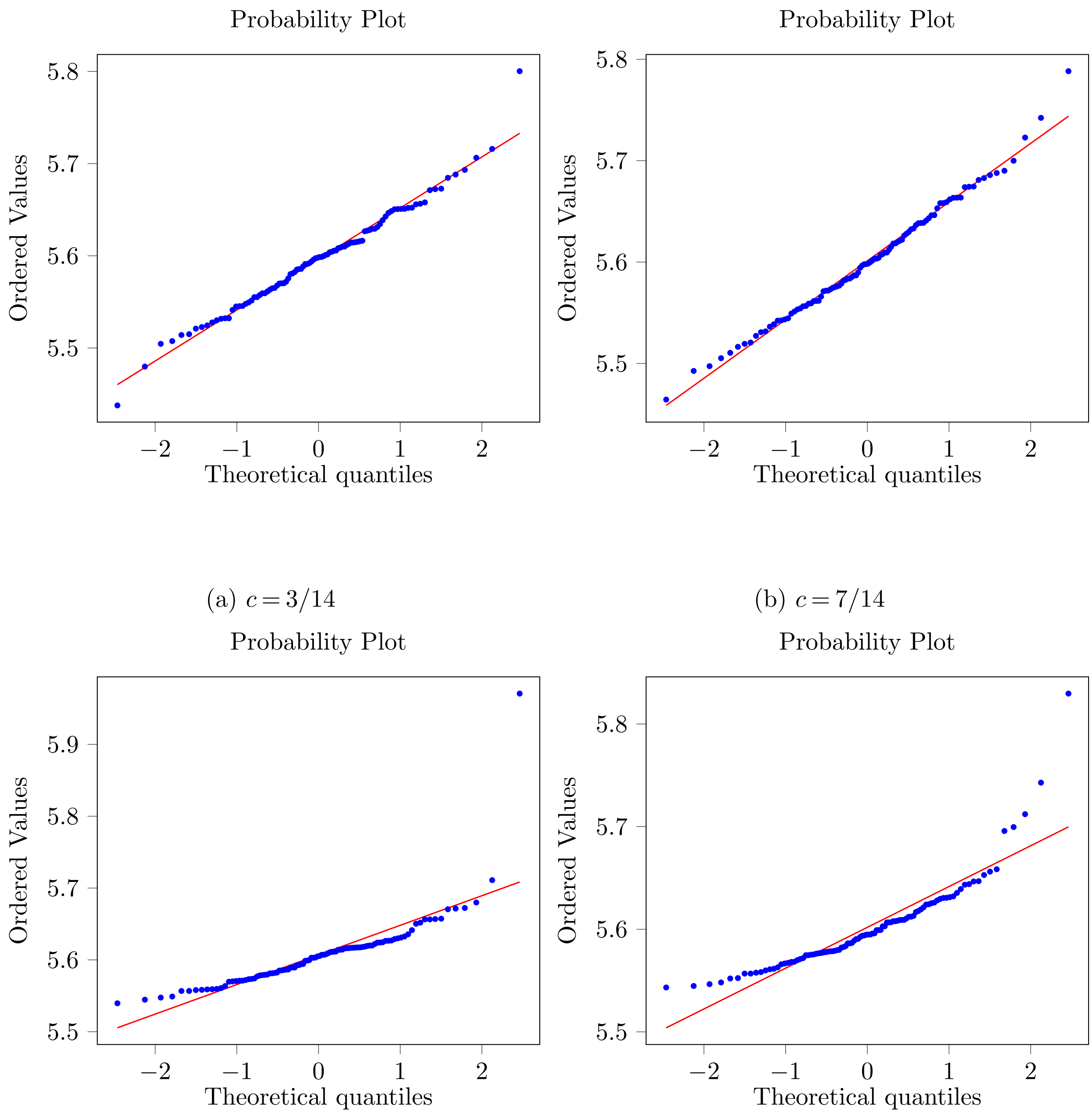
Figure EC.5 p-values of the Shapiro-Wilk test for various values of $c = t/14$ for $2 \le t \le 12$ and for three RQMC techniques with $n = 2^{14}$, from 100 independent estimations. The considered integrand is h_{v} (top left), h_{b} (top right) and $h_{\mathrm{ub},0.35}$ (bottom left), each for dimension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

$\alpha_* = 1.35$ as an approximation from Table EC.1, we get $c_2(\alpha_*) \approx 0.588$ and $c_7(\alpha_*) \approx 0.417$, smaller than the 0.8 from the numerical experiments. This perhaps indicates that the sufficient condition (23) of Corollary 2 is stronger than necessary to ensure a CLT.

- 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 8 apply, and taking $\alpha_* = 0.65$ (approximately what Table EC.1 shows) results in $c_3(\alpha_*) \approx 0.435$ and $c_8(\alpha_*) \approx 0.278$ by (51) and (57). This may indicate that the sufficient condition (24) of Corollary 3 is stronger than necessary to ensure a CLT. Using a larger value of n would be of interest, but the computations already take hours for $n = 2^{14}$.

- For function $h_{\mathrm{ub},0.35}$, p-values are generally very low, except for some c for the lattices with RS. This generally indicates non-normal behavior. The function $h_{\mathrm{ub},0.35}$ 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 EC.1 (for Lattice with RS but using other values do do significantly change the results) gives $c_4(0.5437, 1/0.35 - 2) \approx 0.216$ from (53). 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 8 is not applicable.)

As further illustration, we provide the Q-Q plots in Figure EC.6 for several values of c when $n = 2^{14}$ for function $h_{\mathrm{ub},0.35}$ 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 EC.5. In Figure EC.6, all but the upper left 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.



(a) $c = 3/14$

(b) $c = 7/14$

(c) $c = 11/14$

(d) $c = 12/14$

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