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

MARVIN K. NAKAYAMA, New Jersey Institute of Technology, USA
BRUNO TUFFIN, Inria, Univ Rennes, CNRS, IRISA, France

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 for triangular arrays 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.

CCS Concepts: • **Mathematics of computing → Numerical analysis**; **Probabilistic algorithms**; • **Computing methodologies → Simulation evaluation**.

Additional Key Words and Phrases: randomized quasi-Monte Carlo, central limit theorems, confidence intervals

**ACM Reference Format:**
Marvin K. Nakayama and Bruno Tuffin. 2023. Sufficient Conditions for Central Limit Theorems and Confidence Intervals for Randomized Quasi-Monte Carlo Methods. *ACM Trans. Model. Comput. Simul.* 00, 0, Article 0 (January 2023), 38 pages. https://doi.org/XXXXXXX.XXXXXXX

## 1 INTRODUCTION

Research and analysis of complicated problems across diverse fields of science, engineering, business, etc., often entail computing integrals, as in molecular dynamics, queueing systems, or pricing of financial instruments. The integral frequently corresponds to the mean of a stochastic model, and model complexity usually precludes analytically evaluating the integral, so numerical methods are employed. Monte Carlo (MC) methods are computational algorithms based on random sampling that can be applied to estimate a mean (integral); see, e.g., [1, 17] 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) assumed to hold as the sample size $n \to \infty$.

Authors' addresses: Marvin K. Nakayama, marvin@njit.edu, New Jersey Institute of Technology, Computer Science Department, Newark, New Jersey, USA, 07102; Bruno Tuffin, Inria, Univ Rennes, CNRS, IRISA, Campus de Beaulieu, 263 Avenue Général Leclerc, 35042 Rennes, France, bruno.tuffin@inria.fr.


ACM 1049-3301/2023/1-ART0
https://doi.org/XXXXXXX.XXXXXXX

Quasi-Monte Carlo (QMC) methods [27, 31] 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 [15], the most famous being the Koksma-Hlawka inequality [31, Section 2.2]. A drawback of QMC stems from the fact that such error bounds, while theoretically valuable in proving asymptotic properties, are in most cases practically useless, being difficult to compute and grossly overestimating the error for a fixed sample size and specific integrand; see, e.g., [41, 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, [19, 22, 27, 41]. Several types of randomization exist, the main ones being the *random shift* [7], which translates all points of the low-discrepancy sequence by the *same* uniform random vector; the *digital shift* [19], 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 [35, 36].

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, [23] 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 unsteady ground. A CLT, though, has been verified by [28] (as $m \to \infty$ for fixed $r \geq 1$) in the case of digital scrambling of a particular type called *nested uniform scrambling*; [2, 13] similarly prove CLTs for related scrambles. But this scrambling is computationally expensive, which has limited its adoption in practice in the past, although this may be less of an issue with today’s more powerful computers. The last paragraph of Section 17.6 of [37] states, “It seems likely that the other scrambles considered [in his book] do not satisfy a central limit theorem” for a single or fixed number of randomizations.

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)? The ordinary (Lindeberg-Lévy) CLT [3, Theorem 27.1] typically does not apply in our setting because the distribution of the estimator from a single randomization changes as $m$ varies. Instead, we formulate the problem using a so-called *triangular array*, for which a CLT can be obtained under the Lindeberg condition [3, 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 is usually unknown. 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 [3, 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. Appendices contains all proofs (Appendix A), as well as additional analysis when $(m,r) = (n^c, n^{1-c})$ (Appendices B and C) and numerical results (Appendix D). The present work extends our conference paper [30], which presents many of the theorems, corollaries and propositions (but all without proofs), as well as Figures 1, 3, and 4, which are in Appendix C of the current paper. Moreover, [30] focus mainly on the case of $(m,r) = (n^c, n^{1-c})$, whereas the current paper also develops a more general setting.

## 2 NOTATION/FRAMEWORK

Our goal is to estimate

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

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.

Our paper will consider limiting regimes, often in which $n \to \infty$, and we adopt the following asymptotic notation. Consider univariate real-valued functions $f$ and $g$. We write $f(n) = O(g(n))$ (resp., $f(n) = \Theta(g(n))$) as $n \to \infty$ to mean that there exist positive constants $a_0$, $a_1$, and $n_0$ such that $|f(n)| \leq a_1|g(n)|$ (resp., and also $|f(n)| \geq a_0|g(n)|$) for all $n > n_0$. Also, $f(n) = o(g(n))$ as $n \to \infty$ denotes that $f(n)/g(n) \to 0$ as $n \to \infty$, and $f(n) = \omega(g(n))$ as $n \to \infty$ means that $f(n)/g(n) \to \infty$ as $n \to \infty$ (i.e., $g(n) = o(f(n))$).

The following subsections will describe methods for estimating $\mu$ via MC, QMC, and RQMC.

### 2.1 Monte Carlo

The (standard) MC estimator of $\mu$ 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

$$\mathrm{RMSE}\left[\widehat{\mu}_n^{\mathrm{MC}}\right] \equiv \sqrt{\mathbb{E}[(\widehat{\mu}_n^{\mathrm{MC}} - \mu)^2]} = \frac{\psi}{\sqrt{n}} \tag{1}$$

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 $\mu$ that we consider. Moreover, the MC estimator $\widehat{\mu}_n^{\mathrm{MC}}$ obeys a (Gaussian) CLT (e.g., [3, Theorem 27.1]): $\sqrt{n}[\widehat{\mu}_n^{\mathrm{MC}} - \mu]/\psi \Rightarrow \mathcal{N}(0,1)$ as $n \to \infty$, where $\Rightarrow$ denotes convergence in distribution (which is equivalent to convergence in probability when the limit is deterministic; [3, Theorem 25.3]), and $\mathcal{N}(a_1, a_2^2)$ is a normal random variable with mean $a_1$ and variance $a_2^2$. The CLT provides a simple way to construct a (probabilistic) measure of the estimator's error through a confidence interval for $\mu$. 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 $\widehat{\psi}_n^2 = [1/(n-1)]\sum_{i=1}^n [h(U_i) - \widehat{\mu}_n^{\mathrm{MC}}]^2$ is an unbiased estimator of $\psi^2$, $0 < \gamma < 1$ is the specified confidence level (e.g., $\gamma = 0.95$), $z_\gamma$ satisfies $\Phi(z_\gamma) = 1 - (1-\gamma)/2$, and $\Phi$ is the $\mathcal{N}(0,1)$ *cumulative distribution function* (CDF). Then $I_{n,\gamma}^{\mathrm{MC}}$ is an asymptotically valid CI for $\mu$ in the sense that $\lim_{n\to\infty} \mathbb{P}(\mu \in I_{n,\gamma}^{\mathrm{MC}}) = \gamma$ [5, 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* [27, Chapter 5]), leading QMC to approximate $\mu$ 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., [27, Section 5.6.1] or [31, Theorem 2.11]) 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 $\Xi$, 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 $\Xi$. We typically have

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

Thus, when $V_{\mathrm{HK}}(h) < \infty$, which we denote by $h \in \mathrm{BVHK}$ (i.e., bounded variation in the sense of Hardy and Krause), 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)).

Unfortunately, (2) does not yield a practical error bound for QMC. Indeed, $V_{\mathrm{HK}}(h)$ is the sum of $2^s - 1$ Vitali variations [31], each at least as difficult to compute as $\mu$, and the resulting $V_{\mathrm{HK}}(h)$ is additionally potentially very large even for moderate dimensions $s$. 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, [14, 15, 18, 26, 31]), 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 $\Xi$, 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 $U_{i,j}' \in [0,1]^s$ be the $i$-th point of the $j$-th i.i.d. 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. The independence across the $r$ randomizations ensures that $X_j$, $j = 1, 2, \ldots, r$, are i.i.d. From their sample variance $\widehat{\sigma}_{m,r}^2 =$

$\sum_{j=1}^{r}(X_j - \widehat{\mu}_{m,r}^{\mathrm{RQ}})^2/(r-1)$ when $r \geq 2$, we then obtain a potential $\gamma$-level CI for $\mu$ as $I_{m,r,\gamma}^{\mathrm{RQ}} \equiv [\widehat{\mu}_{m,r}^{\mathrm{RQ}} \pm z_\gamma \widehat{\sigma}_{m,r}/\sqrt{r}]$. The hope is that as $m$ or $r$ (or both) grows large, the overall RQMC estimator $\widehat{\mu}_{m,r}^{\mathrm{RQ}}$ obeys a Gaussian CLT (see Section 3) and $I_{m,r,\gamma}^{\mathrm{RQ}}$ is an AVCI (see Section 4). (Section 2.4 discusses how $(m,r) = (m_n, r_n)$ are chosen as functions of an overall computing budget $n$ for the total number of evaluations of $h$, but we use the simpler notation $(m,r)$ here to introduce the ideas before adopting the more complicated notation $(m_n, r_n)$ later.) We next describe some existing 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 $\Xi$, modulo 1, coordinate-wise [7]. Formally, using the first $m$ points of the low-discrepancy sequence $\Xi$ 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 $\mu$ 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 [39, Theorem 2]. As a consequence, if $h \in$ BVHK, 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 $h$.

The convergence speed can even be faster for special classes of functions and specific sequences $\Xi$ called *lattice rules* ([21, 40]) 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 $|\hat{f}(g)| \leq C t(g)^{-\alpha}$ for all $g \in \mathbb{Z}^s$. Then, for each $\alpha > 1$, $C > 0$ and $m \geq 1$, there exists a lattice rule $\Xi^{[m]} = (\xi_i^{[m]})_{i=1,\ldots,m}$ (depending on $m$, as well as $\alpha$, $C$, 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 [23]. 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 $r = 1$), [23] argues 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 [34, 35] introduces *scrambled digital nets* as another form of RQMC. The method scrambles the digits of special low-discrepancy sequences, namely, digital nets [31]. 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 [31, Chapter 4], 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 $\pi_0^k$, $\pi^k_{0,U_i^{(k,1),\prime},\ldots,U_i^{(k,\ell-1),\prime}}$, $\forall k, \ell$, 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 $\xi_i$ is finite, meaning that the required number of permutations is finite, even if large. Other more computationally efficient (but theoretically less powerful) forms of scrambling appear in [16, 19], and [29], including the *linear matrix scramble*.

Scrambled digital nets keep the discrepancy property of the original digital net. Specifically, let $\Xi_\Pi$ denote the scrambling of a digital net $\Xi$ 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 ([19, 34])

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

so the estimator $X_j$ in (4) from a single randomization converges at the same speed as (7) when the integrand $h \in$ BVHK. For special classes of $h$ [36], the RMSE of the quadrature rule based on nested uniform scrambling can be as small as $O\left(m^{-3/2}(\ln m)^{(s-1)/2}\right)$; Appendix B.5 describes similar results.

For RQMC via digital nets and Owen's full nested scrambling, [28] establishes a Gaussian CLT for $r = 1$ as $m \to \infty$ (so the number of involved independent permutations also tends to infinity); [2] and [13] establish extensions. But these CLTs are restricted to this specific scrambling, whose large computational cost has limited its use in practice in the past (although this may be less of an issue with today's more powerful computers); more general CLTs are needed for other forms of RQMC. Another drawback of the CLT of [28] is that it applies to nested scrambling to only a particular class of digital sequences with so-called $t$-value of 0, where lower values of $t$ ensure better equidistribution; see [37, Section 15.7] for details. But restricting to $t = 0$ rules out the popular Sobol' sequence except in very small dimensions $s$. While $t = 0$ allows for Faure sequences, empirical studies seem to indicate that these produce less accurate approximations to $\mu$ than Sobol' sequences.

*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 ($1 \leq k \leq s$) 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 single 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 $U'_{i,j} = (U^{(1),\prime}_{i,j}, \ldots, U^{(s),\prime}_{i,j})$ from the $j$-th randomization, its $k$-th coordinate is defined by

$$U^{(k),\prime}_{i,j} = \sum_{\ell=1}^{\infty} U^{(k,\ell),\prime}_{i,j} b_0^{-\ell}, \quad \text{where} \quad U^{(k,\ell),\prime}_{i,j} = (\xi_i^{(k,\ell)} + U_j^{(k,\ell)}) \bmod b_0.$$

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

$$D^*_m(\Xi^{\mathrm{Dig}}_{U_j}) \le Cm^{-1}(\ln m)^s. \tag{9}$$

### 2.4 Assumptions and Preliminary Results

Recall that for a given computation budget of about $n$ integrand evaluations, we define the RQMC estimator in (4) with $(m, r)$, where $r$ is the number of randomizations and $m$ is the number of points used from each randomized sequence, so the total number of evaluations of the integrand $h$ is $mr$. For a fair comparison with MC, which uses $n$ evaluations of $h$ when the sample size is $n$, we will assume that RQMC has $mr \approx n$, which will be more precisely expressed below. 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 $\mu$ in (4) becomes

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

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 $\widehat{\mu}^{\mathrm{RQ}}_{m_n,r_n}$ obeys a CLT (with Gaussian limit) as $n \to \infty$, and Section 4 derives conditions for an asymptotically valid CI for $\mu$. Other papers (e.g., [8, 11]) adopt a framework similar to Assumption 1.A to study MC methods for analyzing steady-state behavior of a stochastic model (for example, batching with $r_n$ batches, each of size $m_n$).

Assumption 1.A requires $r_n \to \infty$ as $n \to \infty$ because otherwise, a Gaussian CLT may not hold. As noted earlier, [23] 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 $r \ge 1$ (see the discussion at the end of Section 2.3.1), and the only Gaussian CLTs that the authors are aware of are only for the (computationally expensive) nested digital scrambling [2, 13, 28] (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 $m_0 \ge 1$, 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 $c \in (0, 1)$. In particular, $r_n = n^{1-c} \to \infty$ precludes the setting of [23] in which a Gaussian CLT does not hold.
- 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 $(m_n, r_n) = (a_1+d_1 n^c, a_2+d_2 n^{1-c})$ for nonnegative constants $a_1$ and $a_2$ and 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. But allocations of this form can permit the number $m_n$ of points from the QMC sequence to grow more rapidly than $n^c$ for any $c > 0$ as $n \to \infty$, so Assumption 1.A enables taking greater advantage of QMC's fast convergence rate than is possible under Assumption 1.B.
- While Assumption 1.B specifies that $c < 1$, we could also consider the case of $c = 1$, which corresponds to a single randomization or a fixed number $r_0$ of randomizations. Then for fully nested scrambling [2, 13, 28], each randomization satisfies a CLT as $m_n \to \infty$, so an asymptotically valid Student-$t$ confidence interval could be constructed when $r_0 \geq 2$; see Appendix B.5 for related discussions.

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 value of $c$ that maximizes the convergence rate for RMSE$[\widehat{\mu}^{\mathrm{RQ}}_{m_n,r_n}]$ as $n \to \infty$.

Let $\Xi'$ 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 $h \in$ BVHK.

Assumption 2. *For the RQMC method used, there exists a constant $0 < w_0' < \infty$ such that each randomized sequence $\Xi'$ satisfies $D^*_m(\Xi') \leq w_0' m^{-1} (\ln m)^s$ for all $m > 1$, where $w_0'$ depends on the RQMC method but not on the randomization's realization (e.g., of the random uniforms or permutations).*

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 $D^*_m(\Xi) \leq w_0 m^{-1} (\ln m)^s$ whenever $m > 1$. Hence, Assumption 2 holds with $w_0' = 4^s w_0$ for a random shift by (6) (whatever the considered low-discrepancy sequence), and with $w_0' = C$ in (8) and (9) for scrambling and a digital shift.

Our results will further depend on properties of the integrand $h$. We will often consider four alternative conditions on $h$, presented in order of decreasing strength (see Proposition 2.1 below). For fixed $p \geq 1$, we write $h \in \mathcal{L}^p$ when $\int_{[0,1]^s} |h(u)|^p \,\mathrm{d}u < \infty$, and $h \in \mathcal{L}^\infty$ when $h$ is bounded (almost everywhere).

Assumption 3.A. $h \in$ BVHK.

Assumption 3.B. $h \in \mathcal{L}^\infty$.

Assumption 3.C. *$h \in \mathcal{L}^{2+b}$ for some $b > 0$, which ensures that for $U \sim \mathcal{U}[0,1]^s$,*

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

Assumption 3.D. *$h \in \mathcal{L}^2$; i.e., $h(U)$ has finite variance.*

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 2.1. *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 [39, Theorem 2].

Lemma 2.2. *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 $h \in$ BVHK 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 Appendix A) of Proposition 2.1 notes that $V_{\mathrm{HK}}(h) = \infty$ for $s \geq 2$ if $h$ is an indicator function (so $\mu$ 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 Appendix A.)

Lemma 2.3. *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 $\mu$ can be constructed. As in (12), when $h \in$ BVHK 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 [23], as noted earlier (Section 2.3.1).

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 $\mu$. 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* [3, p. 359]. In a triangular array, also called a *double array* (e.g., [38, 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

$$\sigma^2_{n,j} = \mathrm{Var}[X_{n,j}] \equiv \sigma^2_{m_n}, \tag{13}$$

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

$$X_{n,1}, X_{n,2}, \ldots, X_{n,r_n} \text{ are i.i.d., each with some distribution } F_n. \tag{14}$$

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

Although (14) has that the $r_n$ random variables are i.i.d. for each $n$, the general setting for a triangular array, as in [3, 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 [3, 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,

$$\frac{\bar{X}_n - \mu_n}{\sqrt{\mathrm{Var}[\bar{X}_n]}} \Rightarrow \mathcal{N}(0,1), \qquad \text{as } n \to \infty \tag{15}$$

provided that the *Lindeberg condition* holds:

$$\frac{\tau^2_{n,1}(t) + \cdots + \tau^2_{n,r_n}(t)}{\sigma^2_{n,1} + \cdots + \sigma^2_{n,r_n}} \to 0, \qquad \text{as } n \to \infty, \ \forall t > 0. \tag{16}$$

The Lindeberg condition (16) implies that $\max_{1 \leq j \leq r_n} \sigma^2_{n,j}/s_n^2 \to 0$ as $n \to \infty$ (e.g., [25, 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., $X_{n,j} \equiv 0$, $2 \leq j \leq r_n$, so the left side of (15) 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 (16) to study the RQMC estimator $\widehat{\mu}^{\mathrm{RQ}}_{m_n,r_n}$ in (10). By (14),

$$s_n^2 = \sigma^2_{n,1} + \cdots + \sigma^2_{n,r_n} = r_n \sigma^2_{m_n}. \tag{17}$$

Denote the distribution of $Y_{n,j} = X_{n,j} - \mu$ by $G_n$, which does not depend on $j$ by (14). Note that $\sigma^2_{m_n} = \int y^2 \, \mathrm{d}G_n(y)$, and let

$$\tau_n^2(t) = \int_{|y|>ts_n} y^2 \, \mathrm{d}G_n(y), \qquad \text{for } t > 0. \tag{18}$$

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. $\sigma^2_{m_n} > 0$ *for all n large enough.*

We omit the proof of the following, which specializes for RQMC the condition (16) as (19) below.

Theorem 3.1. *Suppose that Assumptions 1.A and 4 hold. If the Lindeberg condition*

$$\frac{\tau_n^2(t)}{\sigma^2_{m_n}} \to 0, \qquad \text{as } n \to \infty, \ \forall t > 0 \tag{19}$$

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

$$\frac{\widehat{\mu}^{RQ}_{m_n,r_n} - \mu}{\sigma_{m_n}/\sqrt{r_n}} \Rightarrow \mathcal{N}(0,1), \qquad \text{as } n \to \infty. \tag{20}$$

From [3, p. 361], condition (19) is even necessary and sufficient for the CLT (20) since for all $j$, $\sigma^2_{n,j}/s_n^2 = 1/r_n \to 0$ as $n \to \infty$ by Assumption 1.A. The Lindeberg condition (19), which imposes restrictions on the tail behavior of $G_n$ through (18), holds under a Lyapounov (moment) condition [3, Theorem 27.3], which (21) below adapts for our RQMC setting.

Theorem 3.2. *Suppose that Assumptions 1.A and 4 hold. Further suppose that there exists d > 0 such that* $\mathbb{E}\left[\left|X_{n,1} - \mu\right|^{2+d}\right] < \infty$ *for each n satisfying Assumption 4 and that*

$$\frac{\mathbb{E}\left[\left|X_{n,1} - \mu\right|^{2+d}\right]}{r_n^{d/2} \sigma^{2+d}_{m_n}} \to 0, \qquad \text{as } n \to \infty. \tag{21}$$

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

### 3.1 Corollaries of Theorems 3.1 and 3.2

We now develop various sufficient conditions that secure CLT (20) through Theorem 3.1 or 3.2. Section 3.2 will compare the conditions under a general allocation $(m_n, r_n)$ of Assumption 1.A, with Table 1 in Section 4.2 outlining the results (also including those for AVCI in Section 4.1). Later in Section 5.8, Table 2 summarizes and compares all of our corollaries under the simple allocations $(m_n, r_n) = (n^c, n^{1-c})$ of Assumption 1.B; Figures 1, 2, and 3 in Appendix C provide graphical comparisons. Our results provide conditions on $(m_n, r_n)$ in terms of $\sigma_{m_n}$, whose exact asymptotic behavior has not been established in the literature for most RQMC methods. One exception that we are aware of is for nested uniform scrambling (Section 2.3.2) under certain restrictions, which we discuss in Appendix B.5. For some other RQMC methods, Appendix D provides numerical results studying the behavior of $\sigma_{m_n}$.

In the following corollary, we define $A_m = \frac{1}{m} \sum_{i=1}^{m} h(U'_i)$ as an estimator based on $m$ points from a single randomization $\Xi' = (U'_i)_{i\geq 1}$, and $\sigma^2_m = \mathrm{Var}[A_m]$. This is to emphasize the dependence of these quantities on the point-set size $m$ but not on the budget $n$ nor the allocation $(m_n, r_n)$.

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 $d > 0$ and $k_1 \equiv k_{1,d} \in (0, \infty)$ such that $\mathbb{E}\left[|A_m - \mu|^{2+d}\right] < \infty$ and*

$$\frac{\mathbb{E}\left[|A_m - \mu|^{2+d}\right]}{\sigma_m^{2+d}} \leq k_1, \qquad \textit{for all } m \textit{ sufficiently large.} \tag{22}$$

*Then the Lyapounov condition* (21) *and CLT* (20) *hold. A sufficient condition for* (22) *is that there exists a constant $k_2 \in (0, \infty)$ such that*

$$\mathbb{P}\left(|A_m - \mu| \leq k_2 \sigma_m\right) = 1, \qquad \textit{for all } m \textit{ sufficiently large.} \tag{23}$$

Condition (22) (resp., (23)) 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 $\sigma_m$ for all large $m$. While Corollary 1 requires $m_n \to \infty$, (22) and (23) do not depend otherwise on the allocation $(m_n, r_n)$, so Corollary 1 can fulfill Assumption 1.A with $r_n$ growing slowly to $\infty$ 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 $n$. However, [37, Chapter 17 end notes] states that under a linear matrix scramble [29], "There is reason to believe that the skewness and kurtosis ... could diverge" as $m \to \infty$ for some integrands $h$, in which case (22) may not hold for $d = 1$ and 2. Appendix D.4 investigates numerically the validity of Condition (22) and illustrates that it appears to be satisfied in certain (but not all) settings.

Establishing (22) or (23) may be difficult, so we next provide other conditions that can be more readily verifiable to ensure CLT (20). 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 2.1).

Corollary 2. *Suppose that Assumption 1.A holds with $m_n > 1$ for all $n$ large enough, along with Assumptions 2, 3.A ($h \in$ BVHK), and 4. Also, suppose that*

$$r_n^{1-\lambda} \left(\frac{m_n \sigma_{m_n}}{(\ln m_n)^s}\right)^2 \to \infty, \qquad \textit{as } n \to \infty, \tag{24}$$

*for some constant $\lambda \in (0, 1)$, which can be chosen arbitrarily small. Then the Lyapounov condition* (21) *and CLT* (20) *hold.*

The following corollary of Theorem 3.1 considers the case when the integrand $h$ is a bounded function.

Corollary 3. *Suppose that Assumptions 1.A, 3.B ($h \in \mathcal{L}^\infty$), and 4 hold. If*

$$s_n^2 = r_n \sigma_{m_n}^2 \to \infty, \qquad \text{as } n \to \infty, \tag{25}$$

*then the Lindeberg condition* (19) *and CLT* (20) *hold.*

As with Corollary 2, the next corollary follows from Theorem 3.2, but it does not require $h \in$ BVHK, precluding the application of Lemma 2.2. It instead employs Lemma 2.3, so it assumes a moment condition on $h(U)$ (Assumption 3.C).

Corollary 4. *Suppose that Assumptions 1.A, 3.C ($h \in \mathcal{L}^{2+b}$ for some $b > 0$), and 4 hold. Also, suppose that*

$$r_n^{b/(2+b)} \sigma_{m_n}^2 \to \infty, \qquad \text{as } n \to \infty. \tag{26}$$

*Then the Lyapounov condition* (21) *and CLT* (20) *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 3.1 to consider the case when $m_n$ is fixed.

Corollary 5. *Suppose that Assumption 3.D ($h \in \mathcal{L}^2$) holds. If*

$$(m_n, r_n) = \left(m_0, \left\lfloor \frac{n}{m_0} \right\rfloor\right) \text{ for all } n\text{, for some fixed } m_0 \geq 1 \text{ with } \sigma_{m_0}^2 > 0, \tag{27}$$

*which implies Assumptions 1.A and 4, then the Lindeberg condition* (19) *and CLT* (20) *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 (20). Proposition 2.1 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 $(m_n, r_n)$, demonstrating tradeoffs in our assumptions. (Section 5.8 will provide further comparisons under Assumption 1.B.)

Proposition 3.3. *If Assumption 1.A holds, then condition* (27) *of Corollary 5 implies condition* (26) *of Corollary 4, and* (26) *in turn implies condition* (25) *of Corollary 3. If in addition*

$$\frac{1}{r_n^\lambda} \left( \frac{m_n}{(\ln m_n)^s} \right)^2 \to d_0 \in (0, \infty], \qquad \text{as } n \to \infty, \text{ for } \lambda \in (0,1), \tag{28}$$

*then* (25) *implies condition* (24) *of Corollary 2. Condition* (28) *holds, e.g., under Assumption 1.B.*

The condition (28) in Proposition 3.3 specifies that $r_n$ does not grow too quickly (as $n \to \infty$) compared to $m_n$. While (28) is always true under Assumption 1.B, it does not hold, e.g., for fixed $m_n \equiv m_0$, as in condition (27) 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 enhance the CLTs in Section 3 to construct an asymptotically valid CI for $\mu$ 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 (14), and their sample variance

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

is an unbiased estimator of $\sigma^2_{m_n}$ in (13). For a given desired confidence level $100\gamma\%$, with $0 < \gamma < 1$, we then consider a CI for $\mu$ as

$$I^{\mathrm{RQ}}_{m_n,r_n,\gamma} \equiv \left[\widehat{\mu}^{\mathrm{RQ}}_{m_n,r_n} \pm z_\gamma \widehat{\sigma}_{m_n,r_n}/\sqrt{r_n}\right]. \tag{29}$$

Our goal now is to provide conditions ensuring that $I_{n,\gamma}$ is an AVCI, i.e., (31) below holds.

Theorem 4.1. *Suppose that Assumptions 1.A and 4 hold. If the CLT* (20) *holds and if*

$$\frac{\widehat{\sigma}^2_{m_n,r_n}}{\sigma^2_{m_n}} \Rightarrow 1, \qquad \text{as } n \to \infty, \tag{30}$$

*then*

$$\lim_{n\to\infty} \mathbb{P}(\mu \in I^{RQ}_{m_n,r_n,\gamma}) = \gamma. \tag{31}$$

We now want a sufficient condition ensuring that (30) holds to secure AVCI (31).

Theorem 4.2. *Suppose that Assumptions 1.A and 4 hold. If* $\mathbb{E}\left[(X_{n,1}-\mu)^4\right] < \infty$ *and*

$$\frac{\mathbb{E}\left[(X_{n,1}-\mu)^4\right]}{r_n\sigma^4_{m_n}} \to 0, \qquad \text{as } n \to \infty, \tag{32}$$

*then the CLT* (20), (30), *and AVCI* (31) *hold.*

To summarize, we see that Theorem 4.1 imposes an extra condition (30) to those ensuring a CLT through Theorems 3.1 or 3.2 to further guarantee an AVCI. Theorem 4.2 then shows that assuming the Lyapounov condition (21) of Theorem 3.2 holds for $d = 2$ (i.e., (32)) is sufficient to secure (30).

### 4.1 Specific Sufficient Conditions for AVCI

We next establish AVCI (31) (often through Theorem 4.2) under various conditions.

Corollary 6. *Suppose that Assumptions 1.A and 4 hold. If* (22) *holds with $d = 2$, then the CLT* (20), (30), *and AVCI* (31) *hold. A sufficient condition to ensure* (22) *with $d = 2$ is that* (23) *holds.*

By Corollaries 1 and 6, the condition (23) ensures both the CLT (20) and AVCI (31). But as we will see in Section 5, securing AVCI often uses stronger conditions than ensuring a CLT.

We next establish AVCI (31) under alternative Assumptions 3.A ($h \in$ BVHK), 3.C ($h \in \mathcal{L}^{2+b}$), and 3.D ($h \in \mathcal{L}^2$), which were previously used to establish CLTs in Corollaries 2, 4, and 5, respectively. The following AVCI result imposes Assumption 3.A, with the condition (33) below replacing (24) of Corollary 2.

Corollary 7. *Suppose that Assumption 1.A holds with $m_n > 1$ for all $n$ large enough, along with Assumptions 2, 3.A ($h \in$ BVHK), and 4. If*

$$r_n \left(\frac{m_n\sigma_{m_n}}{(\ln m_n)^s}\right)^4 \to \infty, \qquad \text{as } n \to \infty, \tag{33}$$

*then the CLT* (20), (30), *and AVCI* (31) *hold.*

As with Corollary 4, the following result uses Assumption 3.C ($h \in \mathcal{L}^{2+b}$) but requires $b = 2$ rather than $b > 0$, which also corresponds to replacing condition (26) with (34) below.

Corollary 8. *Suppose that Assumptions 1.A and 4 hold, along with Assumption 3.C ($h \in \mathcal{L}^4$). If*

$$r_n\sigma^4_{m_n} \to \infty, \qquad \text{as } n \to \infty, \tag{34}$$

*then the CLT* (20), (30), *and AVCI* (31) *hold.*

The next result considers the case when $m_n \equiv m_0$ is fixed, as in condition (27) of Corollary 5.

COROLLARY 9. *The same conditions as in Corollary 5 imply that* (30) *and AVCI* (31) *hold.*

Corollaries 5 and 9 assume the same conditions, where the former establishes the CLT (20), and the latter proves that (29) is AVCI (31). 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 (31) 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 (31). By Proposition 2.1, Assumption 3.A ($h \in$ BVHK) in Corollary 7 is strictly stronger than Assumption 3.C ($h \in \mathcal{L}^4$) in Corollary 8, and the latter restriction is strictly stronger than Assumption 3.D ($h \in \mathcal{L}^2$) of 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 4.3. *Under Assumption 1.A, condition* (27) *in Corollary 9 implies condition* (34) *of Corollary 8. If also $m_n > 1$ for all $n$ sufficiently large, then* (34) *implies condition* (33) *of Corollary 7.*

Table 1 summarizes the findings of Propositions 3.3 and 4.3. Comparing the conditions securing CLT (20) and AVCI (31) provides further insights. Corollary 1 (resp., 6) ensures the CLT (resp., AVCI) when condition (22) holds for $d > 0$ (resp., $d = 2$), 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 (23). Corollaries 5 and 9 assume the same conditions, so when $m_n = m_0$ is fixed, AVCI (31) does not require any additional conditions beyond that for the CLT (20). Theorem 4.2 and Corollaries 7 and 8 provide sufficient conditions for AVCI, which also imply CLT (20). Ideally, we would have that AVCI is true whenever the CLT holds without any additional restrictions (although this may not be possible, e.g., [9, Example 3.4.13] provides an example of i.i.d. summands having infinite variance, and a CLT holds but with a nonstandard scaling). To study this, we could compare (the conditions of) Theorem 3.2 to Theorem 4.2, 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 (see Table 2), which will lead to simpler and more intuitive results. Compared to the broad generality of Assumption 1.A used by Table 1, the additional structure of Assumption 1.B enables sharper conclusions in Table 2.

# 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 (20) or AVCI (31), and which of those $c$ lead to the best rates of convergence for $\mathrm{RMSE}[\widehat{\mu}^{\mathrm{RQ}}_{m_n,r_n}]$ as $n \to \infty$. Table 2 at the end of this section will summarize the results under Assumption 1.B.

With RQMC, we typically have $\sigma_m$, defined before Corollary 1, is $O(m^{-\alpha})$ as $m \to \infty$ with $\alpha > 1/2$ (e.g., see (12) when $h \in$ BVHK). 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.

Throughout the remainder of this section, we will make the following assumption.

| Cor. $k$ | Ensures | Assumption on only $h$ | Other Key Condition | Implications |
|---|---|---|---|---|
| 1 | **CLT** | | Eq. (22) for some $d > 0$ | |
| 2 | **CLT** | 3.A ($h \in$ BVHK) | Eq. (24) | |
| 3 | **CLT** | 3.B ($h \in \mathcal{L}^\infty$) | Eq. (25) | $[(25) \wedge (28)] \Rrightarrow (24)$ |
| 4 | **CLT** | 3.C ($h \in \mathcal{L}^{2+b}$ for some $b > 0$) | Eq. (26) | $(26) \Rrightarrow (25)$ |
| 5 | **CLT** | 3.D ($h \in \mathcal{L}^2$) | Eq. (27) | $(27) \Rrightarrow (26)$ |
| 6 | **AVCI** | | Eq. (22) for $d = 2$ | |
| 7 | **AVCI** | 3.A ($h \in$ BVHK) | Eq. (33) | |
| 8 | **AVCI** | 3.C ($h \in \mathcal{L}^4$) | Eq. (34) | $(34) \Rrightarrow (33)$ |
| 9 | **AVCI** | 3.D ($h \in \mathcal{L}^2$) | Eq. (27) | $(27) \Rrightarrow (34)$ |

Table 1. Summary of Corollaries 1–9 under Assumption 1.A for any allocations $(m_n, r_n)$ when $m_n > 1$ for all $n$ sufficiently large, where in the last column, "$\wedge$" denotes logical AND, and "$p \Rrightarrow q$" means that $p$ implies $q$.

Assumption 5. *The decreasing rate limit*

$$\alpha_* = -\lim_{m\to\infty} \frac{\ln(\sigma_m)}{\ln(m)}, \tag{35}$$

*exists.*

In this assumption, $\alpha_*$ is the only constant such that the rate (as $m \to \infty$) at which $\sigma_m$ decreases is strictly faster than $m^{-\alpha_*+\epsilon}$ and strictly slower than $m^{-\alpha_*-\epsilon}$ for all $\epsilon > 0$; i.e.,

$$\sigma_m = o(m^{-\alpha_*+\epsilon}) \text{ and } \sigma_m = \omega(m^{-\alpha_*-\epsilon}) \text{ as } m \to \infty \text{ for any } \epsilon > 0, \tag{36}$$

which we denote by $\sigma_m \simeq m^{-\alpha_*}$ as $m \to \infty$. By (12) with $q = 2$, we see that

$$\alpha_* \geq 1 \quad \text{for RQMC when } h \in \text{BVHK} \tag{37}$$

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

$$\alpha_* > \frac{1}{2}. \tag{38}$$

The value of $\alpha_*$ 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.

Note that while the limit in $m$ in Assumption 5 in (35) may not exist for certain RQMC methods, it may instead hold for particular subsequences $m_*$ of $m$. This is often assumed in the QMC and RQMC literature when investigating the methods' efficiency through numerical experiments. Using the particular structure of so-called $(0, k_*, s)$-nets in base $b_0 \geq 2$ for example, such limits are established [28, Theorem 1] for nested scrambling and sufficiently smooth $h$ with $m_* = b_0^{k_*}$ as $k_* \to \infty$, which we will further analyze in Appendix B.5, comparing the convergence rates to some of our corollaries from Sections 3 and 4. Specifically, in the particular case of nested scrambling applied to a smooth function with Lipschitz continuous mixed partial of order $s$, [28, Theorem 1] obtains an exact convergence rate for the variance, which Appendix B.5 uses to get $\alpha_* = 3/2$ along the subsequence $m_* = b_0^{k_*}$ as $k_* \to \infty$. Even if we may need to consider the limit in (35) along such a subsequence $m_*$, we will consider it in $m$ to simplify the notation.

(Appendix D.1 presents results from numerical experiments to estimate $\alpha_*$ for different RQMC methods and various integrands corresponding to Assumptions 3.A ($h \in$ BVHK), 3.B ($h \in \mathcal{L}^\infty$), and 3.C ($h \in \mathcal{L}^{2+b}$). In particular, Table 3 shows that for all randomization methods considered, (37) seems to hold for the integrands satisfying Assumption 3.A, and the other integrands appear to be consistent with (38).)

For $m_n = n^c$, we get from (36) that (as $c$ is bounded)

$$\sigma_{m_n} = \omega(n^{-c\alpha_* - \epsilon}), \quad \text{as } n \to \infty, \text{ for any } \epsilon > 0. \tag{39}$$

Therefore, using $r_n = n^{1-c}$ leads to

$$\text{RMSE}\left[\widehat{\mu}^{\text{RQ}}_{m_n,r_n}\right] = \frac{\sigma_{m_n}}{\sqrt{r_n}} = \omega\left(n^{-v(\alpha_*,c)-\epsilon}\right), \quad \text{as } n \to \infty, \text{ for any } \epsilon > 0, \tag{40}$$

where $v(\alpha_*, c) \equiv c\left[\alpha_* - \frac{1}{2}\right] + \frac{1}{2}$. Similarly, (36) also yields $\text{RMSE}\left[\widehat{\mu}^{\text{RQ}}_{m_n,r_n}\right] = o\left(n^{-v(\alpha_*,c)+\epsilon}\right)$ as $n \to \infty$ for all $\epsilon > 0$. As $\epsilon$ can be arbitrarily small, $\text{RMSE}\left[\widehat{\mu}^{\text{RQ}}_{m_n,r_n}\right]$ then decreases at about rate (ignoring the leading coefficient and lower-order terms) $n^{-v(\alpha_*,c)}$ as $n \to \infty$. Our assumption (38) ensures $v(\alpha_*, c) > 1/2$ for every $c \in (0, 1)$. Consequently, for *any* choice of $c$, the RQMC estimator's RMSE shrinks faster than the MC estimator's RMSE, which from (1) decreases at rate $n^{-v_{\text{MC}}}$ as $n \to \infty$, where

$$v_{\text{MC}} \equiv \frac{1}{2}. \tag{41}$$

Moreover, for any fixed $\alpha_*$ satisfying (38), $v(\alpha_*, c)$ 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 (20) or AVCI (31).

As will be shown below in Sections 5.1–5.6, the corollaries guaranteeing CLT (20) or AVCI (31) in Sections 3 and 4 will typically lead to imposing restrictions on $c$ of the form

$$c < c_k(\alpha_*) \tag{42}$$

for some $0 < c_k(\alpha_*) \le 1$ depending on the particular Corollary $k$ considered. (The only exceptions to constraints as in (42) are Corollaries 5 and 9, which essentially have $c = 0$ because they assume that $n^c = m_n = m_0$ is fixed in (27); see Section 5.7 for more details.) We will see that except for $k = 1$ and 6, each upper bound $c_k(\alpha_*)$ in (42) is strictly decreasing in $\alpha_*$. Thus, as $\alpha_*$ gets larger (i.e., better RQMC convergence rate for a single randomization), the choices for $c$ ensuring CLT (20) or AVCI (31) 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 $n$. In this sense, when RQMC does better on a single randomization, our sufficient conditions handicap the method by choosing a budget allocation to employ more randomizations $r_n$ (i.e., smaller $c$) to secure a Gaussian limit. Appendix D.2 presents numerical results studying this, showing that for a fixed budget $n$ and varying $c$, CI coverage degrades when $c$ is too large.

The "largest" possible $c$ satisfying (42) 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 (40) then becomes $-[(c_k(\alpha_*) - \delta)(\alpha_* - 1/2) + 1/2] - \epsilon = -v_k(\alpha_*) - t_1(\epsilon, \delta)$, where

$$v_k(\alpha_*) \equiv c_k(\alpha_*)\left(\alpha_* - \frac{1}{2}\right) + \frac{1}{2} \tag{43}$$

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 (38). We then get by (40) that

$$\text{RMSE}\left[\widehat{\mu}^{\text{RQ}}_{m_n,r_n}\right] = \omega\left(n^{-v_k(\alpha_*)-t_1(\epsilon,\delta)}\right) \quad \text{as } n \to \infty, \text{ for any } \epsilon > 0 \text{ and any } \delta \in (0, \delta_0(\epsilon)). \tag{44}$$

By (36), we also have that $\sigma_m = o(m^{-\alpha_*+\epsilon'})$ as $m \to \infty$ for all arbitrarily small $\epsilon' > 0$, yielding $\sigma_{m_n} = o(n^{-c\alpha_*+\epsilon})$ as $n \to \infty$ for all $\epsilon > 0$. This then leads to $\text{RMSE}\left[\widehat{\mu}^{\text{RQ}}_{m_n,r_n}\right] = \sigma_{m_n}/\sqrt{r_n} =$

$o(n^{-[c(\alpha_*-1/2)+1/2]+\epsilon})$ as $n \longrightarrow \infty$, and using $c = c_k(\alpha_*) - \delta$ results in

$$\text{RMSE}\left[\widehat{\mu}^{\text{RQ}}_{m_n,r_n}\right] = o\left(n^{-v_k(\alpha_*)+t_2(\epsilon,\delta)}\right) \qquad \text{as } n \longrightarrow \infty,\ \text{for any } \epsilon > 0 \text{ and any } \delta \in (0, \delta_0(\epsilon)), \tag{45}$$

where $t_2(\epsilon, \delta) \equiv \epsilon + \delta[\alpha_* - 1/2] > 0$ under our assumption (38). Therefore, because $t_1(\epsilon, \delta) > 0$ and $t_2(\epsilon, \delta) > 0$ in (44) and (45) can be made arbitrarily small by taking $\epsilon \longrightarrow 0$ (which also ensures $\delta \longrightarrow 0$), the optimal rate at which RMSE $\left[\widehat{\mu}^{\text{RQ}}_{m_n,r_n}\right]$ decreases under (42) is about

$$\text{RMSE}\left[\widehat{\mu}^{\text{RQ}}_{m_n,r_n}\right] \simeq n^{-v_k(\alpha_*)} \qquad \text{as } n \longrightarrow \infty, \tag{46}$$

for $v_k(\alpha_*)$ from (43), where "$\simeq$" is as defined after (36).

We see by (43) that $v_k(\alpha_*) > 1/2$ under our assumption that $\alpha_* > 1/2$ in (38) because $c_k(\alpha_*) > 0$ always holds in (42). Hence, the optimal rate at which the RQMC estimator's RMSE shrinks is faster than MC's rate exponent $v_{\text{MC}} = 1/2$. For any Corollaries $k$ and $k'$, (43) also implies that for a fixed $\alpha_*$ under our assumption (38),

$$v_k(\alpha_*) > v_{k'}(\alpha_*) \quad \text{if and only if} \quad c_k(\alpha_*) > c_{k'}(\alpha_*), \tag{47}$$

so expanding the range of valid values for $c$ in (42) corresponds to better optimal approximate RMSE convergence rate by (46). If the constraint (42) has $c_k(\alpha_*) = 1$, which is the largest possible upper bound, then (43) shows that $v_k(\alpha_*) = \alpha_*$; thus in this case, even with the number of randomizations slowly growing large (i.e., $c = 1 - \delta$ for infinitesimally small $\delta > 0$, which we denote as $c \cong 1$), RMSE $\left[\widehat{\mu}^{\text{RQ}}_{m_n,r_n}\right]$ 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 (42) and $v_k(\alpha_*)$ in (43) for Corollaries $k = 1$, 2, 3, 4, 6, 7, and 8. Appendices B and C compare the resulting values analytically and graphically, respectively.

### 5.1 CLT and AVCI Conditions Under Corollaries 1 and 6

For $m_n \longrightarrow \infty$ as $n \longrightarrow \infty$, Corollary 1 ensures CLT (20) under condition (22) for some $d > 0$ (used for the order-$(2 + d)$ absolute central moment of $A_m$); Corollary 6 secures AVCI (31) using condition (22) with $d = 2$. As (22) 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 (42) and $v_1(\alpha_*)$ and $v_6(\alpha_*)$ in (43) as

$$c_1(\alpha_*) = c_6(\alpha_*) \equiv 1 \quad \text{and} \quad v_1(\alpha_*) \equiv c_1(\alpha_*)\left(\alpha_* - \frac{1}{2}\right) + \frac{1}{2} = \alpha_* \equiv v_6(\alpha_*). \tag{48}$$

Thus, as noted in the discussion after (47), because $v_1(\alpha_*) = v_6(\alpha_*) = \alpha_*$, the RMSE of $\widehat{\mu}^{\text{RQ}}_{m_n,r_n}$ as $n \longrightarrow \infty$ with the number $r_n$ of randomizations slowly growing large (i.e., $c < 1$ with $c \cong 1$) decreases at about the same rate $n^{-\alpha_*}$ as for an RQMC estimator with a *fixed* number $r_0 \geq 1$ of randomizations. This is perhaps the best that one can hope for to ensure a CLT or AVCI with RQMC. Moreover, as previously noted in Section 3.1, if we instead consider the more general allocation $(m_n, r_n)$ of Assumption 1.A, Corollaries 1 and 6 allow the number $r_n$ of randomizations to grow more slowly than $n^c$ for any $c > 0$, e.g., $r_n = \lfloor \ln n \rfloor$, permitting RQMC to even further exploit QMC's fast convergence.

## 5.2 Corollary 2: CLT Conditions When $h \in$ BVHK

We now derive a constraint on $c$ as in (42) to ensure CLT (20) through Corollary 2, which requires $h \in$ BVHK (Assumption 3.A) and that $(m_n, r_n) = (n^c, n^{1-c})$ satisfies (24) for some (small) $\lambda \in (0, 1)$. Then the left side of (24) becomes $n^{(1-c)(1-\lambda)} \left( \frac{n^c \sigma_{m_n}}{(c \ln n)^s} \right)^2$. Thus, squaring (39) implies that (24) holds when $(1-c)(1-\lambda) + 2c - 2c\alpha_* > 0$ for some $\lambda \in (0, 1)$. This is equivalent to

$$c < \frac{1-\lambda}{2\alpha_* - 1 - \lambda} \equiv c_2'(\alpha_*, \lambda). \tag{49}$$

Because Corollary 2 assumes that $h \in$ BVHK so $\alpha_* \geq 1$ by (37), we get $c_2'(\alpha_*, \lambda) = 1$ when $\alpha_* = 1$. Now consider any fixed $\alpha_* > 1$. As $c_2'(\alpha_*, \lambda)$ is strictly decreasing in $\lambda \in (0, 1)$, choosing $\lambda$ smaller leads to a looser constraint (i.e., more possible choices for $c$ satisfying (49)), and Corollary 2 allows taking $\lambda \in (0, 1)$ in (24) to be arbitrarily small. But as $c_2'(\alpha_*, \lambda)$ is continuous in $\lambda \in (0, 1)$ and because (49) has a strict inequality, we can replace (49) with the constraint

$$c < \frac{1}{2\alpha_* - 1} \equiv c_2(\alpha_*), \qquad \text{which satisfies} \qquad 0 < c_2(\alpha_*) \leq 1 \tag{50}$$

since $\alpha_* \geq 1$. (To see why (49) can be replaced by the constraint in (50), note that if $c < c_2(\alpha_*)$, then (49) 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 (50). Also, (37) ensures that $\lambda_0 \leq 1$.) If $\alpha_* = 1$, then $c_2(\alpha_*) = 1$, making (50) the weakest possible constraint on $c$. 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

$$v_2(\alpha_*) \equiv c_2(\alpha_*) \left( \alpha_* - \frac{1}{2} \right) + \frac{1}{2} = 1 \tag{51}$$

by (43), (50) and (37), so $v_2(\alpha_*) > v_{\mathrm{MC}} = 1/2$.

## 5.3 Corollary 3: CLT Conditions When $h \in \mathcal{L}^\infty$

Note that (39) implies that $(m_n, r_n) = (n^c, n^{1-c})$ satisfies condition (25) of Corollary 3 if $r_n \sigma^2_{m_n} = \omega\left(n^{1-c} n^{-2c\alpha_* - 2\epsilon}\right) \to \infty$ as $n \to \infty$ for all sufficiently small $\epsilon > 0$, which is true when $1 - c - 2c\alpha_* > 0$, or equivalently,

$$c < \frac{1}{2\alpha_* + 1} \equiv c_3(\alpha_*), \qquad \text{which satisfies} \qquad 0 < c_3(\alpha_*) < \frac{1}{2} \tag{52}$$

under the assumption that $\alpha_* > 1/2$ in (38). If $\alpha_* \geq 1$, as (37) ensures when $h \in$ BVHK, which is not required by Corollary 3, then $0 < c_3(\alpha_*) \leq 1/3$.

By (43) and (52), 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

$$v_3(\alpha_*) \equiv c_3(\alpha_*) \left( \alpha_* - \frac{1}{2} \right) + \frac{1}{2} = \frac{2\alpha_*}{2\alpha_* + 1}, \qquad \text{which satisfies} \qquad \frac{1}{2} < v_3(\alpha_*) < 1 \tag{53}$$

when $\alpha_* > 1/2$. (If $\alpha_* \geq 1$, as in (37), then $2/3 \leq v_3(\alpha_*) < 1$.) Hence, under Corollary 3, the RQMC estimator's RMSE shrinks faster than the MC estimator's RMSE by (41).

## 5.4 Corollary 4: CLT Conditions When $h \in \mathcal{L}^{2+b}$ for Some $b > 0$

We now derive a constraint on $c$ as in (42) to ensure CLT (20) through Corollary 4, which assumes that $h \in \mathcal{L}^{2+b}$ (Assumption 3.C) and (26) both hold for some $b > 0$. In (26), 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 (39). Thus, $(m_n, r_n) = (n^c, n^{1-c})$ satisfies (26) if $r_n^{b/(2+b)}\sigma^2_{m_n} = \omega\left(n^{(1-c)\frac{b}{2+b}-2c\alpha_*-2\epsilon}\right) \to \infty$ as $n \to \infty$, which holds when $(1-c)\frac{b}{2+b} - 2c\alpha_* > 0$, or equivalently,

$$c < \frac{1}{2\alpha_*(1+\frac{2}{b})+1} \equiv c_4(\alpha_*, b), \qquad \text{which satisfies} \qquad 0 < c_4(\alpha_*, b) < \frac{1}{2} \tag{54}$$

for each $b > 0$ when $\alpha_* > 1/2$, as assumed in (38). If $\alpha_* \geq 1$, as (37) ensures when $h \in$ BVHK (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 (54).

By (43), under Corollary 4, the optimal rate at which RMSE$[\widehat{\mu}^{\mathrm{RQ}}_{m_n,r_n}]$ decreases (as $n \to \infty$) is about $n^{-v_4(\alpha_*,b)}$, where for any $b > 0$,

$$v_4(\alpha_*, b) \equiv c_4(\alpha_*, b)\left(\alpha_* - \frac{1}{2}\right) + \frac{1}{2} = \frac{2\alpha_*(1+\frac{1}{b})}{2\alpha_*(1+\frac{2}{b})+1}, \qquad \text{which satisfies} \qquad \frac{1}{2} < v_4(\alpha_*, b) < 1 \tag{55}$$

when $\alpha_* > 1/2$. Hence, we get $v_4(\alpha_*, b) > v_{\mathrm{MC}} = 1/2$ by (41) 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 RMSE$[\widehat{\mu}^{\mathrm{RQ}}_{m_n,r_n}]$ under Corollary 4 improves as $h(U)$ has more finite absolute moments.

It is interesting to note that $\lim_{b\to\infty} c_4(\alpha_*, b) = c_3(\alpha_*)$, the latter from (52) 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 $b \to \infty$.

### 5.5 Corollary 7: AVCI Conditions When $h \in$ BVHK

For $I^{\mathrm{RQ}}_{m_n,r_n,\gamma}$ in (29) to be AVCI as in (31), we assumed, in addition to the CLT in (20), that (30) holds. Corollary 7 ensures (30) is satisfied when $h \in$ BVHK (Assumption 3.A) and condition (33) holds, which is the same as the square of CLT condition (24) for $\lambda = 1/2$. In the setting of Assumption 1.B, (24) holds for $\lambda = 1/2$ if $c$ satisfies (49) with $\lambda = 1/2$, so AVCI condition (33) is true when

$$c < \frac{1}{4\alpha_* - 3} \equiv c_7(\alpha_*), \qquad \text{which satisfies} \qquad 0 < c_7(\alpha_*) \leq 1 \tag{56}$$

because $\alpha_* \geq 1$ by (37). Note that $c_7(\alpha_*) = 1$ when $\alpha_* = 1$, and $c_7(\alpha_*) < 1$ for $\alpha_* > 1$.

By (43), 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

$$v_7(\alpha_*) \equiv c_7(\alpha_*)\left(\alpha_* - \frac{1}{2}\right) + \frac{1}{2} = \frac{3\alpha_* - 2}{4\alpha_* - 3}, \qquad \text{which satisfies} \qquad \frac{3}{4} < v_7(\alpha_*) \leq 1 \tag{57}$$

since $\alpha_* \geq 1$ by (37). Note that (57) implies that $v_7(\alpha_*) > v_{\mathrm{MC}} = 1/2$. Also, $v_7(\alpha_*) = 1$ when $\alpha_* = 1$, and $v_7(\alpha_*) < 1$ for $\alpha_* > 1$. Thus, while (57) implies that RMSE$[\widehat{\mu}^{\mathrm{RQ}}_{m_n,r_n}]$ can at best decrease at about rate $n^{-1}$ as $n \to \infty$ under Corollary 7, we can do better through Corollary 6 when its conditions hold, as discussed in Section 5.1. In the latter case, (48) shows that RMSE$[\widehat{\mu}^{\mathrm{RQ}}_{m_n,r_n}]$ shrinks at about rate $n^{-\alpha_*}$, which is better than Corollary 7 for $\alpha_* > 1$.

### 5.6 Corollary 8: AVCI Conditions When $h \in \mathcal{L}^4$

By Corollary 8, which does not require $h \in$ BVHK, the combination of $h \in \mathcal{L}^4$ (Assumption 3.C with $b = 2$) and condition (34) implies (30), then ensuring AVCI (31). Under Assumption 1.B, we have in (34) that $\sigma^4_{m_n} = \omega(n^{-4c\alpha_*-4\epsilon})$ as $n \to \infty$ for any $\epsilon > 0$ by (39) and $r_n = n^{1-c}$. Thus, (34) 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,

$$c < \frac{1}{4\alpha_* + 1} \equiv c_8(\alpha_*), \qquad \text{which satisfies} \qquad 0 < c_8(\alpha_*) < \frac{1}{3} \tag{58}$$

| Cor. $k$ | Ensures | Key Assumption | $c$ upper bd $c_k(\alpha_*)$ | | RMSE rate exp $v_k(\alpha_*)$ | |
|---|---|---|---|---|---|---|
| 1 | **CLT** | Eq. (22) for some $d > 0$ | $1$ | $\geq$ | $\alpha_*$ | |
| 2 | **CLT** | 3.A ($h \in$ BVHK) | $\frac{1}{2\alpha_*-1}$ | $>$ | $1$ | $>$ |
| 3 | **CLT** | 3.B ($h \in \mathcal{L}^\infty$) | $\frac{1}{2\alpha_*+1}$ | $>$ | $\frac{2\alpha_*}{2\alpha_*+1}$ | $>$ |
| 4 | **CLT** | 3.C ($h \in \mathcal{L}^{2+b}$ for some $b > 0$) | $\frac{1}{2\alpha_*(1+\frac{2}{b})+1}$ | $>$ | $\frac{2\alpha_*(1+\frac{1}{b})}{2\alpha_*(1+\frac{2}{b})+1}$ | $>$ |
| 5 | **CLT** | 3.D ($h \in \mathcal{L}^2$) | $0$ | | $\frac{1}{2}$ | |
| 6 | **AVCI** | Eq. (22) for $d = 2$ | $1$ | $\geq$ | $\alpha_*$ | |
| 7 | **AVCI** | 3.A ($h \in$ BVHK) | $\frac{1}{4\alpha_*-3}$ | $>$ | $\frac{3\alpha_*-2}{4\alpha_*-3}$ | $>$ |
| 8 | **AVCI** | 3.C ($h \in \mathcal{L}^4$) | $\frac{1}{4\alpha_*+1}$ | $>$ | $\frac{3\alpha_*}{4\alpha_*+1}$ | $>$ |
| 9 | **AVCI** | 3.D ($h \in \mathcal{L}^2$) | $0$ | | $\frac{1}{2}$ | |

Table 2. Summary of Corollaries 1–9 under Assumption 1.B (i.e., $(m_n, r_n) = (n^c, n^{1-c})$), where the inequalities in the last two columns compare entries in successive rows within the same column.

when $\alpha_* > 1/2$, as assumed in (38). If $\alpha_* \geq 1$, then $0 < c_8(\alpha_*) \leq 1/5$.

By (43), 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

$$v_8(\alpha_*) \equiv c_8(\alpha_*)\left(\alpha_* - \frac{1}{2}\right) + \frac{1}{2} = \frac{3\alpha_*}{4\alpha_* + 1}, \quad \text{which satisfies} \quad \frac{1}{2} < v_8(\alpha_*) < \frac{3}{4} \tag{59}$$

when $\alpha_* > 1/2$. Thus, we have $v_8(\alpha_*) > v_{\mathrm{MC}} = 1/2$ by (41).

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

As noted before after (42), Corollaries 5 and 9 essentially have $c = 0$ as they assume that $n^c = m_n = m_0$ is fixed in (27). Thus, under Corollaries 5 and 9, which require Assumption 3.D ($h \in \mathcal{L}^2$), 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 (43) and (46) by setting $c_k(\alpha_*) = 0$. Hence, the RMSE for this case of RQMC shrinks at the same rate as for MC by (41), but RQMC, which can be viewed as a variance-reduction technique (e.g., see [19, 21]), typically has a smaller leading coefficient for its rate.

To facilitate comparisons with the other corollaries, we define $c_5(\alpha_*) = c_9(\alpha_*) = 0$, and we take $c = 0$ for $k \in \{5, 9\}$. Also, we let $v_5(\alpha_*) = v_9(\alpha_*) = 1/2$, consistent with (43) and (46). By (54), we see that $\lim_{b\to 0} c_4(\alpha_*, b) = 0 = c_5(\alpha_*)$, so the allocation condition of Corollary 4 (which assumes $h \in \mathcal{L}^{2+b}$) reduces to that of Corollary 5 (which assumes $h \in \mathcal{L}^2$) as $b \to 0$. Similarly, we have that $\lim_{b\to 0} v_4(\alpha_*, b) = 1/2 = v_5(\alpha_*)$. Thus, the tradeoffs of the conditions of Corollaries 4 and 5 disappear as $b \to 0$.

### 5.8 Summary of Results Under Assumption 1.B

Appendix B (resp., C) analytical (resp., graphical) comparisons of some of the $c_k(\alpha_*)$ and $v_k(\alpha_*)$ values for the various Corollaries $k$.

Table 2 summarizes the most important assumptions and conclusions of Corollaries 1–9 under Assumption 1.B, which enables sharper results compared to those in Table 1 under Assumption 1.A. The table shows that strengthening the assumptions ensures that RQMC's RMSE shrinks at a faster rate.

## 6 CONCLUSIONS

RQMC methods provide powerful simulation tools accelerating the convergence rate with respect to MC. A standard way to estimate the RQMC error in practice exploits an assumed CLT over $r \geq 2$ i.i.d. randomizations, but typically restricting 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. A common rule of thumb suggests a CLT roughly holds for fixed $r \approx 30$; however, this heuristic has lacked rigorous theoretical support for most RQMC methods. The only existing CLT results that the authors are aware of as $m$ increases and fixed $r < \infty$ [2, 13, 28] are for scrambled digital nets with nested uniform scrambling, but this RQMC technique is computationally expensive, which has limited its adoption in the past, although this may be less of an issue with today's more powerful computers. In contrast, [23] proves that increasing $m$ for any fixed number of randomizations can lead to a non-normal limiting distribution. To our knowledge, no theoretical result covering a broad class of RQMC methods has ever been published in the literature guaranteeing a CLT for $m \to \infty$. Our paper provides 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 (at least along a specified subsequence). 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.

Many of our corollaries in Sections 3.1 and 4.1 specify that an allocation $(m, n) = (m_n, r_n)$ has that $m_n$ does not grow too quickly relative to $r_n$ as the computing budget $n \to \infty$. For these results, if the variance of an RQMC estimator from a single randomization decreases too quickly, we may not be able to ensure that an RQMC estimator from $r_n$ randomizations obeys a CLT with a Gaussian limit nor yields an AVCI; see Appendix D.2 for numerical results studying this. But an exception to this is Corollary 1 (resp., 6), which secures a CLT (resp., AVCI) for $r_n \to \infty$ at *any* arbitrarily slow rate as $n \to \infty$ when condition (22) holds for some $d > 0$ (resp., $d = 2$). We are currently investigating alternatives to condition (22) 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 $\alpha_*$ in (35), 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 $\alpha_*$. Given that we provide *sufficient* conditions for a CLT or AVCI, 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 $\mu$ based on a CLT, we are additionally investigating instead employing resampling methods, such as the bootstrap $t$ [37, Chapter 17 end notes].

## ACKNOWLEDGMENT

This material is based upon work supported in part by the National Science Foundation under Grant No. CMMI-1537322. Additional support came from an NJIT seed grant. The authors would like to thank Art Owen for pointing out references, used in the proof of Proposition 2.1, showing that if the integrand $h$ has bounded Hardy-Krause variation, then $h$ is bounded.

## A PROOFS

Proof of Proposition 2.1. As shown in [33], functions of bounded Hardy-Krause variation are Riemann integrable, and therefore bounded, which [32] also uses. Thus, Assumption 3.A (integrand $h \in$ BVHK) is stronger than Assumption 3.B ($h \in \mathcal{L}^\infty$). To show that the relation is strict, for any dimension $s \geq 2$, consider $h$ to be an indicator function (so $\mu$ is a probability) with discontinuities not lining up with the axes: it is bounded but has $V_{\mathrm{HK}}(h) = \infty$. An explicit example [33] is $h(u) = \mathcal{I}(\sum_{i=1}^{s} u_i \leq 1)$, 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 ($h \in \mathcal{L}^{2+b}$ for some $b > 0$). For $h \in \mathcal{L}^\infty$, there exists $t_0 < \infty$ such that $|h(u) - \mu| \leq 2t_0$ for almost every $u \in [0,1]^s$. It follows that $\mathbb{E}[|h(U) - \mu|^q] \leq (2t_0)^q$ for all $q > 0$, so $h \in \mathcal{L}^q$ for all $q > 0$. To show the converse is not true, consider $h(u) = \Phi^{\leftarrow}(u)$ for $s = 1$ (recall $\Phi$ is the $\mathcal{N}(0,1)$ CDF), where $F^{\leftarrow}(u) = \inf\{x : F(x) \geq u\}$ is the (generalized) inverse of a CDF $F$, so $h \notin \mathcal{L}^\infty$ but $h \in \mathcal{L}^q$ for all $q \geq 1$.

Assumption 3.C implies Assumption 3.D ($h \in \mathcal{L}^2$) by Lyapounov's inequality [3, pp. 81 and 277]. To show that the relation is strict, we modify an example of [9, p. 366]. For $s = 1$, let $h(u) = F^{\leftarrow}(u)$, where $F$ is the CDF of the density function $f(x) = k_0 \mathcal{I}(x \geq e)/[x^3(\ln x)^2]$ with $k_0 = (\int_e^\infty \mathrm{d}x/[x^3(\ln x)^2])^{-1} \doteq 26.64$. 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 $\mathbb{E}[h(U)^{2+b}] = \int_e^\infty \frac{k_0 x^{2+b}}{x^3(\ln x)^2}\,\mathrm{d}x = k_0 \int_e^\infty \frac{1}{x^{1-b}(\ln x)^2}\,\mathrm{d}x \geq (k_0/k_b)\int_e^\infty \frac{1}{x^{1-b/2}}\,\mathrm{d}x = \infty$ for $k_b = [4/(be)]^2$ because $(\ln x)^2 \leq k_b x^{b/2}$ for all $x \geq e$. ■

Proof of Lemma 2.3. As seen by (10), $X_{n,1}$ averages $m_n$ *dependent* terms, which we will handle through Minkowski's inequality [3, eq. (5.40)]:

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

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

Proof of Theorem 3.2. Fix any $t > 0$ for $\tau_n^2(t)$ in (18), and we will bound $\tau_n^2(t)$ as in the proof of the Lyapounov CLT in [3, Theorem 27.3]. Specifically, the condition $|y| > ts_n$ in (18) implies that $|y|^d/(t^d s_n^d) > 1$, so

$$\tau_n^2(t) \leq \frac{1}{t^d s_n^d}\int_{|y|>ts_n} y^2|y|^d\,\mathrm{d}G_n(y) \leq \frac{1}{t^d s_n^d}\mathbb{E}\left[|Y_{n,1}|^{2+d}\right] = \frac{1}{t^d r_n^{d/2}\sigma_{m_n}^d}\mathbb{E}\left[|X_{n,1} - \mu|^{2+d}\right] \tag{60}$$

by (17), with $\mathbb{E}\left[|X_{n,1} - \mu|^{2+d}\right] < \infty$ by assumption, which ensures $\sigma_{m_n}^2 < \infty$. Thus, as in (19), dividing throughout eq. (60) by $\sigma_{m_n}^2$, which is strictly positive for all $n$ large enough (Assumption 4), shows that (21) guarantees (19) because $t > 0$ is fixed and arbitrary, so CLT (20) holds by Theorem 3.1. ■

Proof of Corollary 1. The moments of $X_{n,1}$ in (21) are the same as the moments of $A_{m_n}$ in (22) with $m = m_n$. As a consequence, because $m_n \to \infty$ as $n \to \infty$, (22) yields

$$\frac{\mathbb{E}\left[\left|X_{n,1} - \mu\right|^{2+d}\right]}{\sigma_{m_n}^{2+d}} \leq k_1, \qquad \text{for all } n \text{ sufficiently large.} \tag{61}$$

For all $n$ satisfying eq. (61) and Assumption 4, the left side of (21) is bounded above by $k_1/r_n^{d/2}$, which vanishes as $n \to \infty$ because $d > 0$ and $r_n \to \infty$ as $n \to \infty$ under Assumption 1.A. Thus, (21) holds, so CLT (20) follows from Theorem 3.2. Under (23), we have $\mathbb{E}\left[|X_{n,1} - \mu|^{2+d}\right] \leq k_2^{2+d}\sigma_{m_n}^{2+d}$ for all $n$ sufficiently large because $m_n \to \infty$, securing (22) with $k_1 = k_2^{2+d}$. ■

Proof of Corollary 2. We will verify the conditions of Theorem 3.2 for $d = 2(1-\lambda)/\lambda > 0$. 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 $h \in$ BVHK, Lemma 2.2 guarantees that $\mathbb{E}[|X_{n,1} - \mu|^{2+d}] < \infty$ by (12) with $q = 2 + d$, so $0 < \sigma_{m_n}^2 < \infty$. Also, applying (12) of Lemma 2.2 ensures that the left side of (21) satisfies

$$\frac{\mathbb{E}\left[\left|X_{n,1} - \mu\right|^{2+d}\right]}{r_n^{d/2}\sigma_{m_n}^{2+d}} \leq \frac{1}{r_n^{d/2}}\left(\frac{w_0' V_{\mathrm{HK}}(h)(\ln m_n)^s}{m_n \sigma_{m_n}}\right)^{2+d}. \tag{62}$$

Raising the right side of eq. (62) to the $2/(2+d)$ power shows that it vanishes as $n \to \infty$ if and only if

$$r_n^{d/(2+d)}\left(\frac{m_n \sigma_{m_n}}{(\ln m_n)^s}\right)^2 \to \infty, \qquad \text{as } n \to \infty. \tag{63}$$

But $d/(2+d) = 1 - \lambda$, so (24) and (63) imply (21), and (19) and (20) follow from Theorem 3.2. ■

Proof of Corollary 3. As $h \in \mathcal{L}^\infty$ (Assumption 3.B), there exists a constant $t_0 \in (0, \infty)$ such that $\mathbb{P}(|Y_{n,j}| \leq 2t_0) = 1$ for all $n$ and $j$. Hence, (18) 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 (25). Thus, the numerator of (19) is zero for all $n > n_0$, so (20) follows from Theorem 3.1. ■

Proof of Corollary 4. Because $b > 0$, raising (26), which has nonnegative left side, to the $(2+b)/2$ power is equivalent to

$$r_n^{b/2}\sigma_{m_n}^{2+b} \to \infty, \qquad \text{as } n \to \infty. \tag{64}$$

By (11), Lemma 2.3 with $q = 2 + b$ then implies that the left side of (21) satisfies

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

by eq. (64), so (21) holds. Hence, (19) and (20) follow from Theorem 3.2. ■

Proof of Corollary 5. For $n \geq m_0$, we have by (27) 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 $\mu$ and variance $\sigma_{m_0}^2 = \mathrm{Var}[X_1] = \mathrm{Var}[X_{n,1}]$, which Lemma 2.3 with $q = 2$ ensures is finite by Assumption 3.D. For $G$ as the distribution of $X_1 - \mu$, (19) 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 3.1 yields (20). ■

Proof of Proposition 3.3. We first compare the restrictions on $(m_n, r_n)$ in Corollaries 4 and 5. If (27) holds, then the left side of (26) 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 (27) implies (26).

We next analyze the requirements on $(m_n, r_n)$ in Corollaries 3 and 4. Expressing the left side of (25) as $r_n\sigma_{m_n}^2 = r_n^{2/(2+b)} r_n^{b/(2+b)}\sigma_{m_n}^2$ shows that (26) implies (25) since $r_n^{2/(2+b)} \to \infty$ as $n \to \infty$ by Assumption 1.A.

We finally compare condition (24) of Corollary 2 with condition (25) of Corollary 3. Writing the left side of (24) 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 (25) ensures that (24) is satisfied when (28) holds. To verify (28) 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 (28) 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 $\infty$ 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.) ■

Proof of Theorem 4.1. Note that (30) implies that $\sigma_{m_n}/\widehat{\sigma}_{m_n,r_n} \Rightarrow 1$ as $n \to \infty$ by the continuous-mapping theorem (e.g., [3, Theorem 25.7]). Thus, (20) ensures that

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

by Slutsky's theorem [38, Theorem 1.5.4]. Hence,

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

by the portmanteau theorem [3, Theorem 25.8], establishing (31). ■

Proof of Theorem 4.2. As (32) is equivalent to (21) for $d = 2$, Theorem 3.2 guarantees CLT (20) because we assumed that $\mathbb{E}\left[(X_{n,1} - \mu)^4\right] < \infty$. Thus, if (30) holds, then Theorem 4.1 will imply (31).

We will prove (30) 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 [3, eq. (21.13)] and because $\mathbb{E}[\widehat{\sigma}^2_{m_n,r_n}] = \sigma^2_{m_n}$ [38, p. 173], we get

$$\begin{aligned} p_n &\le \frac{1}{v^2}\mathbb{E}\left[\left(\frac{\widehat{\sigma}^2_{m_n,r_n}}{\sigma^2_{m_n}} - 1\right)^2\right] = \frac{1}{v^2\sigma^4_{m_n}}\mathbb{E}\left[\left(\widehat{\sigma}^2_{m_n,r_n} - \sigma^2_{m_n}\right)^2\right] = \frac{1}{v^2\sigma^4_{m_n}}\mathrm{Var}\left[\widehat{\sigma}^2_{m_n,r_n}\right] \\ &= \frac{1}{v^2 r_n \sigma^4_{m_n}}\left(\mathbb{E}\left[\left(X_{n,1} - \mu\right)^4\right] - \frac{r_n - 3}{r_n - 1}\sigma^4_{m_n}\right), \end{aligned} \tag{65}$$

where eq. (65) follows from [38, 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, eq. (65) vanishes as $n \to \infty$ by (32) since $v > 0$ is fixed, thus verifying (30). ■

Proof of Corollary 6. As shown in eq. (61), because $m_n \to \infty$, (22) with $d = 2$ implies that for all $n$ sufficiently large, the left side of (32) 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, (32) holds, so CLT (20), (30), and AVCI (31) follow from Theorem 4.2. Under (23), we have $\mathbb{E}\left[|X_{n,1} - \mu|^4\right] \le k_2^4\sigma^4_{m_n}$ for all $n$ sufficiently large because $m_n \to \infty$, securing (61) for $d = 2$ with $k_1 = k_2^4$. ■

Proof of Corollary 7. As the left side of (33) is nonnegative, taking the square-root of (33) shows that it is equivalent to (24) with $\lambda = 1/2$, so (33) guarantees CLT (20) by Corollary 2. We will next establish (30) and (31) by verifying the conditions of Theorem 4.2. 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 $h \in \mathrm{BVHK}$ from Assumption 3.A, Lemma 2.2 implies that $\mathbb{E}\left[\left(X_{n,1} - \mu\right)^4\right] < \infty$ by (12) with $q = 4$. Also, (12) ensures that the left side of (32) satisfies

$$\frac{\mathbb{E}\left[\left(X_{n,1} - \mu\right)^4\right]}{r_n\sigma^4_{m_n}} \le \frac{1}{r_n}\left(\frac{w_0' V_{\mathrm{HK}}(h)(\ln m_n)^s}{m_n\sigma_{m_n}}\right)^4. \tag{66}$$

Using (33) in eq. (66) verifies (32), so (30) and (31) follow from Theorem 4.2. ■

Proof of Corollary 8. As the left side of (34) is nonnegative, taking the square-root of (34) shows it is equivalent to (26) with $b = 2$, so (34) ensures CLT (20) by Corollary 4 because we further assumed that (11) holds for $b = 2$. We will next establish (30) and (31) by verifying the conditions of Theorem 4.2. By Lemma 2.3 with $q = 4$, we see that the numerator in (32) 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, (34) implies (32), so Theorem 4.2 yields (30) and (31). ■

Proof of Corollary 9. Corollary 5 ensures the CLT (20) is true. Also, as each $X_{n,j} = X_j$ in (14) does not depend on $n$, the triangular array reduces to a single i.i.d. sequence. Consequently, standard arguments (e.g., [38, Theorem 2.2.3A]) show that (30) holds, so Theorem 4.1 implies (31). ■

Proof of Proposition 4.3. If (27) 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 (27) implies (34). To compare (33) and (34), 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 (33) is weaker than (34). ■

## B 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 (42) for $c \in (0, 1)$ in Assumption 1.B and the optimal approximate rates $v_k(\alpha_*)$ in (43) from Sections 5.1–5.6. While many of the corollaries involve conditions that depend on the interaction of the integrand and the RQMC method through only $\sigma_m = \left(\mathrm{Var}[\frac{1}{m}\sum_{i=1}^m h(U'_{i,1})]\right)^{1/2}$, the cases $k = 1$ and 6 instead impose other requirements, (22) or (23), that involve higher-order absolute central moment or almost-sure behavior, which may be difficult to verify in practice. The different forms of the conditions for $k = 1$ and 6 complicate their comparisons with the other corollaries, so we mostly omit $k = 1$ and 6 from the following comparisons. 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 $k \notin \{1, 6\}$ when $\alpha_* \geq 1$. As shown before in Propositions 2.1 and 3.3 under Assumption 1.A, we will see trade-offs in the conditions that ensure CLT (20) 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 (42). A similar situation will also hold for guaranteeing AVCI (31), as we saw before in Proposition 4.3 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.

### B.1 $k = 2$ vs. $k = 3$ and $k = 4$

Comparing $c_2(\alpha_*)$ in (50) of Corollary 2 with $c_3(\alpha_*)$ from (52) for Corollary 3 shows that $c_2(\alpha_*) > c_3(\alpha_*)$; also see Proposition 3.3. This then implies that $v_2(\alpha_*) > v_3(\alpha_*)$ by (47). But recall that Corollary 2 required that $h \in$ BVHK (Assumption 3.A), which is stronger (Proposition 2.1) than restricting to $h \in \mathcal{L}^\infty$ (Assumption 3.B), as Corollary 3 imposed.

We now compare $c_2(\alpha_*)$ in (50) from Corollary 2 to $c_4(\alpha_*, b)$ from (54) of Corollary 4 when $h \in$ BVHK (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 (50) is (substantially) less restrictive on our choices for $c$ than (54) for each $b > 0$. This further implies that $v_2(\alpha_*) > v_4(\alpha_*, b)$ by (47).

### B.2 $k = 3$ vs. $k = 4$

Observe that $c_3(\alpha_*) > c_4(\alpha_*, b)$ for each $b > 0$ by (52) and (54), so the condition (54) for Corollary 4 restricts our choices for $c$ more than condition (52) from Corollary 3, but (52) was obtained under a stronger assumption ($h \in \mathcal{L}^\infty$, i.e., Assumption 3.B) than requiring that $h \in \mathcal{L}^{2+b}$ for some $b > 0$ (Assumption 3.C), used to get (54) for Corollary 4. Because $c_4(\alpha_*, b)$ is strictly increasing in $b$, the constraint on $c$ from condition (54) 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 (54). Similarly, by (47), the exponent $v_4(\alpha_*, b)$ from (55) 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 (53) under Corollary 3.

Section 5.4 previously noted that $\lim_{b\to\infty} c_4(\alpha_*, b) = c_3(\alpha_*)$ and $\lim_{b\to\infty} v_4(\alpha_*, b) = v_3(\alpha_*)$. Thus, the tradeoffs of the conditions of Corollaries 3 and 4 (i.e., $h \in \mathcal{L}^\infty$ vs. $h \in \mathcal{L}^{2+b}$) disappear as $b \to \infty$. While this did not necessarily have to happen (e.g., as in the proof of Proposition 2.1, consider $h(u) = \Phi^{\leftarrow}(u)$, which is unbounded but $h(U)$ has finite moments of all orders, where we recall that $\Phi^{\leftarrow}$ denotes the quantile function of $\Phi$), 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.

### B.3 $k = 7$ vs. $k = 2$ and $k = 8$

We first compare $c_7(\alpha_*)$ from (56) and $c_8(\alpha_*)$ in (58), each of which is an upper bound for $c$ to ensure AVCI (31). It is clear that $c_7(\alpha_*) > c_8(\alpha_*)$, so condition (56), obtained under $h \in$ BVHK (Assumption 3.A), is a strictly weaker restriction on the choice of $c$ than the condition (58), derived under $h \in \mathcal{L}^{2+b}$ (Assumption 3.C) but without requiring $h \in$ BVHK; also see Proposition 4.3. Thus, assuming the stricter condition $h \in$ BVHK (see Proposition 2.1) allows us to expand the values of $c$ that ensure AVCI. Moreover, we have that $v_7(\alpha_*) > v_8(\alpha_*)$ by (47), so when $h \in$ BVHK, 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 (56) is more restrictive than the upper bound $c_2(\alpha_*)$ from CLT condition (50), where both were obtained under the assumption that $h \in$ BVHK. 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 (20) and AVCI (31). But for $\alpha_* > 1$, ensuring AVCI (31) requires restricting $c$ more than what is needed to make sure CLT (20) holds.

### B.4 $k = 8$ vs. $k = 3$ and $k = 4$

Note that $c_8(\alpha_*) < c_3(\alpha_*)$ for all $\alpha_* > 0$ by (52) and (58), so AVCI condition (58) obtained from Corollary 8 restricts our choices for $c$ more than the CLT condition (52) under Corollary 3. Hence, ensuring AVCI (31) under Corollary 8 ($h \in \mathcal{L}^4$) requires further constraining $c$ compared to obtaining CLT (20) under Corollary 3 ($h \in \mathcal{L}^\infty$). Moreover, by (47), it follows that $v_8(\alpha_*) < v_3(\alpha_*)$.

By Corollary 4, CLT (20) also is secured when (54) and (11) hold for some $b > 0$, as discussed in Section 5.4. To compare the AVCI upper bound $c_8(\alpha_*)$ in (58) to CLT bound $c_4(\alpha_*)$ in (54), which has $c_4(\alpha_*, b) = 1/[1 + 2\alpha_*(1 + \frac{2}{b})]$, note that $c_4(\alpha_*, b) \le c_8(\alpha_*)$ if and only if $b \le 2$. Thus, if we select $c < c_4(\alpha_*, b)$ for some $b \le 2$ to ensure the CLT (20) under Corollary 4, then the same $c$ also yields AVCI (31).

### B.5 Comparisons with Loh's CLT

For a smooth function with Lipschitz continuous mixed partial of order $s$, Loh [28, Theorem 1] shows that for a $(0, k_*, s)$-net in base $b_0$, the variance $\sigma_m^2$ from a single randomization of $m = b_0^{k_*}$ points satisfies $\sigma_m^2 = \Theta(k_*^{s-1} b_0^{-3k_*}) = \Theta((\ln m/\ln b_0)^{s-1} m^{-3})$ as $k_* \to \infty$, so the standard deviation from a single randomization of $m = b_0^{k_*}$ points satisfies $\sigma_m = \Theta((\ln m/\ln b_0)^{(s-1)/2} m^{-3/2})$ along the subsequence $m = b_0^{k_*}$ with $k_* \to \infty$. (Here, we simplify notation by denoting the subsequence as $m$

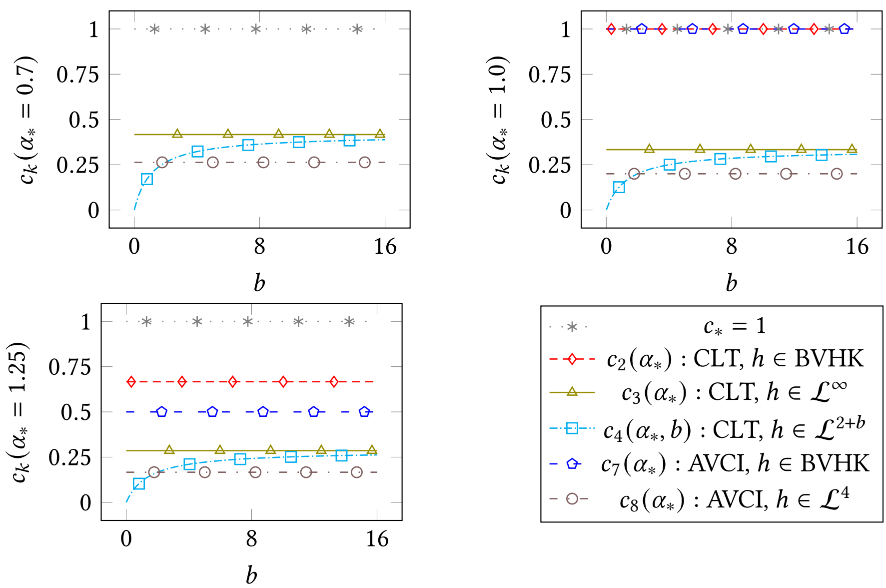


Fig. 1. Plots of the upper bounds $c_k(\alpha_*)$ in (42) of $c$ in Assumption 1.B for different Corollaries $k$. The plots display the $c_k(\alpha_*)$ as functions of $b$ from Assumption 3.C for different fixed values of $\alpha_*$. The upper left panel does not include $c_2(\alpha_*)$ and $c_7(\alpha_*)$ because these require $h \in$ BVHK, which then implies $\alpha_* \geq 1$ by (37). The plots show that stronger restrictions on the integrand $h$ lead to larger $c_k(\alpha_*)$.

rather than $m_*$, as in the discussion in the paragraph after (38).) Thus, taking the limit in (35) along the subsequence $m = b_0^{k_*}$ with $k_* \to \infty$, we get

$$\alpha_* = -\lim_{k_* \to \infty} \frac{\ln(\sigma_{b_0^{k_*}})}{\ln(b_0^{k_*})} = 3/2.$$

Now take the total computing budget as $n = b_0^{k_*}$ letting $k_* \to \infty$, and use an allocation $(m_n, r_n)$ with the number $r_n$ of randomizations growing very slowly to $\infty$, e.g., $r_n = n^{1-c}$ for $c \cong 1$ or $r_n = \ln n$, or even fixed $r_n = r_0$. Then under the conditions of [28, Theorem 1], the RMSE decreases at about rate $n^{-\alpha_*} = n^{-3/2}$. This rate is better than the rate of about $n^{-1}$ when $h \in$ BVHK from (51) through the CLT of Corollary 2, or the rate of about $n^{-5/6}$ from (57) through the AVCI of Corollary 7. But if we instead assume condition (22), our Corollaries 1 and 6 lead to the RMSE decreasing at about rate $n^{-3/2}$ by (48), the same as for [28, Theorem 1].

## C GRAPHICAL COMPARISONS OF THE $c_k(\alpha_*)$ AND THE $v_k(\alpha_*)$ FROM SECTION 5

Figure 1 shows the upper bounds $c_k(\alpha_*)$ in (42) for $c$ in Assumption 1.B, where $c_k(\alpha_*)$ is given by $c_2(\alpha_*)$ in (50) for Corollary 2, $c_3(\alpha_*)$ in (52) for Corollary 3, $c_4(\alpha_*, b)$ in (54) for Corollary 4, $c_7(\alpha_*)$ in (56) for Corollary 7, and $c_8(\alpha_*)$ in (58) for Corollary 8, where we recall that Table 2 of Section 5.8 provides a short summary of the corollaries. The plots display, for various fixed values of $\alpha_*$, the upper bounds as functions of $b > 0$ from Assumption 3.C ($h \in \mathcal{L}^{2+b}$). We further plot the upper bound $c_* = 1$ for reference because Assumption 1.B requires $c \in (0, 1)$, which is also seen through (50), (52), (54), (56), and (58). Note that $c_* = 1$ also corresponds to $c_1(\alpha_*) = c_6(\alpha_*) = 1$ in (48). We do not include $v_1(\alpha_*) = v_6(\alpha_*) = \alpha_*$ from (48) in the graphs to better see the differences among the other corollaries.

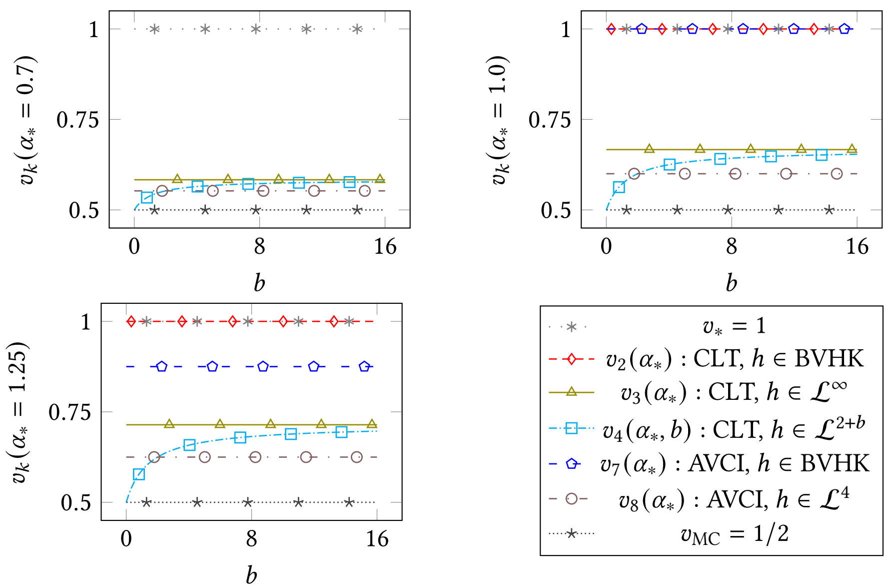


Fig. 2. Plots of the negative exponent $v_k(\alpha_*)$ of the optimal rate at which the estimator RMSE decreases as functions of $b$ for different values of $\alpha_*$. The upper left panel does not include $v_2(\alpha_*)$ and $v_7(\alpha_*)$ because these require $h \in$ BVHK, which then implies $\alpha_* \geq 1$ by (37). The plots show that stronger restrictions on the integrand $h$ lead to larger $v_k(\alpha_*)$.

Figure 1 corroborates the results obtained in Section B and summarized in Table 2. We can readily check the following:

- $c_4(\alpha_*, b) < c_3(\alpha_*) < c_2(\alpha_*)$ ($c_2(\alpha_*)$ being valid only when $\alpha_* \geq 1$), illustrating that the stricter the condition of the integrand $h$ (see Proposition 2.1), the larger the possible value of $c$ to ensure CLT (20) (Proposition 3.3).
- $c_8(\alpha_*) < c_7(\alpha_*)$, so we similarly see that a stronger condition on $h$ leads to larger range of values of $c$ that ensure AVCI (31); also see Proposition 4.3.
- $c_4(\alpha_*, b)$ approaches $c_3(\alpha_*)$ as $b$ grows large in Assumption 3.C, which agrees with the principle that having $h \in \mathcal{L}^{2+b}$ as $b \to \infty$ is "close" to meaning $h \in \mathcal{L}^\infty$; see the related discussion at the end of Section B.2. Similarly, as $b \to 0$, $c_4(\alpha_*, b)$ approaches $0 = c_5(\alpha_*)$ for Corollary 5, which assumes Assumption 3.D ($h \in \mathcal{L}^2$).
- 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 (31) 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 2 plots, as functions of $b > 0$ from Assumption 3.C, the (negative) exponent $v_k(\alpha_*)$ from (43) of the optimal ral rate at which the RQMC estimator's RMSE decreases, given by $v_2(\alpha_*)$ in (51) for Corollary 2, $v_3(\alpha_*)$ in (53) for Corollary 3, $v_4(\alpha_*, b)$ in (55) for Corollary 4, $v_7(\alpha_*)$ in (57) for Corollary 7, and $v_8(\alpha_*)$ in (59) for Corollary 8. Each of these exponents is at most $v_* = 1$, which is also plotted for reference. The figure further includes $v_{\mathrm{MC}} = 1/2$ from (41) for comparison, and all of these $v_k(\alpha_*)$ are strictly greater than $v_{\mathrm{MC}}$, so RQMC with $c > 0$ always has that the optimal rate

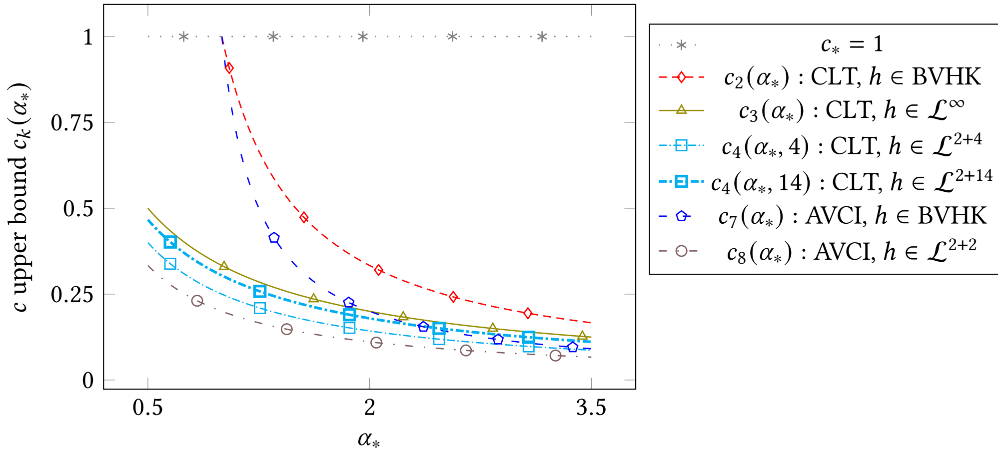


Fig. 3. Plots of the upper bounds $c_k(\alpha_*)$ of $c$ as functions of $\alpha_*$. Functions $c_2(\alpha_*)$ and $c_7(\alpha_*)$ require $h \in$ BVHK, so they are shown for only $\alpha_* \geq 1$ because of (37).

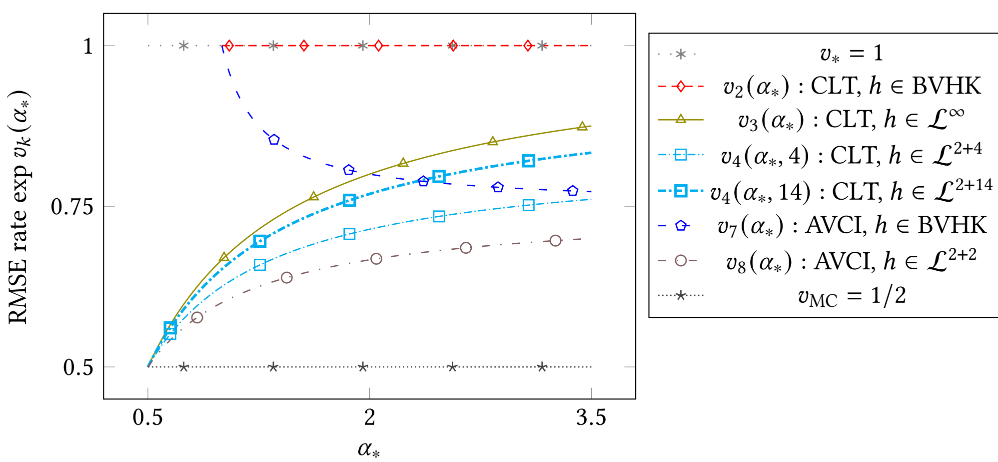


Fig. 4. Plots of the negative exponent $v_k(\alpha_*)$ of the optimal rate at which the estimator RMSE decreases as functions of $\alpha_*$. Functions $v_2(\alpha_*)$ and $v_7(\alpha_*)$ require $h \in$ BVHK, so they are shown for only $\alpha_* \geq 1$ because of (37).

at which its RMSE decreases is better than for MC. By (47), the comparisons and ordering of RMSE rate exponents $v_k(\alpha_*)$ are the same as the ones obtained before for the $c_k(\alpha_*)$.

Comparing across the different panels in Figure 1, we see that each upper bound $c_k(\alpha_*)$ in (42) on $c$ decreases as $\alpha_*$ increases. To investigate this further, Figures 3 and 4 plot the $c_k(\alpha_*)$ and $v_k(\alpha_*)$, respectively, as functions of $\alpha_*$. We can then see the differences between the various corollaries and assumptions as the QMC method improves (i.e., $\alpha_*$ increases).

Figure 3 more clearly illustrates that as $\alpha_*$ grows, each upper bound $c_k(\alpha_*)$ in (42) on $c$ decreases, so guaranteeing a CLT (20) or AVCI (31) 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 $\alpha_*$. However, Figure 4 shows that this is not the case for all CLT and most AVCI conditions: the RMSE optimal convergence speed determined by $v_k(\alpha_*)$, as in (43), is generally increasing in $\alpha_*$. The one exception is $v_7(\alpha_*)$, which determines the optimal RMSE convergence rate to ensure AVCI when $h \in \mathrm{BVHK}$. In this case, Figure 3 shows that as $\alpha_*$ increases, $c_7(\alpha_*)$ drops off quickly, so the number of randomizations must grow rapidly as $\alpha_*$ increases to ensure AVCI when $h \in \mathrm{BVHK}$. But even so, we still have $v_7(\alpha_*) > v_8(\alpha_*)$ for all $\alpha_*$, where $v_8(\alpha_*)$ in (59) is the optimal RMSE rate exponent for AVCI obtained under the moment condition ($h \in \mathcal{L}^4$) of Corollary 8.

## D 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 or AVCI seem 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 or close to nominal coverage for only finitely many 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 or a valid CI. 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 or AVCI. We will investigate this point.

We implemented our experiments in `python`, with the RQMC sequences generated using the `QMCPy` library [6]. We consider three different randomized sequences: Sobol' sequences with digital shift (DS) [19], scrambled Sobol' sequence with linear matrix scrambling (LMS) [16, 29], and Lattice rules with random shift (RS) [7, 21, 40]. 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 2.1 established a strict ordering of the restrictions in Assumption 3 on the integrand $h$. We now consider various functions associated with three of the cases of Assumption 3.

(1) For $u = (u_1, \dots, u_s) \in [0, 1]^s$, function

$$h_{\mathrm{v}}(u) = \prod_{l=1}^{s} \left(u_l - \frac{1}{2}\right)^2 \tag{67}$$

is of bounded Hardy-Krause variation (as $h_{\mathrm{v}}$ is a product of functions of bounded variation [4, 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 [4, Example 2.1.9]) and to be bounded. The product function for $u = (u_1, \dots, u_s) \in [0, 1]^s$ is

$$h_{\mathrm{b}}(u) = \prod_{l=1}^{s} g_{\mathrm{b}}(u_l), \tag{68}$$

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

$$h_{\mathrm{ub},\theta}(u) = \prod_{l=1}^{s} g_{\mathrm{ub},\theta}(u_l) \tag{69}$$

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 $\mathbb{E}[(h_{\mathrm{ub},\theta}(U))^{2+b}] = \left[\int_0^1 (u_1)^{-\theta(2+b)}\,\mathrm{d}u_1\right]^s$, 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.

(4) For $u = (u_1, \ldots, u_s) \in [0,1]^s$, function

$$h_{\mathrm{sum}}(u) = -s + \sum_{l=1}^{s} u_l \exp(u_l) \tag{70}$$

is smooth and additive, making it very easy for RQMC methods to integrate, as explained in Section 17.2 of [37]. This function is of bounded Hardy-Krause variation. In contrast to the other three integrands, $h_{\mathrm{sum}}$ does not have a product form.

To test whether a CLT or AVCI roughly holds for a given allocation $(m_n, r_n)$, we generated 200 independent values of $\widehat{\mu}^{\mathrm{RQ}}_{m_n,r_n}$ from (10) and the nominal 95% confidence interval $I^{\mathrm{RQ}}_{m_n,r_n,\gamma}$ from (29). We then investigated the proportion of the 200 times the true value $\mu$ of the integral was included in $I^{\mathrm{RQ}}_{m_n,r_n,\gamma}$ as an estimate of the CI's true coverage $\mathbb{P}(\mu \in I^{\mathrm{RQ}}_{m_n,r_n,\gamma})$. When the CLT (20) and AVCI (31) hold, we expect the observed coverage to be close to 0.95, with coverage error approximately $\pm\sqrt{0.95(1-0.95)/200} \approx 0.03$ based on a 95% confidence interval for the coverage.

## D.1 Estimation of $\alpha_*$

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 or AVCI, we need to estimate $\alpha_*$ in (35). 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 $\beta$ and $\alpha_*$, 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 $\mu$ from single randomizations of $m^{(i)}$ points. Then, using the simplifying notation $\ell_i = \ln(m^{(i)})$ and $v_i = \ln(\sigma_{m^{(i)}})$ for each $1 \le i \le K$, standard linear regression yields an estimator of $\alpha_*$ as

$$\widehat{\alpha}_* = -\frac{\sum_{i=1}^K (v_i - \hat{v}_K)(\ell_i - \bar{\ell}_K)}{\sum_{i=1}^K (\ell_i - \bar{\ell}_K)^2}, \quad \text{where} \quad \hat{v}_K = \frac{1}{K}\sum_{i=1}^K v_i \quad \text{and} \quad \bar{\ell}_K = \frac{1}{K}\sum_{i=1}^K \ell_i.$$

For our estimations, we used $R_0 = 100$, $K = 17$, and $m^{(i)} = 2^{i+5}$ for $i \in \{1, 2, \ldots, 17\}$.

Table 3 gives the estimated $\alpha_*$ for the three considered RQMC methods (Sobol' sequence with DS, Sobol' sequence with LMS, and Lattice with RS) and the four integrands ($h_{\mathrm{v}}$, $h_{\mathrm{b}}$, $h_{\mathrm{ub},0.35}$, and $h_{\mathrm{sum}}$) in dimensions $s = 3$ and $s = 4$. For the function $h_{\mathrm{v}} \in$ BVHK, the estimated values of $\alpha_*$ are larger than or equal to 1, agreeing with (37). The estimated $\alpha_*$ is very close to 1 for the additive $h_{\mathrm{sum}} \in$ BVHK for Sobol' with DS and Lattice with RS; thus, the bound in (37) appears to be tight. For $h_{\mathrm{sum}}$ using Sobol' with LMS, the regression provides a bigger estimated $\alpha_*$ but also a much larger constant $\beta$: $\beta = 16585$ for LMS in dimension 3 as opposed to $\beta = 1.33$ for Sobol' with DS and Lattice with RS. Figure 7 (also see [24, Section 4.2]) shows extremely spiky histograms for $h_{\mathrm{sum}}$ with Sobol' and LMS, indicating very low variance, which has been similarly observed in [20] in other contexts; but in general, the numerical values should be considered with caution. For functions $h_{\mathrm{b}}$ and $h_{\mathrm{ub},0.35}$, which have infinite Hardy-Krause variation, estimated values of $\alpha_*$ are smaller than 1 but larger than $1/2$ (see (38)), the exponential rate when using MC methods.

| Integrand | Dimension $s$ | Estimated $\alpha_*$ Sobol' DS | Sobol' LMS | Lattice RS |
|---|---|---|---|---|
| $h_{\mathrm{v}}$ | 3 | 1.3420 | 1.3040 | 1.4218 |
| $h_{\mathrm{b}}$ | 3 | 0.6843 | 0.6735 | 0.6173 |
| $h_{\mathrm{ub},0.35}$ | 3 | 0.6103 | 0.6217 | 0.5612 |
| $h_{\mathrm{sum}}$ | 3 | 1.0022 | 2.7554 | 0.9983 |
| $h_{\mathrm{v}}$ | 4 | 1.2620 | 1.1282 | 1.1528 |
| $h_{\mathrm{b}}$ | 4 | 0.6856 | 0.6412 | 0.6035 |
| $h_{\mathrm{ub},0.35}$ | 4 | 0.6115 | 0.5624 | 0.5437 |
| $h_{\mathrm{sum}}$ | 4 | 1.0035 | 2.7675 | 1.0009 |

Table 3. Estimated values of $\alpha_*$ in (35) from log-log regressions for different integrands, dimensions, and RQMC methods.

## D.2 Analysis of the impact of the allocation on coverage

In this subsection, we set the dimension as $s = 3$. 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 $t \in \{2, \ldots, 12\}$ to study a wide range of values of $c = t/14$. Figure 5 displays, as $c$ increases, the estimated coverage values of the produced nominal 95% confidence intervals $I^{\mathrm{RQ}}_{m_n,r_n,\gamma}$ from (29) for the integrands $h_{\mathrm{v}}$, $h_{\mathrm{b}}$, $h_{\mathrm{ub},0.35}$, and $h_{\mathrm{sum}}$ and the three types of randomization considered in Table 3. For the sake of readability, we do not draw the 95% confidence intervals on the coverages generated from the 200 independent experiments, but recall that for a true coverage value 0.95 (resp., 0.85), the resulting error of the coverage is $\pm 0.03$ (resp., $\pm 0.05$) (based on a 95% confidence interval for the coverage).

- For all integrands and randomizations, coverage is close to the expected 0.95 when $c$ is small. Then it drops significantly when $c$ is larger than 0.8 for all functions. Note though that, as expected, the stricter the assumption on the integrand, the larger the threshold value $c_k(\alpha_*)$ above which coverage is statistically significantly below 0.95.
- For function $h_{\mathrm{v}}$ in (67), up to value $c = 0.786$, we do not detect any suspect coverage values, but when $c = 0.857$, the estimated coverage levels drop to 0.815, 0.87 and 0.90 for the various randomizations, which lead to rejecting the null hypothesis of 95% coverage (at 95% level of confidence). Theoretically, as $h_{\mathrm{v}}$ satisfies Assumption 3.A, Corollaries 2 and 7 apply; thus, taking $\alpha_* = 1.35$ as an approximation from Table 3 for (35), we get $c_2(\alpha_*) \approx 0.588$ by (50) and $c_7(\alpha_*) \approx 0.417$ by (56), smaller than the 0.8 from the numerical experiments. This perhaps indicates that the sufficient condition (24) (resp., (33)) of Corollary 2 (resp., 7) is stronger than necessary to ensure a CLT (resp., AVCI).
- For function $h_{\mathrm{b}}$ in (68), which is bounded (so Assumption 3.B holds) but has infinite Hardy-Krause variation, an under-coverage is statistically detected at smaller values of $c$ than for $h_{\mathrm{v}}$, i.e., at $c = 0.786$ for Sobol' sequence with DS or LMS, while lattices with random shift has (statistically) acceptable value for this $c$ but not anymore when $c = 0.857$. Theoretically, as Assumption 3.B holds, Corollaries 3 and 8 apply, and taking $\alpha_* = 0.65$ (approximately what Table 3 shows) results in $c_3(\alpha_*) \approx 0.435$ and $c_8(\alpha_*) \approx 0.278$ by (52) and (58). This may again indicate that the sufficient condition (25) (resp., (34)) of Corollary 3 (resp., 8) is stronger than necessary to ensure a CLT (resp., AVCI). Using a larger value of $n$ would be of interest, but the computations already take hours for $n = 2^{14}$.

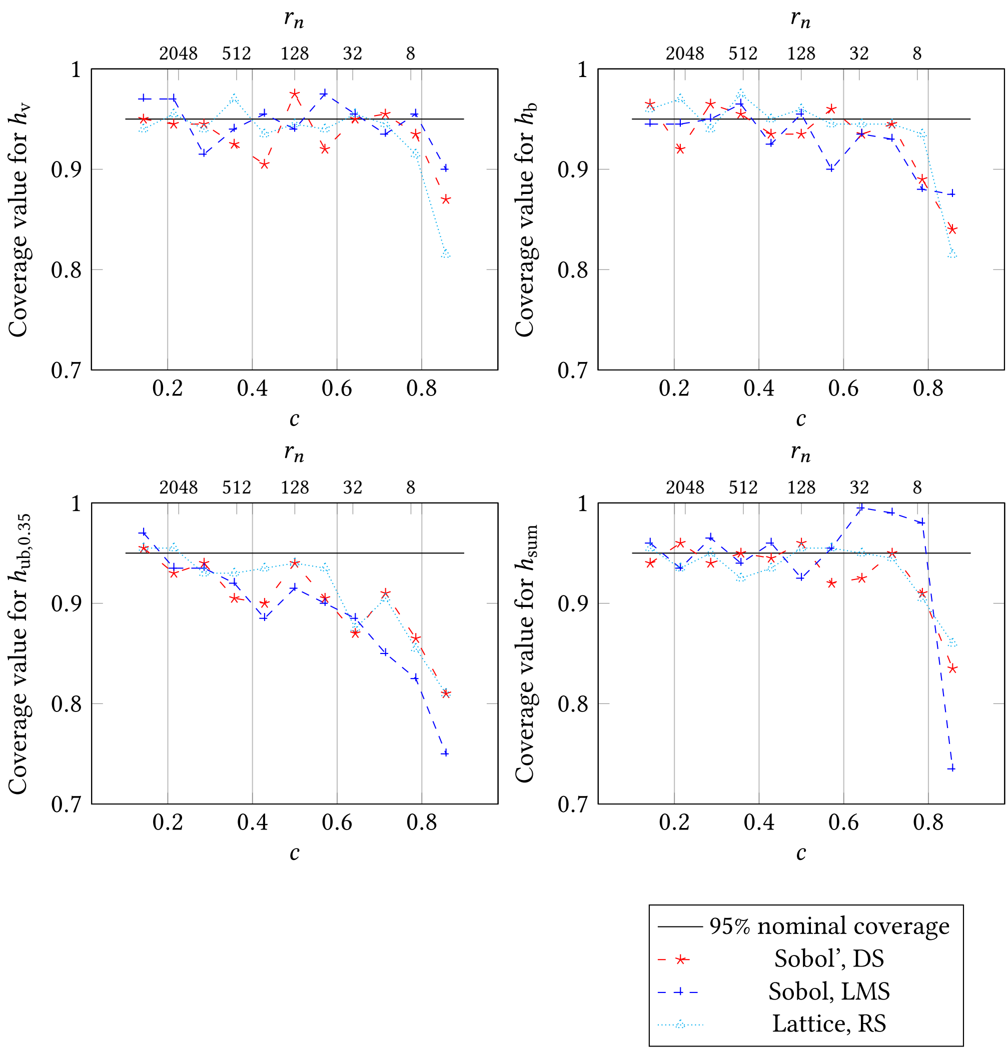


Fig. 5. Coverage values from 200 independent estimations for various values of $c = t/14$ for $2 \le t \le 12$ and for three RQMC techniques with $n = 2^{14}$ and $(m_n, r_n) = (n^t, n^{14-t})$. The considered integrand is $h_v$ (top left), $h_b$ (top right) , $h_{ub,0.35}$ (middle left) and $h_{sum}$ (middle right), each for dimension $s = 3$.

- For function $h_{ub,0.35}$ in (69), coverage is generally below the nominal 0.95 as soon as $c > 0.22$; it seems difficult to numerically determine a threshold value, but coverage is below those for $h_v$ and $h_b$, and becomes as low as 0.75 when $c = 0.857$ for Sobol' sequence and LMS. The function $h_{ub,0.35}$ is unbounded but $h_{ub,0.35} \in \mathcal{L}^{2+b}$ for $b < 1/0.35 - 2 \approx 0.857$, so Assumption 3.C holds for these values of $b$. Corollary 4 then applies, and using, say, $\alpha_* = 0.5612$ from Table 3 (for Lattice with RS but using other values do not significantly change the results) gives $c_4(0.5612, 1/0.35 - 2) \approx 0.211$ from (54). This seems corroborated by the observed coverage values for most values of $c$. (As Assumption 3.C does not hold for $b = 2$, Corollary 8 is not applicable.)

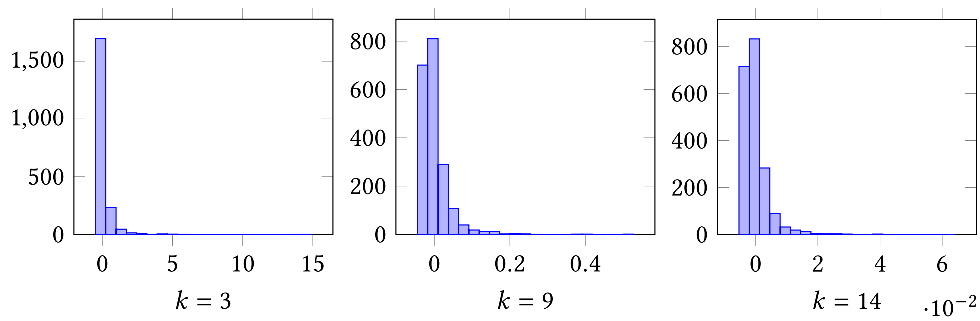


Fig. 6. Histograms of 2000 centered RQMC values $A_m - \mu$ for the different $m = 2^k$ Sobol' points with LMS on function $h_{\mathrm{ub},0.35}$

- For function $h_{\mathrm{sum}}$ in (70), no suspect coverage is detected until $c = 0.857$, the coverage dropping then to values below 0.86. Taking $\alpha_* = 1$ (resp., 2.75) for Sobol' with DS and for Lattice with RS (resp., Sobol' with LMS) from Table 3, we get $c_2(\alpha_*) = 1$ (resp., 0.221) from (50) and $c_7(\alpha_*) = 1$ (resp., 0.124) from (57). Thus, for Sobol' with LMS, our sufficient condition seems to be stronger than necessary. For Sobol' with DS and Lattice with RS for which $r_n$ can increase polynomially as slowly as we wish to infinity with $n$ to get a CLT, it may indicate that the number $r$ of randomizations has to be large enough to obtain a reasonable variance estimate, which may not occur when $r$ is too small (we indeed have $r = 4$ for $c = 0.857$ here).

Remember that for $r$ fixed, the centered and scaled RQMC estimator may not converge (weakly) in general to a normal random variable as $m \to \infty$; e.g., [23] prove a non-Gaussian limit for the random shift with lattices for a single randomization $r = 1$, so the same is true for any fixed $r \geq 1$. To further study the unsatisfactory coverage when $r = r_n$ is small in Figure 5, we also constructed histograms of 2000 values of $A_m$ (defined before Corollary 1 in Section 3.1) for different values of $m = 2^k$ for functions $h_{\mathrm{ub},0.35}$ and $h_{\mathrm{sum}}$, which are the ones in Figure 5 for which the coverage suffers the most for small $r_n$, and our three types of randomization considered in Table 3. Figures 6, 7 and 8 show a representative subset of the results. Although the plots within each figure often share similar shapes, the ranges on the horizontal axes become much smaller as $k$ increases, demonstrating RQMC's greatly increased precision for larger $k$. But Figures 6 and 7 exhibit extreme skewness or kurtosis, which, as seen through Edgeworth expansions [12, Example 2.1], can lead to a CLT providing a poor normal approximation and sizable CI coverage error when the sample size is not large; this coincides with the substandard behavior for small $r_n$ in Figure 5.

For example, Figure 6 presents histograms for the integrand $h_{\mathrm{ub},0.35}$ when using Sobol' with LMS; the plots for the other randomization methods with $h_{\mathrm{ub},0.35}$ are similar so are not shown. The graphs illustrate that $A_m$ has a non-normal and highly skewed distribution for function $h_{\mathrm{ub},0.35}$ whatever the randomization type, even for large $m$. This aligns with the degrading coverage as $r_n$ shrinks in the middle left panel of Figure 5.

Figures 7 and 8 display histograms for $h_{\mathrm{sum}}$ when applying Sobol' with LMS and DS, respectively; the plots for $h_{\mathrm{sum}}$ when applying Lattice with RS are similar to those in Figure 8, so are omitted. Figure 7 shows increasing spikiness (growing kurtosis) as $m = 2^k$ increases, with extreme skewness for the largest $m$. (Additional omitted plots with other values of $k$ reveal that the skewness becomes apparent starting around $k = 11$ and gets more pronounced as $k$ grows.) Correspondingly, the middle right panel of Figure 5 indicates that the coverage for Sobol'-LMS suffers greatly when $r_n$ is small (so $m_n$ is large). In contrast, the histograms for Sobol' with DS (Figure 8) and Lattice with RS

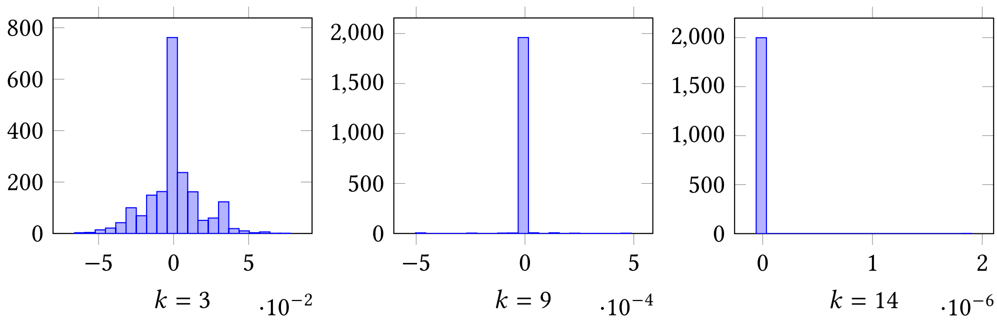


Fig. 7. Histograms of 2000 centered RQMC values $A_m - \mu$ for the different $m = 2^k$ Sobol' points with LMS on function $h_{\text{sum}}$

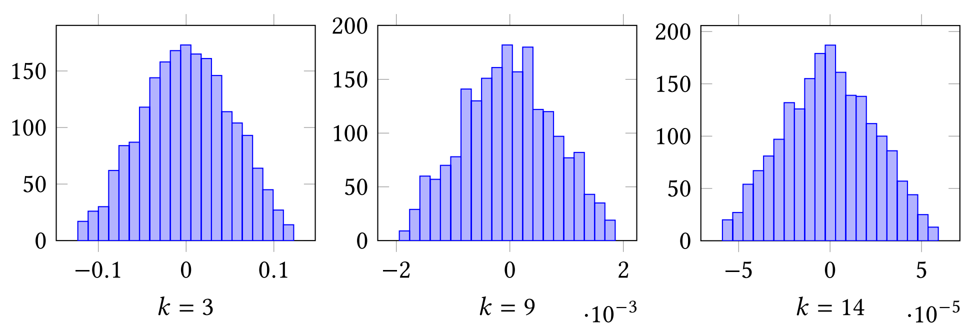


Fig. 8. Histograms of 2000 centered RQMC values $A_m - \mu$ for the different $m = 2^k$ Sobol' points with DS on function $h_{\text{sum}}$

are symmetric and have somewhat of a Gaussian appearance (although for large $k$, the tails seem to end more abruptly than for a normal and perhaps more akin to a triangular distribution). The minimal skewness and kurtosis in Figure 8 is consistent with the middle right panel of Figure 5, where coverage with small $r_n$ does not suffer as much for these RQMC methods as for Sobol' with LMS.

### D.3 Analysis in terms of the dimension

We aim here at investigating the coverage in terms of the dimension $s$ for two integrands, $h_{\text{ub},0.35}$ and $h_{\text{sum}}$, respectively a product and a sum of functions of individual coordinates. RQMC efficiency is expected to deteriorate as the dimension increases for $h_{\text{ub},0.35}$ and to be independent of the dimension for $h_{\text{sum}}$ because the *effective dimension* [19, 37] increases for the former and not for the latter. The effective dimension measures the contribution to the variance from low-dimensional subsets of coordinates, but it is not clear if the metric can be used to say anything about the quality of a normal approximation or a CI having close to nominal coverage.

Figure 9 displays the evolution of the coverage in terms of the dimension $s$ when we fix $r = 2^5$ and $m = 2^9$, so that $c \approx 0.643$ and to consider $n = 2^{14} = 16384$ as in Section D.2. We again consider 200 independent experiments for the estimation of the coverage. Coverage seems independent of the dimension for $h_{\text{sum}}$ but to decrease with it for $h_{\text{ub},0.35}$. It suggests that a low-effective dimension might have some connections with asymptotic coverage properties, but further theoretical study is needed.

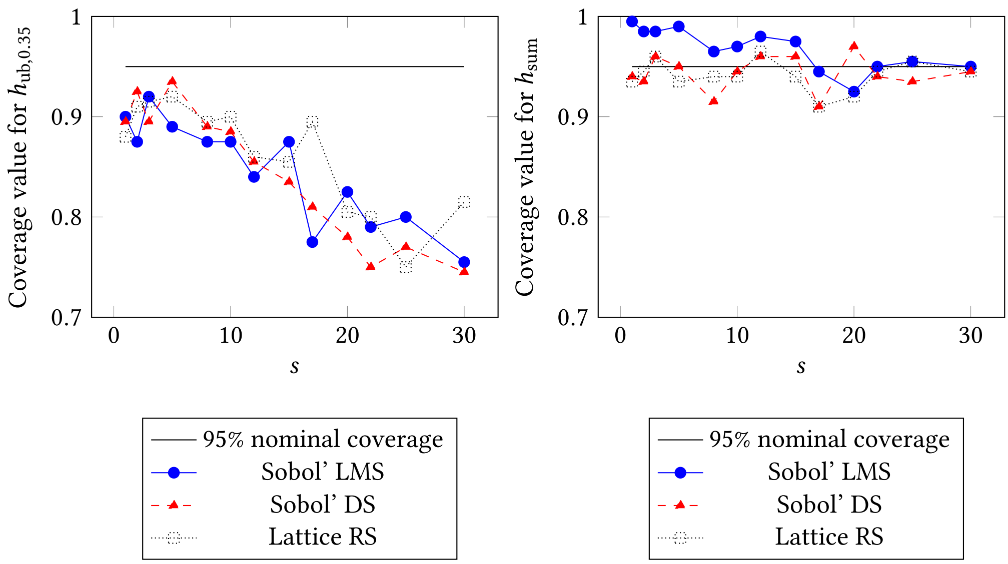


Fig. 9. Evolution of the coverage in terms of the dimension $s$ for the three randomization types and the two integrands $h_{\mathrm{ub},0.35}$ (left panel) and $h_{\mathrm{sum}}$ (right panel). The estimation is based on $r = 2^5$ and $m = 2^9$.

### D.4 On the validity of Condition (22)

If the conditions of Corollary 1 are satisfied, then a CLT holds for $r_n$ growing arbitrarily slowly to infinity. Unfortunately, it appears difficult to prove rigorously that Conditions (22) or (23) holds for specific functions, including those described in Section D. Instead, our goal here is to investigate *numerically* whether (22) appears to hold or not. For integrands $h_{\mathrm{ub},0.35}$ and $h_{\mathrm{sum}}$ in respectively (69) and (70), we plot estimates of $\frac{\mathbb{E}\left[|A_m-\mu|^3\right]}{\sigma_m^3}$ (i.e., with $d = 1$ in (22)) as $m$ increases to see if it seems to be bounded in $m$. Both $\mathbb{E}\left[|A_m - \mu|^3\right]$ and $\sigma_m^2$ are estimated by considering $r_0 = 2^{12}$ independent replications of $A_m$ used for both the numerator and the denominator, using the known expected value $\mu$. Figure 10 displays the results, which provide *suggestive*, but by no means definitive, evidence. It is not so easy to discern trends in the plots for $h_{\mathrm{ub},0.35}$ in the left panel of Figure 10 because they exhibit somewhat chaotic behavior, which perhaps can be explained as follows. The variance of an estimate of the $q$th-order moment depends on the $2q$th moment, which can make estimating higher-order moments difficult. As we are considering the 3rd absolute central moment in the numerator, this leads to noisy estimates of the ratio $\mathbb{E}\left[|A_m - \mu|^3\right]/\sigma_m^3$, even though the sample size $r_0 = 4096$ is reasonably large.

The plots for $h_{\mathrm{sum}}$ in the right panel of Figure 10 possess some features that are easier to interpret. First, the curves for Sobol' DS and Lattice RS are roughly constant, suggesting that (22) holds. But for the Sobol' LMS randomization, we see that the estimation smoothly increases up to a certain value of $m$, at which point it suddenly drops. This type of behavior often occurs in rare-event simulation for plots of estimated relative errors of estimators without bounded relative error; see [10, Section 4.1.1]. Specifically, when showing a curve of the empirical relative error in terms of a rarity parameter for a fixed sample size, we are able to obtain observations of the rare event for non-extreme values of the rarity parameter, and the curve increases smoothly for those values. But then at some threshold, the rare event is no longer observed in a sample, leading to an estimator that is orders of magnitude off from the true value of the estimand. It is in our context characteristic

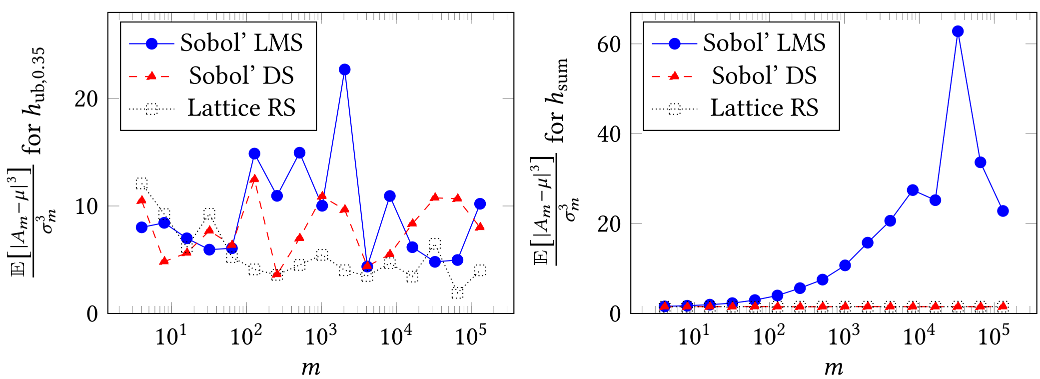


Fig. 10. Estimates of $\frac{\mathbb{E}[|A_m-\mu|^3]}{\sigma_m^3}$ for various $m$ for the three considered randomization types and the two integrands $h_{\mathrm{ub},0.35}$ (left panel) and $h_{\mathrm{sum}}$ (right panel). Each estimate is based on $r_0 = 2^{12}$ independent observations of $A_m$.

of an unbounded $\frac{\mathbb{E}[|A_m-\mu|^3]}{\sigma_m^3}$, as also suggested by the extreme spikiness and skewness of the histogram for large $m = 2^k$ in Figure 7, so that (22)) does *not* appear to be satisfied for Sobol' with LMS. This is in line with our discussion after Corollary 1 that according to [37, Chapter 17 end notes], skewness and kurtosis could diverge as $m \longrightarrow \infty$.

## REFERENCES


[1] S. Asmussen and P. Glynn. 2007. *Stochastic Simulation: Algorithms and Analysis.* Springer, New York.

[2] Kinjal Basu and Rajarshi Mukherjee. 2017. Asymptotic Normality of Scrambled Geometric Net Quadrature. *Annals of Statistics* 45, 4 (aug 2017), 1759–1788. https://doi.org/10.1214/16-aos1508

[3] P. Billingsley. 1995. *Probability and Measure* (3rd ed.). John Wiley and Sons, New York.

[4] S. Breneis. 2020. *Functions of Bounded Variation in One and Multiple Dimensions.* Master's thesis. Johannes Kepler Universitat Linz, Austria. https://epub.jku.at/obvulihs/download/pdf/4951527?originalFilename=true

[5] G. Casella and R. L. Berger. 2002. *Statistical Inference* (2nd ed.). Duxbury, Pacific Grove, Calif.

[6] S.-C. T. Choi, F. J. Hickernell, M. McCourt, and A. Sorokin. 2020+. QMCPy: A quasi-Monte Carlo Python Library. https://github.com/QMCSoftware/QMCSoftware

[7] R. Cranley and T. N. L. Patterson. 1976. Randomization of Number Theoretic Methods for Multiple Integration. *SIAM J. Numer. Anal.* 13, 6 (1976), 904–914.

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

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

[10] P.W. Glynn, G. Rubino, and B. Tuffin. 2009. Robustness properties and confidence interval reliability issues. In *Rare Event Simulation Using Monte Carlo Methods*, G. Rubino and B. Tuffin (Eds.). John Wiley & Sons Ltd., Chichester, UK, 63–84.

[11] P. W. Glynn. 1987. Limit Theorems for the Method of Replication. *Stochastic Models* 3, 3 (1987), 343–355.

[12] P. Hall. 1992. *The Bootstrap and Edgeworth Expansions.* Springer, New York.

[13] Zhijian He and Lingjiong Zhu. 2017. Asymptotic Normality of Extensible Grid Sampling. *Statistics and Computing* 29, 1 (nov 2017), 53–65. https://doi.org/10.1007/s11222-017-9794-y

[14] F. J. Hickernell. 1998. A Generalized Discrepancy and Quadrature Error Bound. *Math. Comp.* 67, 221 (1998), 299–322.

[15] F. J. Hickernell. 2018. Error Analysis for Quasi-Monte Carlo Methods. In *Monte Carlo and Quasi-Monte Carlo Methods: MCQMC 2016*, P. W. Glynn and A. B. Owen (Eds.). Springer, Berlin, 3–27.

[16] H. S. Hong and F. H. Hickernell. 2003. Algorithm 823: Implementing Scrambled Digital Sequences. *ACM Trans. Math. Software* 29 (2003), 95–109.

[17] D. P. Kroese, T. Taimre, and Z. I. Botev. 2011. *Handbook of Monte Carlo Methods.* John Wiley and Sons, New York.

[18] L. Kuipers and H. Niederreiter. 1974. *Uniform Distribution of Sequences.* John Wiley, New York.

[19] P. L'Ecuyer. 2018. Randomized Quasi-Monte Carlo: An Introduction for Practitioners. In *Monte Carlo and Quasi-Monte Carlo Methods: MCQMC 2016*, P. W. Glynn and A. B. Owen (Eds.). Springer, Berlin, 29–52.

[20] P. L'Ecuyer, V. Demers, and B. Tuffin. 2007. Rare Events, Splitting, and quasi-Monte Carlo. *ACM Transactions on Modeling and Computer Simulation* 17, 2 (apr 2007), Article 9. https://doi.org/10.1145/1225275.1225280

[21] P. L'Ecuyer and C. Lemieux. 2000. Variance Reduction via Lattice Rules. *Management Science* 46, 9 (2000), 1214–1235.

[22] P. L'Ecuyer and C. Lemieux. 2002. Recent Advances in Randomized Quasi-Monte Carlo Methods. In *Modeling Uncertainty: An Examination of Stochastic Theory, Methods, and Applications*, M. Dror, P. L'Ecuyer, and F. Szidarovszky (Eds.). Kluwer Academic, Boston, 419–474.

[23] P. L'Ecuyer, D. Munger, and B. Tuffin. 2010. On the Distribution of Integration Error by Randomly-Shifted Lattice Rules. *Electronic Journal of Statistics* 4 (2010), 950–993.

[24] P. L'Ecuyer, M. K. Nakayama, A. B. Owen, and B. Tuffin. 2023. Confidence Intervals for Randomized Quasi-Monte Carlo Estimators. In *Proceedings of the 2023 Winter Simulation Conference*, C.G. Corlu, S.R. Hunter, H. Lam, B.S. Onggo, J. Shortle, and B. Biller (Eds.). Institute of Electrical and Electronics Engineers, Piscataway, New Jersey.

[25] E. L. Lehmann. 1999. *Elements of Large-Sample Theory.* Springer, New York.

[26] C. Lemieux. 2006. Quasi-Random Number Techniques. In *Simulation*, S. G. Henderson and B. L. Nelson (Eds.). Elsevier, Amsterdam, The Netherlands, 351–380. Chapter 12.

[27] C. Lemieux. 2009. *Monte Carlo and Quasi-Monte Carlo Sampling.* Springer, New York.

[28] W.-L. Loh. 2003. On the Asymptotic Distribution of Scrambled Net Quadrature. *Annals of Statistics* 31, 4 (2003), 1282–1324.

[29] J. Matoušek. 1998. On the $L_2$-discrepancy for Anchored Boxes. *J. of Complexity* 14 (1998), 527–556.

[30] M. K. Nakayama and B. Tuffin. 2021. Sufficient Conditions for a Central Limit Theorem to Assess the Error of Randomized Quasi-Monte Carlo Methods. In *Proceedings of the 2021 Winter Simulation Conference.* IEEE. Accepted..

[31] H. Niederreiter. 1992. *Random Number Generation and Quasi-Monte Carlo Methods.* Vol. 63. SIAM, Philadelphia.

[32] A.B. Owen. 2005. Multidimensional Variation for Quasi-Monte Carlo. In *Contemporary Multivariate Analysis and Design of Experiments.* World Scientific, 49–74. https://doi.org/10.1142/9789812567765_0004

[33] A.B. Owen and D. Rudolf. 2020. A strong law of large numbers for scrambled net integration. (2020). arXiv:2002.07859 [math.NA] https://arxiv.org/abs/2002.07859.

[34] A. B. Owen. 1995. Randomly Permuted (t,m,s)-nets and (t,s)-sequences. In *Monte Carlo and Quasi-Monte Carlo Methods in Scientific Computing: Lecture Notes in Statistics.* Vol. 106. Springer, 299–317.

[35] A. B. Owen. 1997. Monte Carlo variance of scrambled net quadrature. *SIAM Journal of Numerical Analysis* 34 (1997), 1884–1910.

[36] A. B. Owen. 1997. Scrambled net variance for integrals of smooth functions. *Annals of Statistics* 25, 4 (1997), 1541–1562.

[37] A. B. Owen. 2019. *Monte Carlo Theory, Methods and Examples.* Draft available at http://statweb.stanford.edu/~owen/mc/. In preparation.

[38] R. J. Serfling. 1980. *Approximation Theorems of Mathematical Statistics.* John Wiley and Sons, New York.

[39] B. Tuffin. 1997. Variance Reduction Applied to Product-Form Multi-Class Queuing Networks. *ACM Transactions on Modeling and Computer Simulation* 7, 4 (1997), 478–500.

[40] B. Tuffin. 1998. Variance Reduction Order Using good Lattice points in Monte Carlo Methods. *Computing* 61, 4 (1998), 371–378.

[41] B. Tuffin. 2004. Randomization of Quasi-Monte Carlo Methods for Error Estimation: Survey and Normal Approximation. *Monte Carlo Methods and Applications* 10, 3-4 (2004), 617–628.