



**HAL**  
open science

# On a Central Limit Theorem and a Confidence Interval for Randomized Quasi-Monte Carlo Methods

Marvin Nakayama, Bruno Tuffin

► **To cite this version:**

Marvin Nakayama, Bruno Tuffin. On a Central Limit Theorem and a Confidence Interval for Randomized Quasi-Monte Carlo Methods. Stochastic Numerics and Statistical Learning: Theory and Applications Workshop, KAUST, May 2022, Djeddah, Saudi Arabia. pp.1-34. hal-03678150

**HAL Id: hal-03678150**

**<https://inria.hal.science/hal-03678150>**

Submitted on 25 May 2022

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

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



Distributed under a Creative Commons Attribution 4.0 International License

# On a Central Limit Theorem and a Confidence Interval for Randomized Quasi-Monte Carlo Methods

Marvin K. Nakayama<sup>1</sup>, and Bruno Tuffin<sup>2</sup>

<sup>1</sup>New Jersey Institute of Technology

<sup>2</sup>Inria, Rennes, France



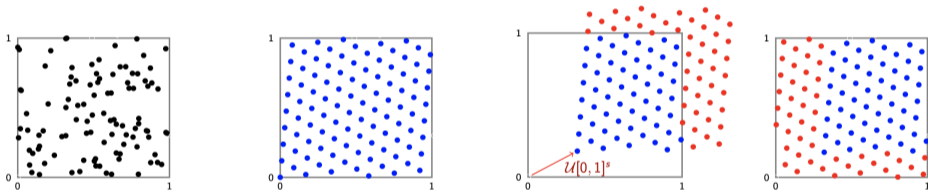
Stochastic Numerics and Statistical Learning workshop  
May 2022

## Motivation

- Estimate  $\mu$  and provide **computable measure of error**, where

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

- Monte Carlo (MC), Quasi-Monte Carlo (QMC), **Randomized QMC (RQMC)**



- Central limit theorem (CLT) not proven for many RQMC settings.**
- Goal:** General conditions for RQMC to ensure
  - CLT
  - Asymptotically valid confidence interval (AVCI)**

## Table of contents

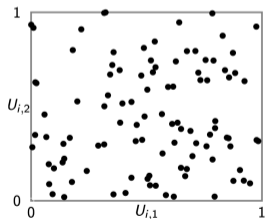
- 1 Review: MC, QMC, RQMC
- 2 RQMC CLT and AVCI
- 3 Corollaries Ensuring CLT and AVCI: Tradeoffs
- 4 Concluding Remarks

## Review: Monte Carlo (MC)

- MC: **random sampling** to estimate  $\mu = \mathbb{E}[h(U)]$

$$\hat{\mu}_n^{\text{MC}} = \frac{1}{n} \sum_{i=1}^n h(U_i)$$

- $U_1, U_2, \dots, U_n$  i.i.d.  $\mathcal{U}[0, 1]^s$



## Review: MC — Error Estimation Easy, But Slow Convergence

- MC estimator:  $\hat{\mu}_n^{\text{MC}} = \frac{1}{n} \sum_{i=1}^n h(U_i)$
- **CLT**: If  $\psi^2 \equiv \text{Var}[h(U)] \in (0, \infty)$ , then [Billingsley 1995]

$$\sqrt{\frac{n}{\psi^2}} \left[ \hat{\mu}_n^{\text{MC}} - \mu \right] \Rightarrow \mathcal{N}(0, 1) \quad \text{as } n \rightarrow \infty$$

- Approximate  $100\gamma\%$  confidence interval (CI) for  $\mu$ :

$$I_{n,\gamma}^{\text{MC}} \equiv \left[ \hat{\mu}_n^{\text{MC}} \pm z_\gamma \frac{\hat{\psi}_n}{\sqrt{n}} \right]$$

▶  $\hat{\psi}_n^2 = \frac{1}{n-1} \sum_{i=1}^n [h(U_i) - \hat{\mu}_n^{\text{MC}}]^2$  and  $\Phi(z_\gamma) = 1 - (1 - \gamma)/2$ .

- **Asymptotically valid CI (AVCI)**:

$$\mathbb{P}(\mu \in I_{n,\gamma}^{\text{MC}}) \rightarrow \gamma, \quad \text{as } n \rightarrow \infty$$

- Root mean-squared error:  $\text{RMSE}[\hat{\mu}_n^{\text{MC}}] = \frac{\psi}{\sqrt{n}}$

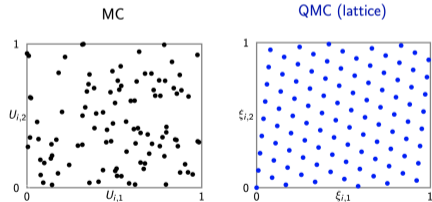
## Review: Quasi-Monte Carlo (QMC)

- QMC: **deterministic points** to estimate  $\mu = \mathbb{E}[h(U)]$

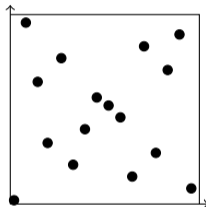
$$\hat{\mu}_n^Q = \frac{1}{n} \sum_{i=1}^n h(\xi_i)$$

- **Low-discrepancy sequence**  $\Xi = (\xi_i : i = 1, 2, \dots)$

- ▶  $\Xi$  is **deterministic** and evenly fill  $[0, 1]^s$
- ▶ **lattices** (e.g., Korobov, ...),



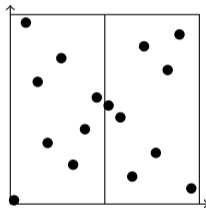
- **Digital nets/sequences** (e.g., Sobol', Faure, ...)



- Is this 16 points sequence well-distributed?

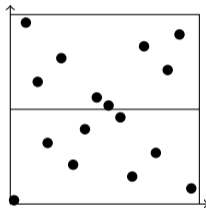


- **Digital nets/sequences** (e.g., Sobol', Faure, ...)



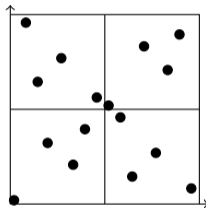
- If we split  $[0, 1]^2$ , we get the right number of points in the rectangles.

- **Digital nets/sequences** (e.g., Sobol', Faure, ...)



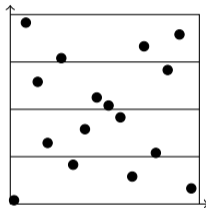
- If we split  $[0, 1]^2$ , we get the right number of points in the rectangles.

- **Digital nets/sequences** (e.g., Sobol', Faure, ...)



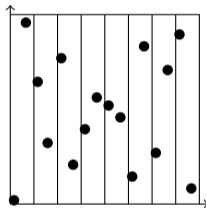
- If we split  $[0, 1]^2$ , we get the right number of points in the rectangles.

- **Digital nets/sequences** (e.g., Sobol', Faure, ...)



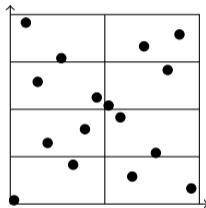
- If we split  $[0, 1]^2$ , we get the right number of points in the rectangles.

- **Digital nets/sequences** (e.g., Sobol', Faure, ...)



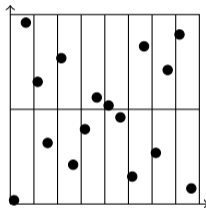
- If we split  $[0, 1]^2$ , we get the right number of points in the rectangles.

- **Digital nets/sequences** (e.g., Sobol', Faure, ...)



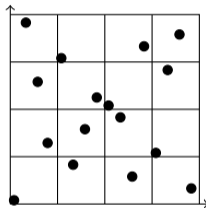
- If we split  $[0, 1]^2$ , we get the right number of points in the rectangles.

- **Digital nets/sequences** (e.g., Sobol', Faure, ...)



- If we split  $[0, 1]^2$ , we get the right number of points in the rectangles.

- **Digital nets/sequences** (e.g., Sobol', Faure, ...)



- If we split  $[0, 1]^2$ , we get the right number of points in the rectangles.



## Review: QMC — Fast Convergence

- QMC: **deterministic points** to estimate  $\mu = \mathbb{E}[h(U)]$

$$\hat{\mu}_n^Q = \frac{1}{n} \sum_{i=1}^n h(\xi_i), \quad \Xi = (\xi_i : i = 1, 2, \dots)$$

- **Koksma-Hlawka (K-H) inequality** [Niederreiter 1992]: for each  $n > 1$ ,

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

- ▶ **Hardy-Krause variation**  $V_{\text{HK}}(h) \in [0, \infty]$ : “roughness” of  $h$
- ▶ **Star-discrepancy**  $D_n^*(\Xi) \in [0, 1]$ : how unevenly first  $n$  points of  $\Xi$  fill  $[0, 1]^s$

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

- ▶ If  $V_{\text{HK}}(h) < \infty$  (BVHK), then K-H bound shrinks at faster rate than MC rate  $\Theta(n^{-1/2})$

$$|\hat{\mu}_n^Q - \mu| \approx O(n^{-1}).$$

- ★ BVHK: “bounded variation in sense of Hardy and Krause”

## Review: QMC — But Error Estimation Difficult

Two components:  $V_{\text{HK}}(h)$  and  $D_n^*(\Xi)$ .

- Variation

- ▶ Difficult to compute: sum of  $2^s - 1$  variations in the sense of Vitali (each difficult to estimate).
- ▶  $V_{\text{HK}}(h) = \infty$  often, even for simple functions such as

$$h(x_1, x_2, x_3) = \min(x_1 + x_2 + x_3, 1)$$

for which QMC converges (faster than for MC).

- Discrepancy

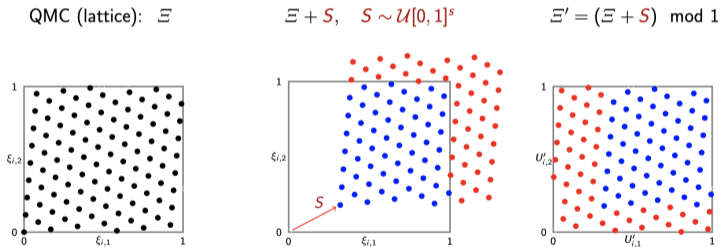
- ▶ Expression  $D_n^*(\Xi) = \sup_{x \in [0,1]^s} \left| \frac{\sum_{i=1}^n \mathbb{1}_{\prod_{k=1}^s [0, x_k]}(\xi_i)}{n} - \prod_{k=1}^s x_k \right|$ .
- ▶ We only have bounds with asymptotic usefulness, not for practice use.
- ▶ Ex: bounds for the  $(0, s)$ -Niederreiter sequence (while we know  $D_n^*(\Xi) < 1!$ )

$s = 5$	$n$	$10^3$	$10^4$	$10^6$	$10^9$	$10^{12}$	$10^{16}$
	Bound	1.50	$3.34 \cdot 10^{-1}$	$2.14 \cdot 10^{-2}$	$1.21 \cdot 10^{-4}$	$5.83 \cdot 10^{-7}$	$1.92 \cdot 10^{-10}$
$s = 25$	$n$	$10^3$	$10^9$	$10^{15}$	$10^{21}$	$10^{35}$	
	Bound	$2.88 \cdot 10^5$	$1.11 \cdot 10^6$	$2.46 \cdot 10^5$	$5.22 \cdot 10^3$	$1.62 \cdot 10^{-3}$	

- Other bounds exist, in general with the same issues.

## Review: Randomized Quasi-Monte Carlo (RQMC)

- i.i.d. randomizations of  $\Xi = (\xi_i : i \geq 1)$ , each yielding  $\Xi' = (U'_i : i \geq 1)$ 
  - ▶ Each  $U'_i \sim \mathcal{U}[0, 1]^s$
  - ▶  $\Xi'$  retains low-discrepancy properties of  $\Xi$
- **Lattice:** random shift [Cranley & Patterson 1976]



- **Digital net:** nested scrambling [Owen 1995], digital shift [L'Ecuyer & Lemieux 2002], ...

## Review: Randomized Quasi-Monte Carlo (RQMC)

- RQMC computation budget of  $n$  evaluations of  $h$  (as for MC)
  - ▶ **allocation**  $(m_n, r_n)$  with  $m_n \times r_n \approx n$
  - ▶  $r_n = \#$  i.i.d. randomizations
  - ▶  $m_n = \#$  points used from  $j$ th randomized sequence  $\Xi'_j = (U'_{i,j} : i \geq 1)$ ,  $j = 1, 2, \dots, r_n$
- RQMC:  $r_n \geq 2$  i.i.d. **randomizations** to estimate  $\mu = \mathbb{E}[h(U)]$

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

- ▶  $X_{n,1}, X_{n,2}, \dots, X_{n,r_n}$  i.i.d.: estimate  $\sigma_{m_n}^2 \equiv \text{Var}[X_{n,1}]$  by

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

- Approx  $\gamma$ -level CI for  $\mu$

$$I_{m_n, r_n, \gamma}^{\text{RQ}} \equiv \left[ \hat{\mu}_{m_n, r_n}^{\text{RQ}} \pm z_\gamma \frac{\hat{\sigma}_{m_n, r_n}}{\sqrt{r_n}} \right]$$

- ▶  $X_{n,1}, X_{n,2}, \dots, X_{n,r_n}$  i.i.d., but distn of each  $X_{n,j}$  depends on  $n$ : **Triangular array.**

## How to choose RQMC Allocation $(m_n, r_n)$ with $m_n \times r_n \approx n$ ?

- **Heuristic:** For given budget  $n$ , choose  $r_n$  small and  $m_n \approx n/r_n$  large to exploit QMC.
  - ▶ CI:  $I_{m_n, r_n, \gamma}^{\text{RQ}} \equiv \left[ \hat{\mu}_{m_n, r_n}^{\text{RQ}} \pm z_\gamma \frac{\hat{\sigma}_{m_n, r_n}}{\sqrt{r_n}} \right]$
  - ▶  $r_n = \#$  i.i.d. randomizations
  - ▶  $m_n = \#$  points used from each randomized sequence
- But heuristic lacks rigorous justification.

## How to choose RQMC Allocation $(m_n, r_n)$ with $m_n \times r_n \approx n$ ?

- **Heuristic:** For given budget  $n$ , choose  $r_n$  small and  $m_n \approx n/r_n$  large to exploit QMC.
  - ▶ CI:  $I_{m_n, r_n, \gamma}^{\text{RQ}} \equiv \left[ \hat{\mu}_{m_n, r_n}^{\text{RQ}} \pm z_\gamma \frac{\hat{\sigma}_{m_n, r_n}}{\sqrt{r_n}} \right]$
  - ▶  $r_n = \#$  i.i.d. randomizations
  - ▶  $m_n = \#$  points used from each randomized sequence
- But heuristic lacks rigorous justification.

## Question and Goal

- Is this heuristic theoretically valid?
- **Goal:** Obtain sufficient conditions to ensure CLT and AVCI (as  $n \approx m_n \times r_n \rightarrow \infty$ ).

## How to choose RQMC Allocation $(m_n, r_n)$ with $m_n \times r_n \approx n$ ?

- AVCI relies on **CLT**: **not established for many RQMC settings**.
  - ▶ Nested scrambling of digital nets: CLT as  $m_n = n \rightarrow \infty$ , **fixed**  $r_n = 1$  [Loh 2003]
  - ▶ Randomly shifted lattices: **no** CLT as  $m_n = n/r_n \rightarrow \infty$ , **fixed**  $r_n \geq 1$  [L'Ecuyer et al. 2010]

## How to choose RQMC Allocation $(m_n, r_n)$ with $m_n \times r_n \approx n$ ?

- AVCI relies on **CLT**: **not established for many RQMC settings**.
  - ▶ Nested scrambling of digital nets: CLT as  $m_n = n \rightarrow \infty$ , **fixed**  $r_n = 1$  [Loh 2003]
  - ▶ Randomly shifted lattices: **no** CLT as  $m_n = n/r_n \rightarrow \infty$ , **fixed**  $r_n \geq 1$  [L'Ecuyer et al. 2010]
- **Assumption 1**. On
  - ▶ **Assumption 1.A**:  $m_n r_n \leq n$  for each  $n \geq 1$ , with  $m_n r_n / n \rightarrow 1$  and  $r_n \rightarrow \infty$  as  $n \rightarrow \infty$ .
  - ▶ **Assumption 1.B**: “Simple allocation”:  $(m_n, r_n) = (n^c, n^{1-c})$  for constant  $c \in (0, 1)$ .
  - ▶ **Main Issue**: How to choose  $c$ ?
  - ▶ More general allocation  $(m_n, r_n)$ :  $r_n \rightarrow \infty$  with  $m_n \times r_n \approx n$  as  $n \rightarrow \infty$ .
- **Assumption 2**.  $\sigma_{m_n}^2 \equiv \text{Var}[X_{n,1}] > 0$  for all  $n$  large enough.
- Results based on CLT for triangular arrays. For a triangular array, the  $r_n$  variables within each row  $n$  are independent, but variables in different rows may be dependent. Here:  
 $X_{n,1}, X_{n,2}, \dots, X_{n,r_n}$  are i.i.d., each with some distribution  $F_n$ .



## Theorem 1

If Assumptions 1 and 2 hold, then RQMC estimator  $\hat{\mu}_{m_n, r_n}^{RQ}$  satisfies **CLT**

$$\sqrt{\frac{r_n}{\sigma_{m_n}^2}} [\hat{\mu}_{m_n, r_n}^{RQ} - \mu] \Rightarrow \mathcal{N}(0, 1), \quad \text{as } n \rightarrow \infty$$

under **either**

**Lindeberg condition:**

$$\frac{\mathbb{E} \left[ (X_{n,1} - \mu)^2; |X_{n,1} - \mu| > t \sqrt{r_n \sigma_{m_n}^2} \right]}{\mathbb{E} \left[ (X_{n,1} - \mu)^2 \right]} \rightarrow 0, \quad \text{as } n \rightarrow \infty, \quad \forall t > 0;$$

or

**Lyapounov condition:**

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

- $\sigma_{m_n}^2 = \mathbb{E}[(X_{n,1} - \mu)^2]$  = variance of estimator  $X_{n,1}$  from single randomization of  $m_n$  points.

## RQMC Asymptotically Valid CI (AVCI)

- Recall **Lyapounov condition**:

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

- $\hat{\sigma}_{m_n, r_n}^2 = \frac{1}{r_n-1} \sum_{j=1}^{r_n} (X_{n,j} - \hat{\mu}_{m_n, r_n}^{\text{RQ}})^2$  is unbiased estimator of  $\sigma_{m_n}^2 = \text{Var}[X_{n,1}]$ .
- Approx.  $\gamma$ -level CI for  $\mu$

$$I_{m_n, r_n, \gamma}^{\text{RQ}} = \left[ \hat{\mu}_{m_n, r_n}^{\text{RQ}} \pm z_\gamma \frac{\hat{\sigma}_{m_n, r_n}}{\sqrt{r_n}} \right]$$

### Theorem 2

If Assumptions 1 and 2 hold, along with Lyapounov condition for  $b' = 2$ , then **CLT**

$$\sqrt{\frac{r_n}{\hat{\sigma}_{m_n, r_n}^2}} [\hat{\mu}_{m_n, r_n}^{\text{RQ}} - \mu] \Rightarrow \mathcal{N}(0, 1), \quad \text{as } n \rightarrow \infty$$

and **AVCI**

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

## Ensuring CLT (with $\sigma_m^2 > 0$ )

Instead of condition (1), impose alternative conditions on integrand  $h$

- **Assumption 3.A:**  $V_{\text{HK}}(h) < \infty$  (BVHK)
- **Assumption 3.B:**  $h$  is bounded
- **Assumption 3.C:**  $\mathbb{E}[|h(U) - \mu|^{2+b}] < \infty$  for some  $b > 0$ , where  $U \sim \mathcal{U}[0, 1]^s$ .

### Proposition 1

- Under Assumption 3.A, if  $\exists \lambda \in (0, 1)$   $r_n^{1-\lambda} \left( \frac{m_n \sigma_{m_n}}{(\ln m_n)^s} \right)^2 \rightarrow \infty$  as  $n \rightarrow \infty$ ,
- Under Assumption 3.B, if  $s_n^2 = r_n \sigma_{m_n}^2 \rightarrow \infty$  as  $n \rightarrow \infty$ ,
- Under Assumption 3.C, if  $r_n^{b/(2+b)} \sigma_{m_n}^2 \rightarrow \infty$  as  $n \rightarrow \infty$ ,

then the CLT holds

### Proposition 2

Assumption 3.A  $\implies$  3.B  $\implies$  3.C, leading reversely strict conditions in the above proposition.

## Ensuring AVCI (with $\sigma_m^2 > 0$ )

Instead of condition (1), impose alternative conditions on integrand  $h$

- **Assumption 3.A:**  $V_{\text{HK}}(h) < \infty$  (BVHK)
- **Assumption 3.B:**  $h$  is bounded
- **Assumption 3.C:**  $\mathbb{E}[|h(U) - \mu|^{2+b}] < \infty$  for some  $b > 0$ , where  $U \sim \mathcal{U}[0, 1]^s$ .

Under Assumption 1.A:

### Proposition 3

- Under Assumption 3.A, if  $r_n \left( \frac{m_n \sigma_{m_n}}{(\ln m_n)^s} \right)^4 \rightarrow \infty$  as  $n \rightarrow \infty$ ,
- Under Assumption 3.C, if  $\frac{\mathbb{E}[|X_{m,1} - \mu|^{2+b'}]}{\sigma_m^{2+b'}} \leq k_1$  with  $b' = 2$ ,

then AVCI holds

## Corollaries Ensuring CLT or AVCI under Assumption 1.B

### Assumption

For estimator  $X_{n,1}$  from **single** randomization of  $m_n$  points,

$$\sigma_{m_n} \equiv \sqrt{\text{Var}[X_{n,1}]} \approx \Theta(m_n^{-\alpha_*}) \quad \text{as } m_n \rightarrow \infty, \quad \text{where } \alpha_* \equiv - \lim_{m_n \rightarrow \infty} \frac{\ln(\sigma_{m_n})}{\ln(m_n)} > \frac{1}{2}$$

- $\alpha_* \geq 1$  when  $V_{\text{HK}}(h) < \infty$  (BVHK).
- Under Assumption 1.B [  $(m_n, r_n) = (n^c, n^{1-c}), c \in (0, 1)$  ],

$$\text{RMSE} [\hat{\mu}_{m_n, r_n}^{\text{RQ}}] = \frac{\sigma_{m_n}}{\sqrt{r_n}} \approx \Theta(n^{-v(\alpha_*, c)}) \quad \text{as } n \rightarrow \infty, \quad \text{with } v(\alpha_*, c) \equiv c \left[ \alpha_* - \frac{1}{2} \right] + \frac{1}{2}.$$

## Corollaries Ensuring CLT or AVCI under Assumption 1.B

### Assumption

For estimator  $X_{n,1}$  from **single** randomization of  $m_n$  points,

$$\sigma_{m_n} \equiv \sqrt{\text{Var}[X_{n,1}]} \approx \Theta(m_n^{-\alpha_*}) \quad \text{as } m_n \rightarrow \infty, \quad \text{where } \alpha_* \equiv -\lim_{m_n \rightarrow \infty} \frac{\ln(\sigma_{m_n})}{\ln(m_n)} > \frac{1}{2}$$

- $\alpha_* \geq 1$  when  $V_{\text{HK}}(h) < \infty$  (BVHK).
- Under Assumption 1.B [  $(m_n, r_n) = (n^c, n^{1-c}), c \in (0, 1)$  ],

$$\text{RMSE} [\hat{\mu}_{m_n, r_n}^{\text{RQ}}] = \frac{\sigma_{m_n}}{\sqrt{r_n}} \approx \Theta(n^{-\nu(\alpha_*, c)}) \quad \text{as } n \rightarrow \infty, \quad \text{with } \nu(\alpha_*, c) \equiv c \left[ \alpha_* - \frac{1}{2} \right] + \frac{1}{2}.$$

- Corollary  $k = 1, 2, \dots, 6$ : ensure CLT or AVCI under constraint  $c < c_k(\alpha_*)$ 
  - ▶  $c_k(\alpha_*) \in (0, 1]$ , sometimes  $c_k(\alpha_*) = 1$ .
  - ▶ Optimal RMSE: take  $c < c_k(\alpha_*)$  with  $c \approx c_k(\alpha_*)$

$$\text{RMSE} [\hat{\mu}_{m_n, r_n}^{\text{RQ}}] \approx \Theta(n^{-\nu_k(\alpha_*)}) \quad \text{as } n \rightarrow \infty, \quad \text{with } \nu_k(\alpha_*, c) \equiv c_k(\alpha_*) \left[ \alpha_* - \frac{1}{2} \right] + \frac{1}{2} > \frac{1}{2}$$

$\implies$  **RQMC better than MC.**

## Corollaries Ensuring CLT or AVCI

### Corollary 1

Suppose that Assumptions 1 and 2 hold, and  $\exists b' > 0$  and  $k_1 \in (0, \infty)$  such that

$$\frac{\mathbb{E}\left[|X_{n,1}-\mu|^{2+b'}\right]}{\sigma_{m_n}^{2+b'}} \leq k_1 \quad \forall m_n \text{ sufficiently large.} \quad (1)$$

Then **CLT** holds for allocation  $(m_n, r_n) = (n^c, n^{1-c})$  with any

$$c < 1 \equiv c_1(\alpha_*),$$

and optimal  $RMSE \approx \Theta(n^{-v_1(\alpha_*)})$  as  $n \rightarrow \infty$  with

$$v_1(\alpha_*) \equiv \alpha_*.$$

If (1) holds for  $b' = 2$ , then **AVCI** holds for  $c < c_1(\alpha_*)$ , and RMSE rate exponent is  $v_1(\alpha_*)$ .

- CLT/AVCI with  $r_n \rightarrow \infty$  since  $c < 1$ .
- RMSE rate same as  $\sigma_{m_n} \approx \Theta(m_n^{-\alpha_*})$  for **single** randomization of  $m_n = n$  points.

## Corollaries Ensuring CLT or AVCI: Tradeoffs

Impose again alternative conditions on integrand  $h$

- **Assumption 3.A:**  $V_{\text{HK}}(h) < \infty$  (BVHK)
- **Assumption 3.B:**  $h$  is bounded
- **Assumption 3.C:**  $\mathbb{E}[|h(U) - \mu|^{2+b}] < \infty$  for some  $b > 0$ , where  $U \sim \mathcal{U}[0, 1]^s$ .

### Proposition 4

- *Assumption 3.A  $\implies$  3.B  $\implies$  3.C,  
leading to successively smaller  $c_k(\alpha_*)$  for Corollaries  $k$  (next slide).*
- *Under Assumption 3.x, for  $c_k(\alpha_*)$  ensuring CLT and  $c_{k'}(\alpha_*)$  ensuring AVCI,*

$$c_k(\alpha_*) \geq c_{k'}(\alpha_*) \quad (\text{often } >).$$

- **Assumption 1:**  $(m_n, r_n) = (n^c, n^{1-c})$ ,  $c \in (0, 1)$
- Corollary  $k$ :  $c < c_k(\alpha_*)$
- $\sigma_{m_n} \approx \Theta(m_n^{-\alpha_*})$ ,  $\alpha_* > 1/2$

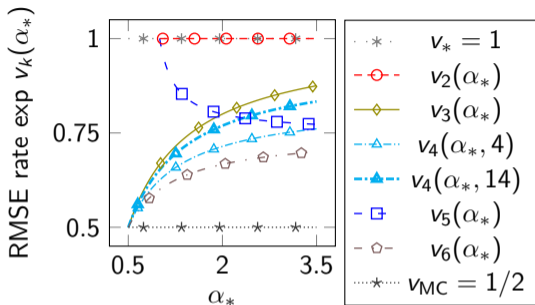
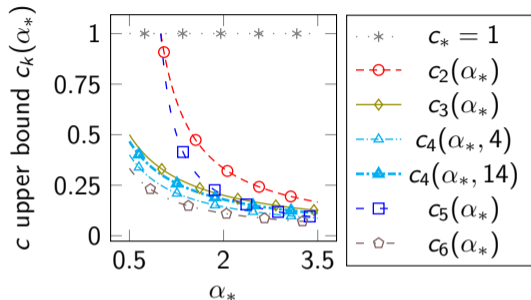


## Corollaries CLT or AVCI: Tradeoffs

Cor. $k$	Ensures	Assumption on $h$	$c$ upper bd $c_k(\alpha_*)$	RMSE rate exp $v_k(\alpha_*)$
2	CLT	3.A (BVHK)	$\frac{1}{2\alpha_* - 1} >$	$1 >$
3	CLT	3.B ( $h$ bdd)	$\frac{1}{2\alpha_* + 1} >$	$\frac{2\alpha_*}{2\alpha_* + 1} >$
4	CLT	3.C ( $b > 0$ )	$\frac{1}{2\alpha_*(1+\frac{2}{b})+1} \in (0, \frac{1}{2})$	$\frac{2\alpha_*(1+\frac{1}{b})}{2\alpha_*(1+\frac{2}{b})+1} > \frac{1}{2}$
5	AVCI	3.A (BVHK)	$\frac{1}{4\alpha_* - 3} >$	$\frac{3\alpha_* - 2}{4\alpha_* - 3} >$
6	AVCI	3.C ( $b = 2$ )	$\frac{1}{4\alpha_* + 1} \in (0, \frac{1}{3})$	$\frac{3\alpha_*}{4\alpha_* + 1} > \frac{1}{2}$

- 3.A  $\implies$  3.B  $\implies$  3.C
  - ▶ **Assumption 3.A:**  $V_{\text{HK}}(h) < \infty$  (BVHK:  $\implies \alpha_* \geq 1$ )
  - ▶ **Assumption 3.B:**  $h$  is bounded.
  - ▶ **Assumption 3.C:**  $\mathbb{E}[|h(U) - \mu|^{2+b}] < \infty$  for some  $b > 0$ , where  $U \sim \mathcal{U}[0, 1]^s$ .
- Comparisons for fixed  $\alpha_* > 1/2$ 
  - ▶  $(m_n, r_n) = (n^c, n^{1-c})$ ,  $c < c_k(\alpha_*)$ ,  $\text{opt RMSE} \approx \Theta(n^{-v_k(\alpha_*)})$ .

## Conditions Ensuring CLT or AVCI: Tradeoffs



- All  $c_k(\alpha_*) \downarrow$  as  $\alpha_* \uparrow$ 
  - ▶ Corollary  $k$ :  $c < c_k(\alpha_*)$  in  $(m_n, r_n) = (n^c, n^{1-c})$ .
  - ▶  $\sigma_{m_n} \approx \Theta(m_n^{-\alpha_*})$ ,  $\alpha_* > 1/2$  ( $\geq 1$  BVHK)
- Most  $v_k(\alpha_*) \uparrow$  as  $\alpha_* \uparrow$ 
  - ▶ Optimal RMSE  $\approx \Theta(n^{-v_k(\alpha_*)})$ ,  $n \rightarrow \infty$
  - ▶ Larger  $\alpha_*$  usually yields better RQMC performance.

## Concluding Remarks

- Sufficient conditions for RQMC CLT and AVCI for  $\mu = \mathbb{E}[h(U)]$ 
  - ▶ Corollary  $k$ :  $c < c_k(\alpha_*)$  for allocation  $(m_n, r_n) = (n^c, n^{1-c})$ .
  - ▶ Tradeoffs: Weaker conditions on integrand  $h$   
→ stronger restrictions on  $c$ .
  - ▶ AVCI imposes (often strictly) stronger restrictions on  $c$  than CLT.
  - ▶ RMSE: RQMC outperforms MC.
- Current work
  - ▶ Estimate  $c_k(\alpha_*)$ .
  - ▶ Weaker conditions.
  - ▶ CLT and AVCI for biased estimators, e.g., quantiles.
  - ▶ Alternative CIs: Bootstrap- $t$ , etc.

**Thank you!**