

# The Array-RQMC method: Review of convergence results

**Pierre L'Ecuyer, Christian Lécot, Bruno Tuffin**

DIRO, Université de Montréal, Canada

LAMA, Université de Savoie, France

Inria–Rennes, France

# Monte Carlo for Markov Chains

**Setting:** A Markov chain with state space  $\mathcal{X} \subseteq \mathbb{R}^{\ell}$ , evolves as

$$X_0 = x_0, \quad X_j = \varphi_j(X_{j-1}, \mathbf{U}_j), \quad j \geq 1,$$

where the  $\mathbf{U}_j$  are i.i.d. uniform r.v.'s over  $(0, 1)^d$ . Want to estimate

$$\mu = \mathbb{E}[Y] \quad \text{where} \quad Y = \sum_{j=1}^{\tau} g_j(X_j)$$

for some fixed time horizon  $\tau$ .

# Monte Carlo for Markov Chains

**Setting:** A Markov chain with state space  $\mathcal{X} \subseteq \mathbb{R}^{\ell}$ , evolves as

$$X_0 = x_0, \quad X_j = \varphi_j(X_{j-1}, \mathbf{U}_j), \quad j \geq 1,$$

where the  $\mathbf{U}_j$  are i.i.d. uniform r.v.'s over  $(0, 1)^d$ . Want to estimate

$$\mu = \mathbb{E}[Y] \quad \text{where} \quad Y = \sum_{j=1}^{\tau} g_j(X_j)$$

for some fixed time horizon  $\tau$ .

**Ordinary MC:** For  $i = 0, \dots, n-1$ , generate  $X_{i,j} = \varphi_j(X_{i,j-1}, \mathbf{U}_{i,j})$ ,  $j = 1, \dots, \tau$ , where the  $\mathbf{U}_{i,j}$ 's are i.i.d.  $U(0, 1)^d$ . Estimate  $\mu$  by

$$\hat{\mu}_n = \frac{1}{n} \sum_{i=1}^n \sum_{j=1}^{\tau} g_j(X_{i,j}) = \frac{1}{n} \sum_{i=1}^n Y_i.$$

## Example: Asian Call Option

Given observation times  $t_1, t_2, \dots, t_\tau$  suppose

$$S(t_j) = S(t_{j-1}) \exp[(r - \sigma^2/2)(t_j - t_{j-1}) + \sigma(t_j - t_{j-1})^{1/2} \Phi^{-1}(U_j)],$$

where  $U_j \sim U[0, 1)$  and  $S(t_0) = s_0$  is fixed.

Running average:  $\bar{S}_j = \frac{1}{j} \sum_{i=1}^j S(t_i)$ .

State:  $X_j = (S(t_j), \bar{S}_j)$ .

Transition:

$$X_j = (S(t_j), \bar{S}_j) = \varphi_j(S(t_{j-1}), \bar{S}_{j-1}, U_j) = \left( S(t_j), \frac{(j-1)\bar{S}_{j-1} + S(t_j)}{j} \right).$$

Payoff at step  $j = \tau$  is  $Y = g_\tau(X_\tau) = \max[0, \bar{S}_\tau - K]$ .

## Plenty of other applications:

Finance

Queueing systems

Inventory, distribution, logistic systems

Reliability models

MCMC in Bayesian statistics

Etc.

# Classical RQMC for Markov Chains

Put  $\mathbf{V}_i = (\mathbf{U}_{i,1}, \dots, \mathbf{U}_{i,\tau}) \in (0, 1)^s = (0, 1)^{d\tau}$ . Estimate  $\mu$  by

$$\hat{\mu}_{\text{rqmc},n} = \frac{1}{n} \sum_{i=1}^n \sum_{j=1}^{\tau} g_j(X_{i,j})$$

where  $P_n = \{\mathbf{V}_0, \dots, \mathbf{V}_{n-1}\} \subset (0, 1)^s$  satisfies:

- (a) each point  $\mathbf{V}_i$  has the **uniform distribution** over  $(0, 1)^s$ ;
- (b)  $P_n$  covers  $(0, 1)^s$  very evenly (i.e., has low discrepancy).

The dimension  $s$  is often very large!

## Array-RQMC for Markov Chains

L., Lécot, Tuffin, et al. [2004, 2006, 2008, etc.]

Simulate an “array” of  $n$  chains in “parallel.”

At each step, use an RQMC point set  $P_n$  to advance all the chains by one step, while inducing global **negative dependence** across the chains.

**Goal:** Want a **small discrepancy** (or “distance”) between the empirical distribution of  $S_{n,j} = \{X_{0,j}, \dots, X_{n-1,j}\}$  and the theoretical distribution of  $X_j$ , for each  $j$ .

If we succeed, these (unbiased) estimators will have small variance:

$$\mu_j = \mathbb{E}[g_j(X_j)] \approx \frac{1}{n} \sum_{i=0}^{n-1} g_j(X_{i,j}) \quad \text{and} \quad \mu = \mathbb{E}[Y] \approx \frac{1}{n} \sum_{i=0}^{n-1} Y_i.$$

## Array-RQMC for Markov Chains

L., Lécot, Tuffin, et al. [2004, 2006, 2008, etc.]

Simulate an “array” of  $n$  chains in “parallel.”

At each step, use an RQMC point set  $P_n$  to advance all the chains by one step, while inducing global **negative dependence** across the chains.

**Goal:** Want a **small discrepancy** (or “distance”) between the empirical distribution of  $S_{n,j} = \{X_{0,j}, \dots, X_{n-1,j}\}$  and the theoretical distribution of  $X_j$ , for each  $j$ .

If we succeed, these (unbiased) estimators will have small variance:

$$\mu_j = \mathbb{E}[g_j(X_j)] \approx \frac{1}{n} \sum_{i=0}^{n-1} g_j(X_{i,j}) \quad \text{and} \quad \mu = \mathbb{E}[Y] \approx \frac{1}{n} \sum_{i=0}^{n-1} Y_i.$$

How can we preserve low-discrepancy of  $S_{n,j}$  as  $j$  increases?

Can we quantify the variance improvement? Convergence rate in  $n$ ?

## Some generalizations

L., Lécot, and Tuffin [2008]:  $\tau$  can be a random stopping time w.r.t. the filtration  $\mathcal{F}\{(j, X_j), j \geq 0\}$ .

L., Demers, and Tuffin [2006, 2007]: Combination with splitting techniques (multilevel and without levels), combination with importance sampling and weight windows. Covers particle filters.

L. and Sanvido [2010]: Combination with coupling from the past for exact sampling.

Dion and L. [2010]: Combination with approximate dynamic programming and for optimal stopping problems.

Gerber and Chopin [2014]: Sequential QMC (yesterday's talk).

# Convergence results and applications

L., Lécot, and Tuffin [2006, 2008]: Special cases: convergence at MC rate, one-dimensional, stratification, etc.

Lécot and Tuffin [2004]: Deterministic, one-dimension, discrete state.

El Haddad, Lécot, L. [2008, 2010]: Deterministic, multidimensional.

Fakherredine, El Haddad, Lécot [2012, 2013, 2014]: LHS, stratification, Sudoku sampling, ...

Wächter and Keller [2008]: Applications in computer graphics.

## Other QMC methods for Markov chains

Interested in **steady-state distribution**. Introduce dependence between the steps  $j$ ; a **single chain visit the state space very uniformly**.

Owen, Tribble, Chen, Dick, Matsumoto, Nishimura, .... [2004–2010]:  
Markov chain quasi-Monte Carlo.

Propp [2012] and earlier: Rotor-router sampling.

To simplify, suppose each  $X_j$  is a uniform r.v. over  $(0, 1)^\ell$ .

Select a discrepancy measure  $D$  for the point set  $S_{n,j} = \{X_{0,j}, \dots, X_{n-1,j}\}$  over  $(0, 1)^\ell$ , and a corresponding measure of variation  $V$ , such that

$$\text{Var}[\hat{\mu}_{\text{rqmc},j,n}] = \mathbb{E}[(\hat{\mu}_{\text{rqmc},j,n} - \mu_j)^2] \leq \mathbb{E}[D^2(S_{n,j})] V^2(g_j).$$

To simplify, suppose each  $X_j$  is a uniform r.v. over  $(0, 1)^\ell$ .

Select a discrepancy measure  $D$  for the point set  $S_{n,j} = \{X_{0,j}, \dots, X_{n-1,j}\}$  over  $(0, 1)^\ell$ , and a corresponding measure of variation  $V$ , such that

$$\text{Var}[\hat{\mu}_{\text{rqmc},j,n}] = \mathbb{E}[(\hat{\mu}_{\text{rqmc},j,n} - \mu_j)^2] \leq \mathbb{E}[D^2(S_{n,j})] V^2(g_j).$$

If  $D$  is defined via a reproducing kernel Hilbert space, then, for some random  $\xi_j$  (that generally depends on  $S_{n,j}$ ),

$$\begin{aligned} \mathbb{E}[D^2(S_{n,j})] &= \text{Var} \left[ \frac{1}{n} \sum_{i=1}^n \xi_j(X_{i,j}) \right] = \text{Var} \left[ \frac{1}{n} \sum_{i=1}^n (\xi_j \circ \varphi_j)(X_{i,j-1}, \mathbf{U}_{i,j}) \right] \\ &\leq \mathbb{E}[D_{(2)}^2(Q_n)] \cdot V_{(2)}^2(\xi_j \circ \varphi_j) \end{aligned}$$

for some other discrepancy  $D_{(2)}$  over  $(0, 1)^{\ell+d}$ , where

$$Q_n = \{(X_{0,j-1}, \mathbf{U}_{0,j}), \dots, (X_{n-1,j-1}, \mathbf{U}_{n-1,j})\}.$$

**Goal:** Under appropriate conditions, to obtain  $V_{(2)}(\xi_j \circ \varphi_j) < \infty$  and  $\mathbb{E}[D_{(2)}^2(Q_n)] = O(n^{-\alpha+\epsilon})$  for some  $\alpha \geq 1$ .

## Discrepancy bounds by induction?

Let  $\ell = d = 1$ ,  $\mathcal{X} = [0, 1]$ , and  $X_j \sim U(0, 1)$ .  $L_2$ -star discrepancy:

$$D^2(x_0, \dots, x_{n-1}) = \frac{1}{12n^2} + \frac{1}{n} \sum_{i=0}^{n-1} (w_i - x_i)^2$$

where  $w_i = (i + 1/2)/n$  and  $0 \leq x_0 \leq x_1 \leq \dots \leq x_{n-1}$ . We have

$$\xi_j(x) = -\frac{1}{n} \sum_{i=1}^{n-1} [\mu(Y_i) + B_2((x - Y_i) \bmod 1) + B_1(x)B_1(Y_i)],$$

where  $B_1(x) = x - 1/2$  and  $B_2(x) = x^2 - x + 1/6$ .

**Problem:** the 2-dim function  $\xi_j \circ \varphi_j$  has mixed derivative that is not square integrable, so it has infinite variation, it seems. Otherwise, we would have a proof that  $\mathbb{E}[D^2(S_{n,j})] = \mathcal{O}(n^{-2})$ . **Help!**

In the points  $(X_{i,j-1}, \mathbf{U}_{i,j})$  of  $Q_n$ , the  $\mathbf{U}_{i,j}$  can be defined via some RQMC scheme, but the  $X_{i,j-1}$  cannot be chosen; they are determined by the history of the chains.

The idea is to select a low-discrepancy point set

$$\tilde{Q}_n = \{(\mathbf{w}_0, \mathbf{U}_0), \dots, (\mathbf{w}_{n-1}, \mathbf{U}_{n-1})\},$$

where the  $\mathbf{w}_i \in [0, 1)^\ell$  are fixed and the  $\mathbf{U}_i \in (0, 1)^d$  are randomized, and then define a bijection between the states  $X_{i,j-1}$  and the  $\mathbf{w}_i$  so that the  $X_{i,j-1}$  are “close” to the  $\mathbf{w}_i$  (small discrepancy between the two sets).

**Example:** If  $\ell = 1$ , can take  $\mathbf{w}_i = (i + 0.5)/n$ .

Bijection defined by a permutation  $\pi_j$  of  $S_{n,j}$ .

In the points  $(X_{i,j-1}, \mathbf{U}_{i,j})$  of  $Q_n$ , the  $\mathbf{U}_{i,j}$  can be defined via some RQMC scheme, but the  $X_{i,j-1}$  cannot be chosen; they are determined by the history of the chains.

The idea is to select a low-discrepancy point set

$$\tilde{Q}_n = \{(\mathbf{w}_0, \mathbf{U}_0), \dots, (\mathbf{w}_{n-1}, \mathbf{U}_{n-1})\},$$

where the  $\mathbf{w}_i \in [0, 1)^\ell$  are fixed and the  $\mathbf{U}_i \in (0, 1)^d$  are randomized, and then define a bijection between the states  $X_{i,j-1}$  and the  $\mathbf{w}_i$  so that the  $X_{i,j-1}$  are “close” to the  $\mathbf{w}_i$  (small discrepancy between the two sets).

**Example:** If  $\ell = 1$ , can take  $\mathbf{w}_i = (i + 0.5)/n$ .

Bijection defined by a permutation  $\pi_j$  of  $S_{n,j}$ .

For state space in  $\mathbb{R}^\ell$ : same algorithm essentially.

## Array-RQMC algorithm

$X_{i,0} \leftarrow x_0$ , for  $i = 0, \dots, n-1$ ;

**for**  $j = 1, 2, \dots, \tau$  **do**

Randomize afresh  $\{\mathbf{U}_{0,j}, \dots, \mathbf{U}_{n-1,j}\}$  in  $\tilde{Q}_n$ ;

$X_{i,j} = \varphi_j(X_{\pi_j(i),j-1}, \mathbf{U}_{i,j})$ , for  $i = 0, \dots, n-1$ ;

Compute the permutation  $\pi_{j+1}$  (sort the states);

**end for**

Estimate  $\mu$  by the average  $\bar{Y}_n = \hat{\mu}_{\text{rqmc},n}$ .

## Array-RQMC algorithm

$X_{i,0} \leftarrow x_0$ , for  $i = 0, \dots, n-1$ ;

**for**  $j = 1, 2, \dots, \tau$  **do**

Randomize afresh  $\{\mathbf{U}_{0,j}, \dots, \mathbf{U}_{n-1,j}\}$  in  $\tilde{Q}_n$ ;

$X_{i,j} = \varphi_j(X_{\pi_j(i),j-1}, \mathbf{U}_{i,j})$ , for  $i = 0, \dots, n-1$ ;

Compute the permutation  $\pi_{j+1}$  (sort the states);

**end for**

Estimate  $\mu$  by the average  $\bar{Y}_n = \hat{\mu}_{\text{rqmc},n}$ .

**Theorem:** The average  $\bar{Y}_n$  is an unbiased estimator of  $\mu$ .

Can estimate  $\text{Var}[\bar{Y}_n]$  by the empirical variance of  $m$  indep. realizations.

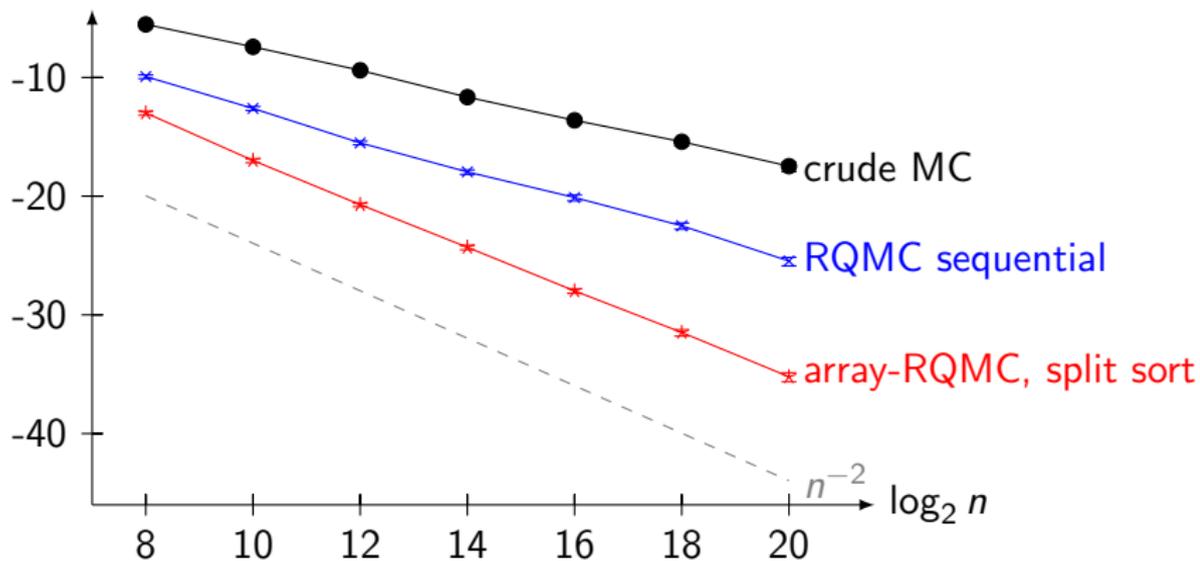
## Example: Asian Call Option

$S(0) = 100$ ,  $K = 100$ ,  $r = 0.05$ ,  $\sigma = 0.15$ ,  $t_j = j/52$ ,  $j = 0, \dots, \tau = 13$ .

RQMC points: Sobol' nets with a linear scrambling + random digital shift, for all the results reported here.

Similar results for randomly-shifted lattice + baker's transform.

$\log_2 \text{Var}[\hat{\mu}_{\text{RQMC},n}]$



## Mapping chains to points

### One possibility: Multivariate sort:

Sort the states (chains) by first coordinate, in  $n_1$  packets of size  $n/n_1$ .

Sort each packet by second coordinate, in  $n_2$  packets of size  $n/n_1 n_2$ .

...

At the last level, sort each packet of size  $n_\ell$  by the last coordinate.

Choice of  $n_1, n_2, \dots, n_\ell$ ?

## Mapping chains to points

### One possibility: Multivariate sort:

Sort the states (chains) by first coordinate, in  $n_1$  packets of size  $n/n_1$ .

Sort each packet by second coordinate, in  $n_2$  packets of size  $n/n_1 n_2$ .

...

At the last level, sort each packet of size  $n_\ell$  by the last coordinate.

### Choice of $n_1, n_2, \dots, n_\ell$ ?

**For large  $\ell$ :** Define a transformation  $v : \mathcal{X} \rightarrow [0, 1)^c$  and do a multivariate sort (in  $c < \ell$  dimensions) of the points  $v(X_{i,j})$ .

Choice of  $v$ : states mapped to nearby values of  $v$  should be nearly equivalent.

## Mapping chains to points

**One possibility: Multivariate sort:**

Sort the states (chains) by first coordinate, in  $n_1$  packets of size  $n/n_1$ .

Sort each packet by second coordinate, in  $n_2$  packets of size  $n/n_1 n_2$ .

...

At the last level, sort each packet of size  $n_\ell$  by the last coordinate.

Choice of  $n_1, n_2, \dots, n_\ell$ ?

**For large  $\ell$ :** Define a transformation  $v : \mathcal{X} \rightarrow [0, 1)^c$  and do a multivariate sort (in  $c < \ell$  dimensions) of the points  $v(X_{i,j})$ .

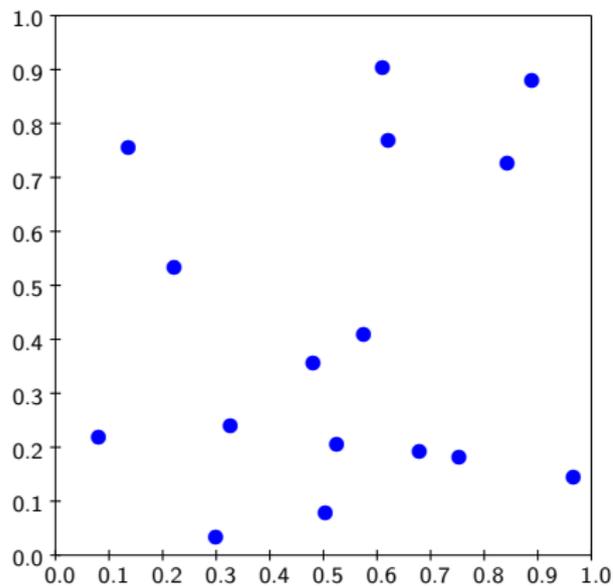
Choice of  $v$ : states mapped to nearby values of  $v$  should be nearly equivalent.

**For  $c = 1$ ,**  $\mathcal{X}$  is mapped to  $[0, 1)$ , which leads to a one-dim sort.

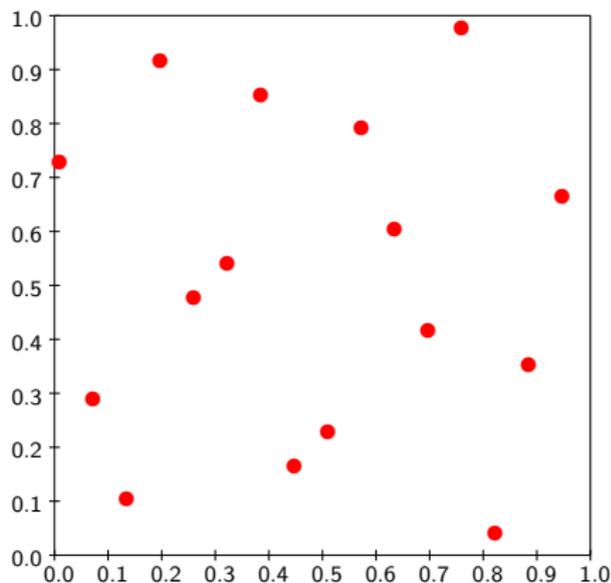
The mapping  $v$  can be based on a space-filling curve: Z-curve, Hilbert curve, etc. See Wächter and Keller [2008], Gerber and Chopin [2014].

# A (4,4) mapping

States of the chains

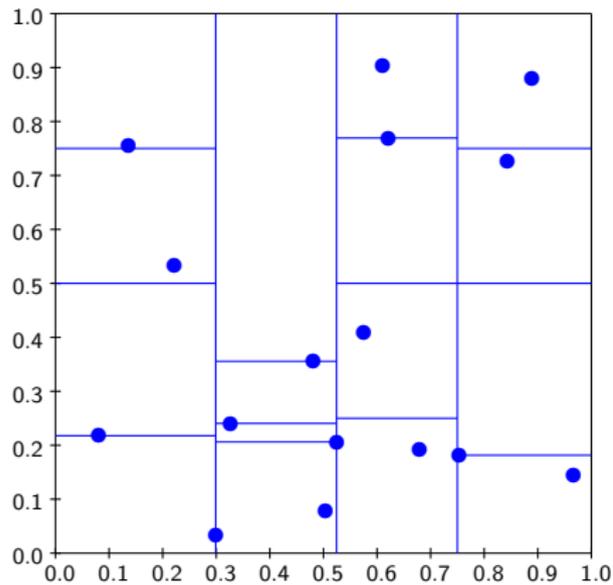


Sobol' net in 2 dimensions with digital shift

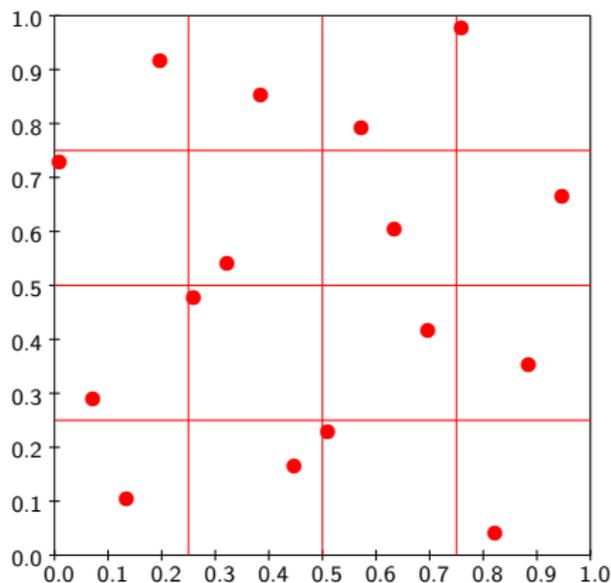


# A (4,4) mapping

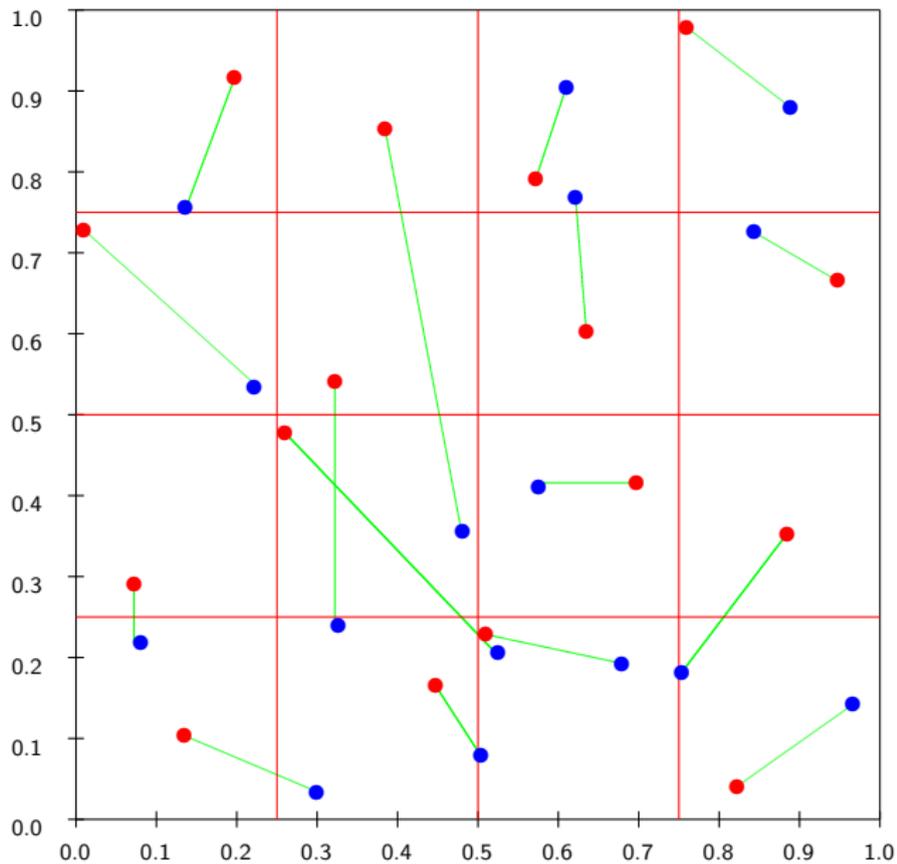
States of the chains



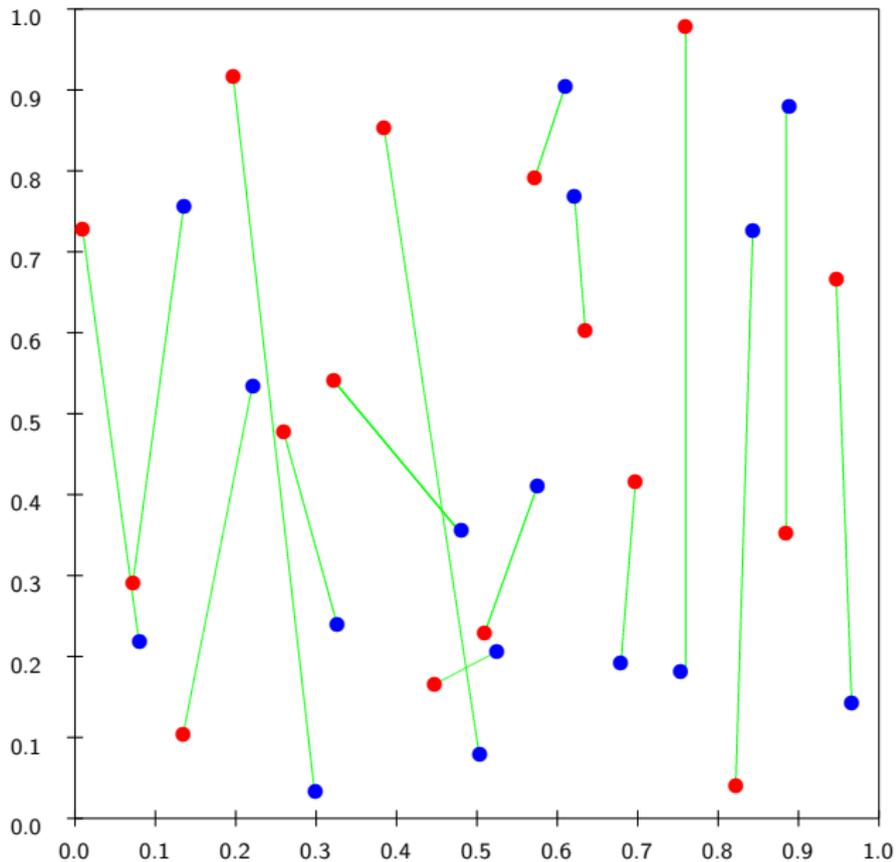
Sobol' net in 2 dimensions with digital shift



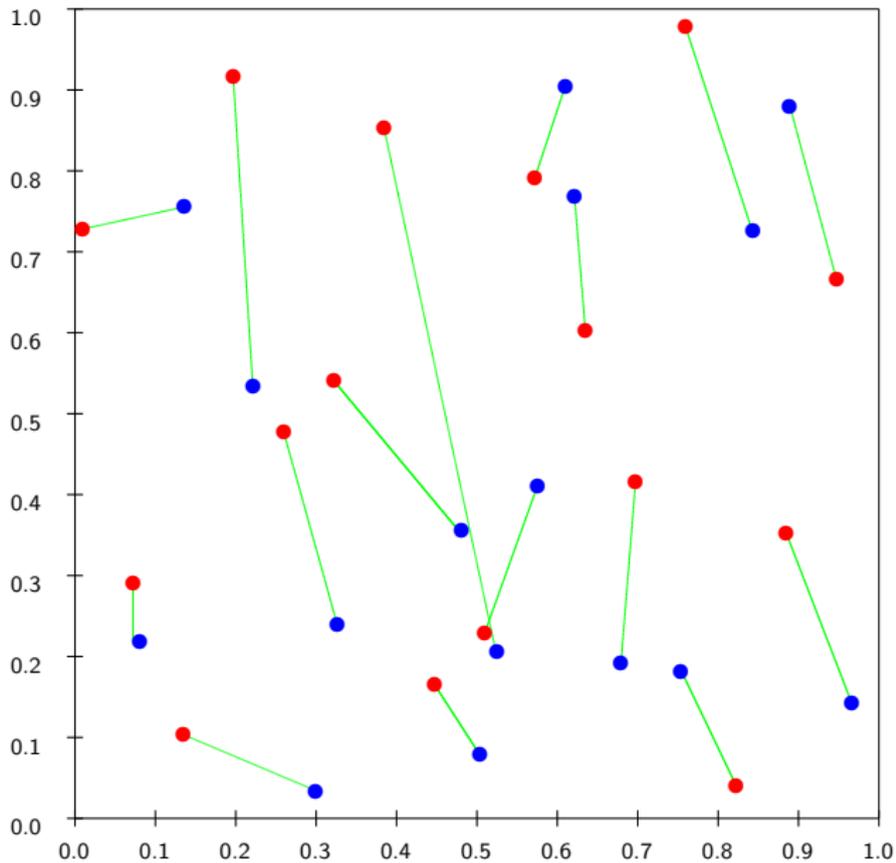
# A (4,4) mapping



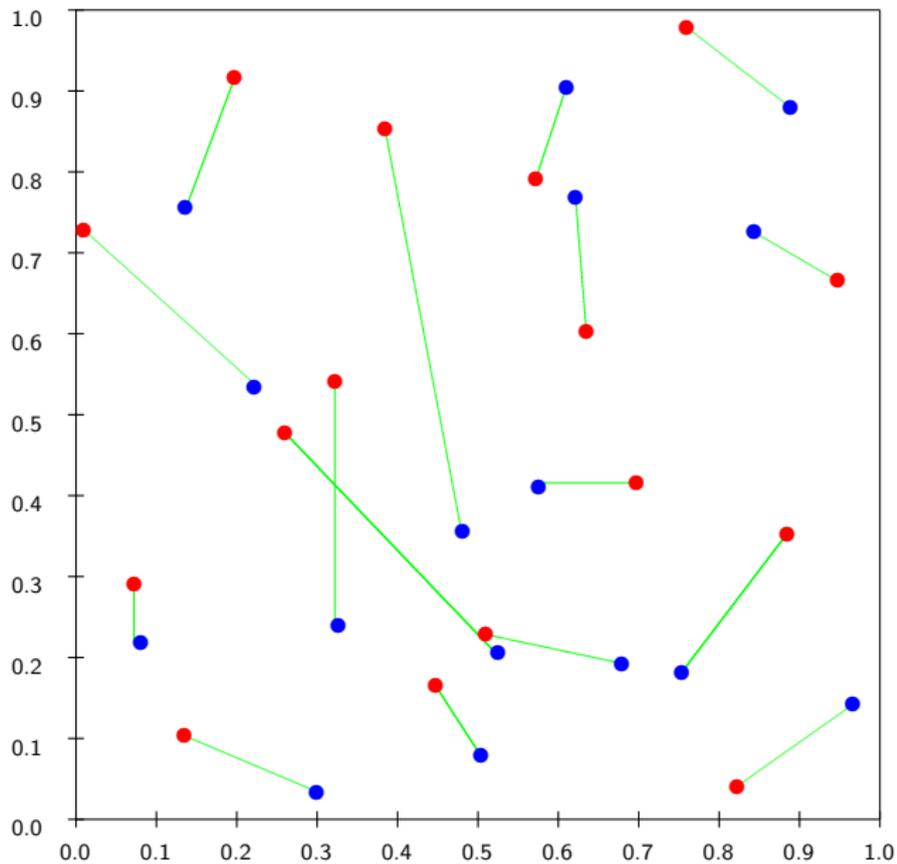
# A (16,1) mapping, sorting along first coordinate



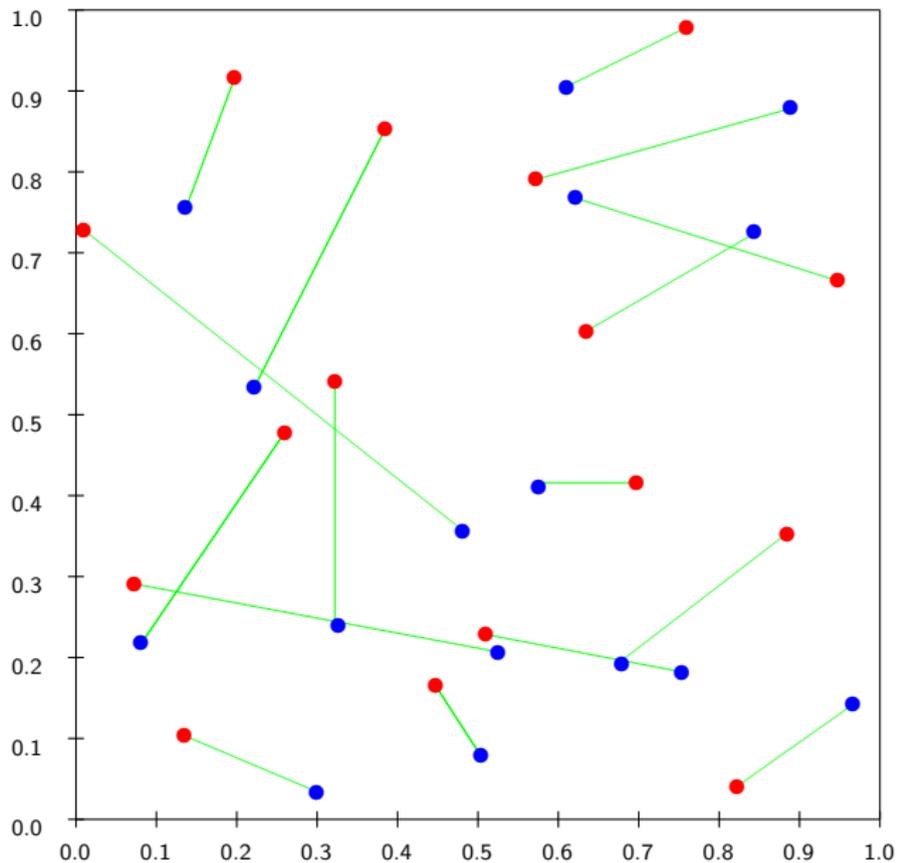
# A (8,2) mapping



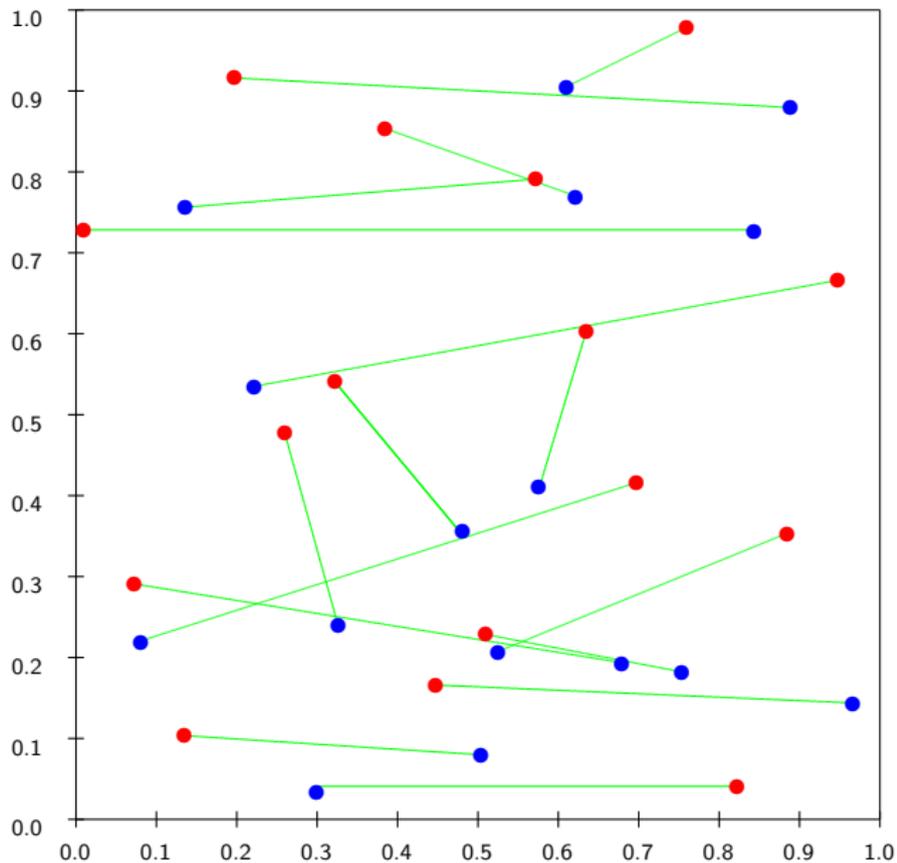
# A (4,4) mapping



# A (2,8) mapping



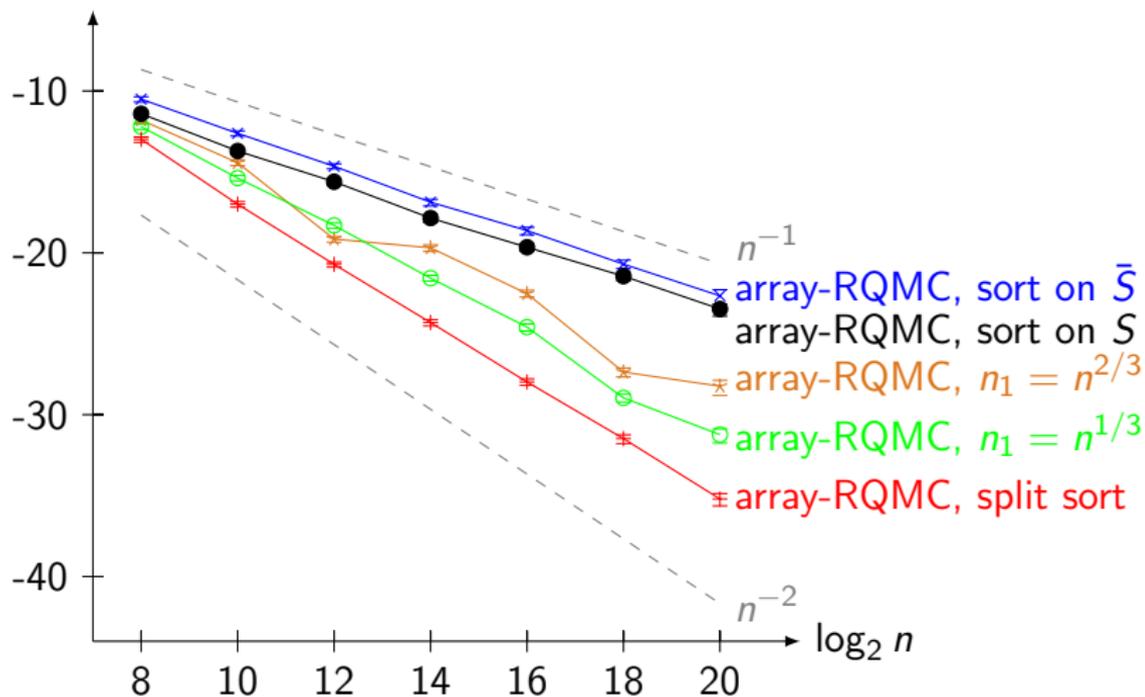
# A (1,16) mapping, sorting along second coordinate



## Sorting strategies for array-RQMC.

State-point **mapping** via two-dimensional sort: sort in  $n_1$  packets based on  $S(t_j)$ , then sort the packets based on  $\bar{S}_j$ . Split sort:  $n_1 = n_2$ .

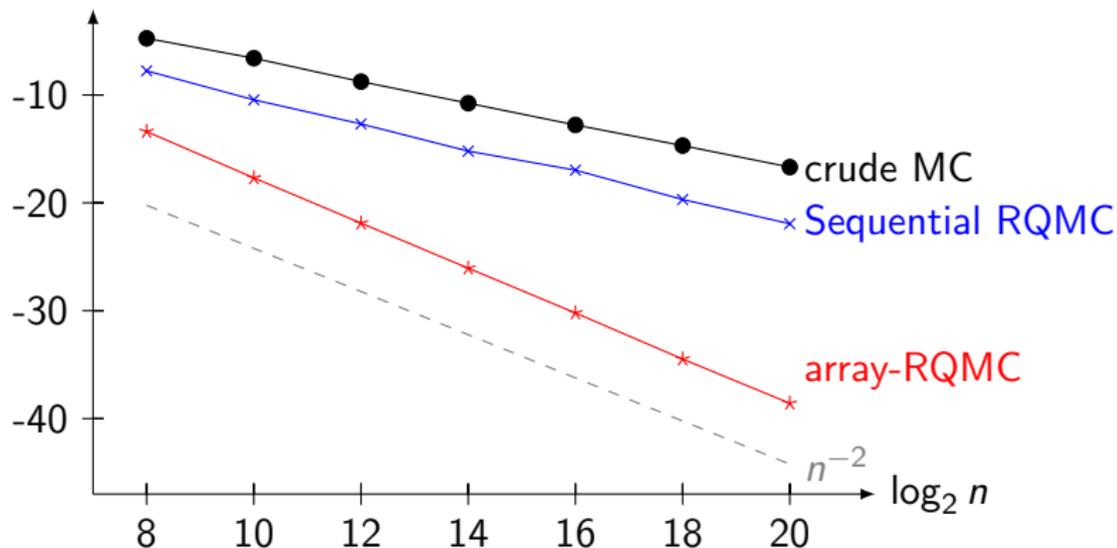
$$\log_2 \text{Var}[\hat{\mu}_{\text{RQMC},n}]$$



## Artificial one-dim example: a simple put option

GBM  $\{S(t), t \geq 0\}$  with drift  $\mu = 0.05$ , volatility  $\sigma = 0.08$ ,  $S(0) = 100$ .  
 Generate  $X_j = S(t_j)$  for  $t_j = j/16, j = 1, \dots, \tau = 16$ , sequentially.  
 Payoff at  $t_{16} = 1$ :  $Y = g_\tau(S(1)) = e^{-0.05} \max(0, 101 - S(1))$ .

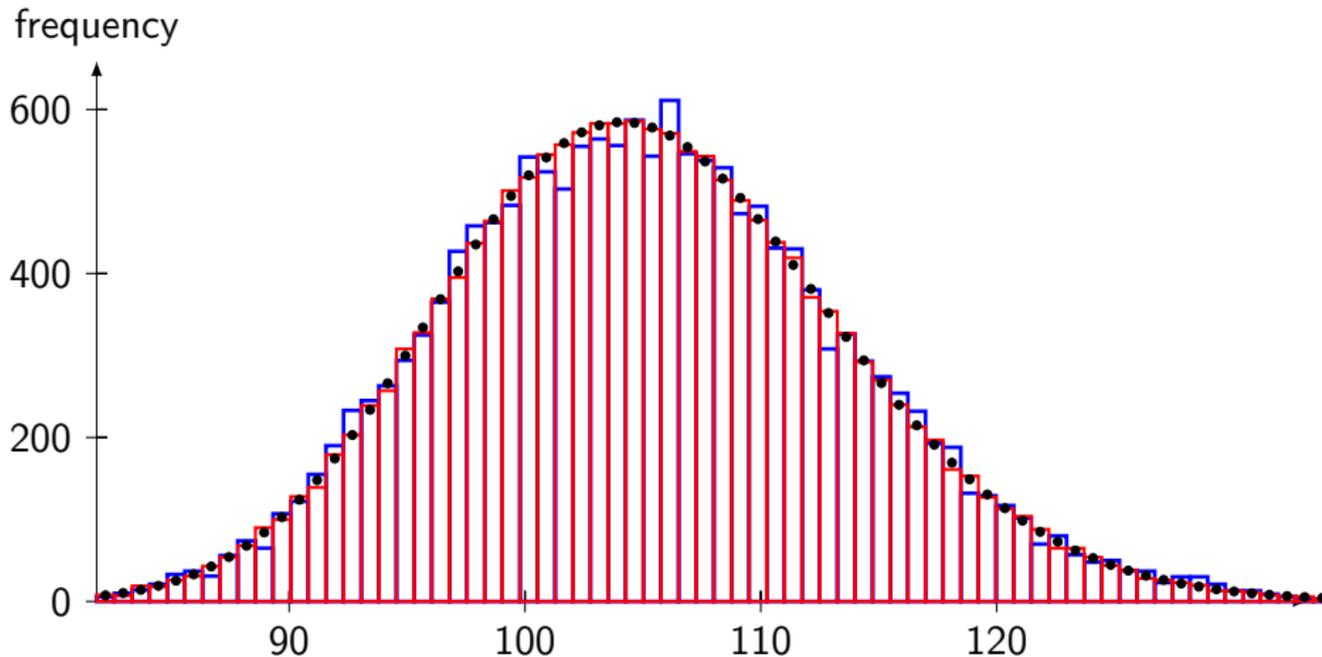
$\log_2 \text{Var}[\hat{\mu}_{\text{RQMC},n}]$



## Histogram of states at step 16

States for **array-RQMC** with  $n = 2^{14}$  in **red** and for **MC** in **blue**.

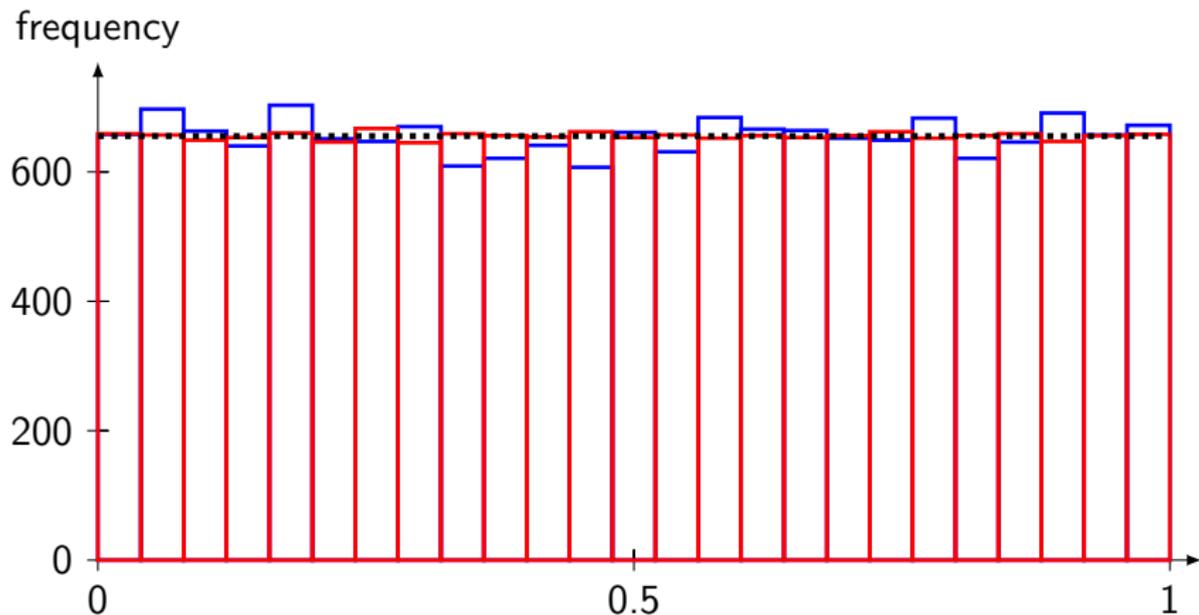
Theoretical dist.: black dots.



## Histogram after transformation to uniforms (applying the cdf).

States for **array-RQMC** with  $n = 2^{14}$  in red and for **MC** in blue.

Theoretical dist. is uniform (black dots).



## Example

Let  $Y = \theta U + (1 - \theta)V$ , where  $U, V$  indep.  $U(0, 1)$  and  $\theta \in [0, 1)$ .  
This  $Y$  has cdf  $G_\theta$ .

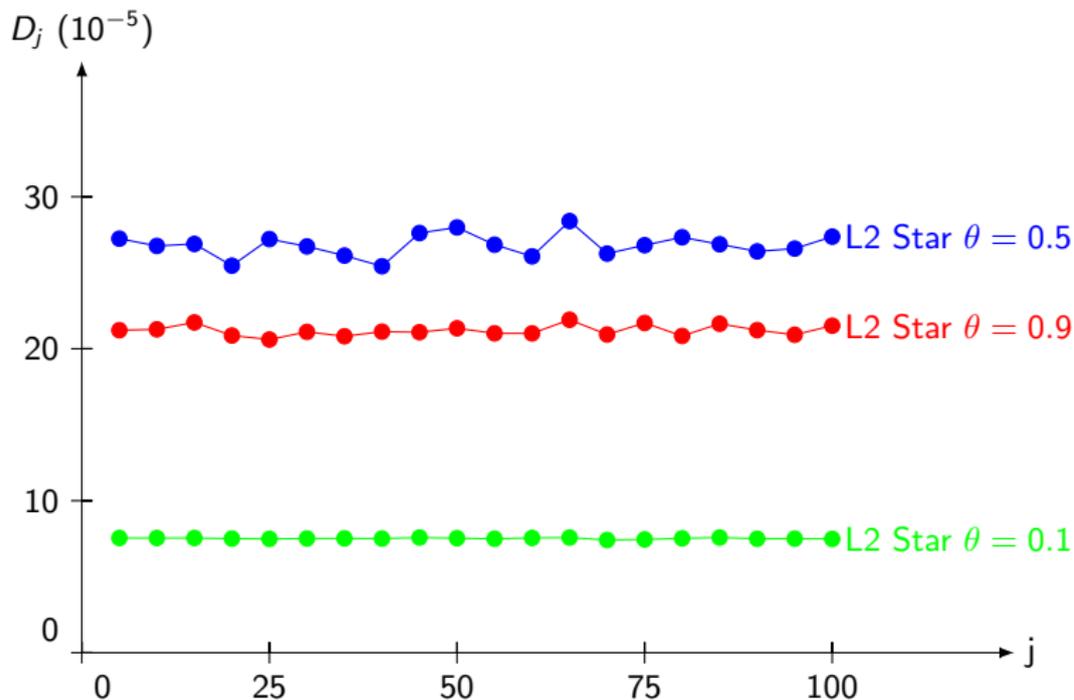
Markov chain is defined by

$$X_0 = U_0; \quad X_j = \varphi_j(X_{j-1}, U_j) = G_\theta(\theta X_{j-1} + (1 - \theta)U_j), j \geq 1$$

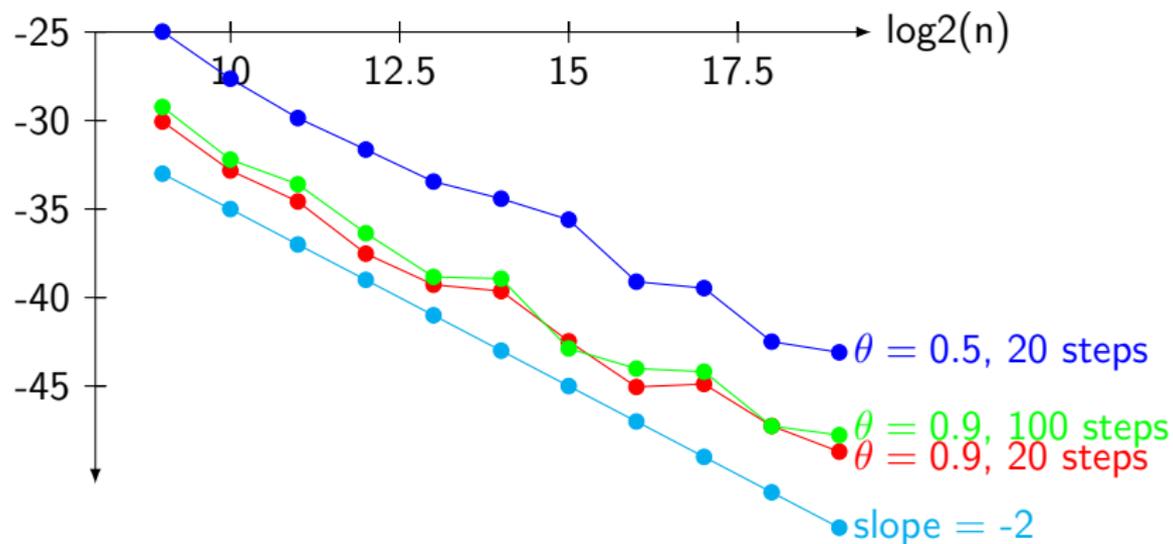
where  $U_j \sim U(0, 1)$ . Then,  $X_j \sim (0, 1)$ .

Define  $g_j(X_j) = X_j$

$\log D_j$  as a function of  $j$ , for  $n = 4093 \approx 2^{12}$



$\log_2 \text{Var}[\hat{\mu}_{\text{rqmc},j,n}]$  as a function of  $\log_2 n$



# Convergence results and proofs

L., Lécot, Tuffin [2008] + some extensions.

Simple case: suppose  $\ell = d = 1$ ,  $\mathcal{X} = [0, 1]$ , and  $X_j \sim U(0, 1)$ . Define

$$\Delta_j = \sup_{x \in \mathcal{X}} |\hat{F}_j(x) - F_j(x)| \quad (\text{discrepancy of states})$$

$$V(g_j) = \int_0^1 |dg_j(x)| \quad (\text{variation of } g_j)$$

**Theorem.**  $|\bar{Y}_{n,j} - E[g_j(X_j)]| \leq \Delta_j V(g_j)$ .

## Convergence results and proofs

**Assumption 1.**  $\varphi_j(x, u)$  non-decreasing in  $u$ . That is, we use inversion to generate next state from cdf  $F_j(z | \cdot) = P[X_j \leq z | X_{j-1} = \cdot]$ .

$$\text{Let } \Lambda_j = \sup_{0 \leq z \leq 1} V(F_j(z | \cdot)).$$

## Convergence results and proofs

**Assumption 1.**  $\varphi_j(x, u)$  non-decreasing in  $u$ . That is, we use inversion to generate next state from cdf  $F_j(z | \cdot) = P[X_j \leq z | X_{j-1} = \cdot]$ .

$$\text{Let } \Lambda_j = \sup_{0 \leq z \leq 1} V(F_j(z | \cdot)).$$

**Assumption 2.** Each square of  $\sqrt{n} \times \sqrt{n}$  grid has one RQMC point.

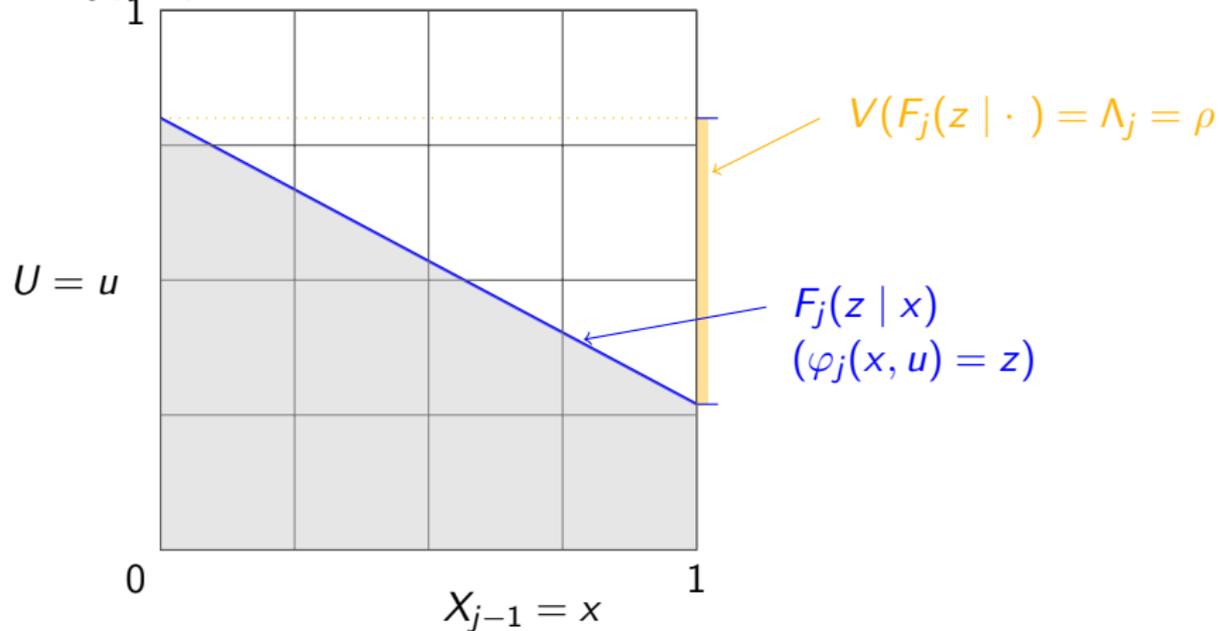
**Proposition.** (Worst-case error.) Under Assumptions 1 and 2,

$$\Delta_j \leq n^{-1/2} \sum_{k=1}^j (\Lambda_k + 1) \prod_{i=k+1}^j \Lambda_i.$$

**Corollary.** If  $\Lambda_j \leq \rho < 1$  for all  $j$ , then

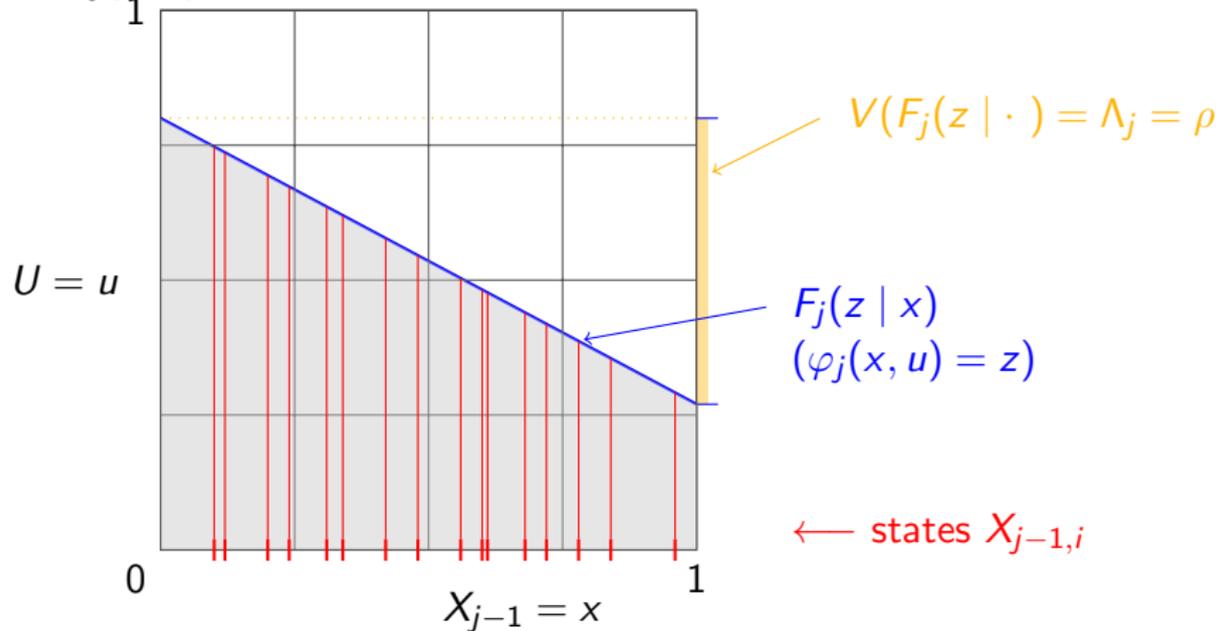
$$\Delta_j \leq \frac{1 + \rho}{1 - \rho} n^{-1/2}.$$

Let  $\varphi_j(x, u)$  non-decreasing in  $x$  and  $u$ . Fix  $z$ .



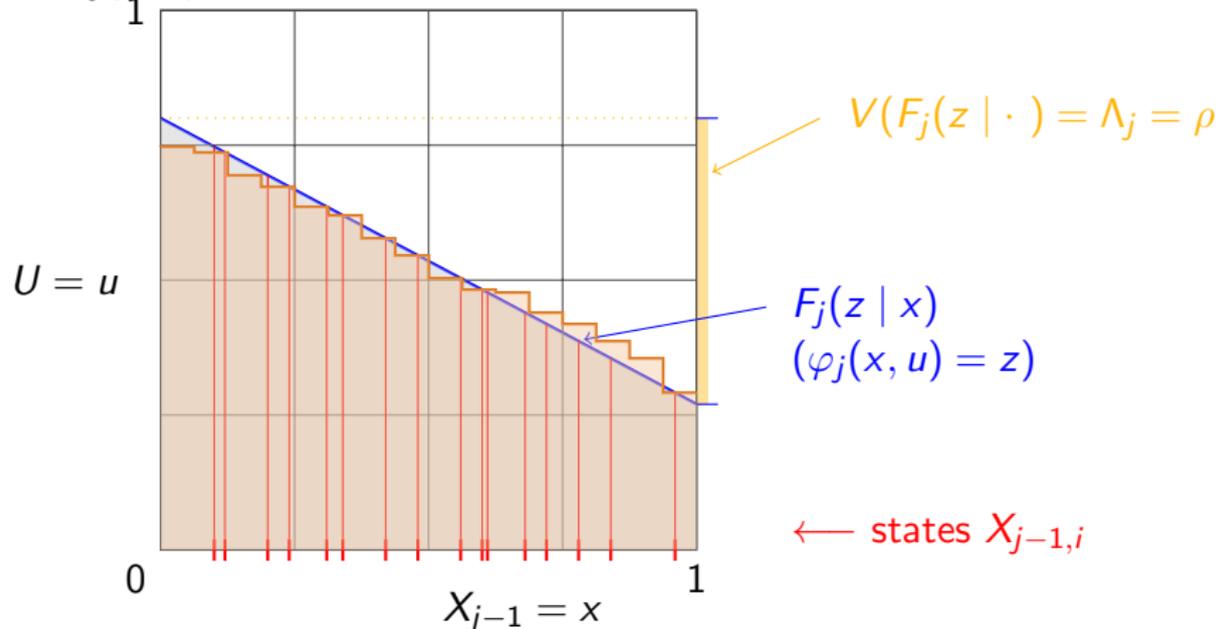
$F_j(z) = \mathbb{P}[X_j \leq z] = \text{size of blue area.}$

Let  $\varphi_j(x, u)$  non-decreasing in  $x$  and  $u$ . Fix  $z$ .



$F_j(z) = \mathbb{P}[X_j \leq z] = \text{size of blue area.}$

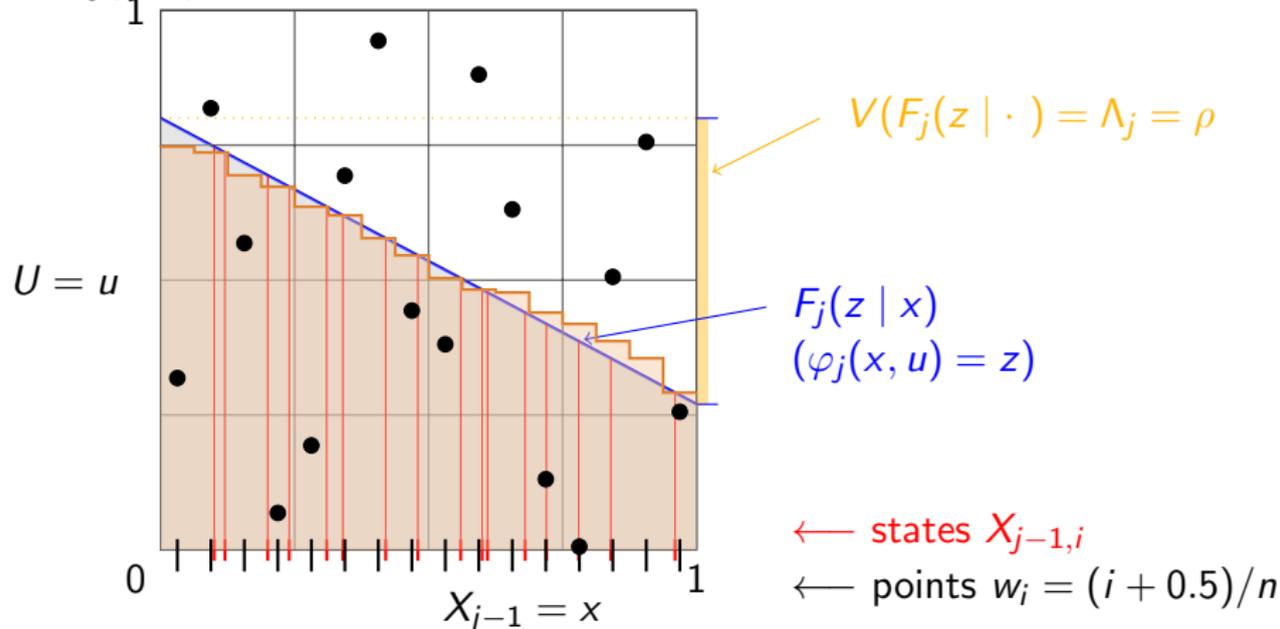
Let  $\varphi_j(x, u)$  non-decreasing in  $x$  and  $u$ . Fix  $z$ .



$F_j(z) = \mathbb{P}[X_j \leq z] = \text{size of blue area.}$

$\tilde{F}_j(z) = \mathbb{P}[X_j \leq z | X_{j-1} \sim \hat{F}_{j-1}] = \text{area of histogram.}$

Let  $\varphi_j(x, u)$  non-decreasing in  $x$  and  $u$ . Fix  $z$ .

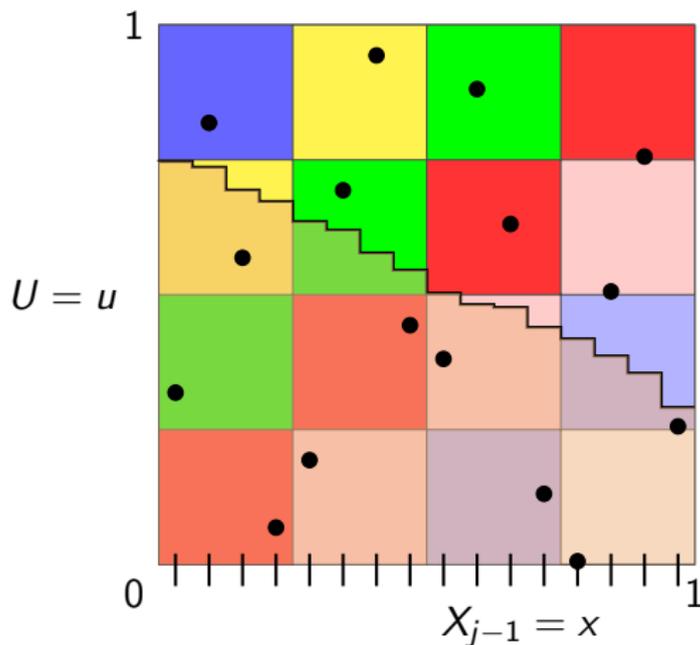


$F_j(z) = \mathbb{P}[X_j \leq z] = \text{size of blue area.}$

$\tilde{F}_j(z) = \mathbb{P}[X_j \leq z | X_{j-1} \sim \hat{F}_{j-1}] = \text{area of histogram.}$

$\hat{F}_j(z) = \text{fraction of the points that fall in histogram.}$

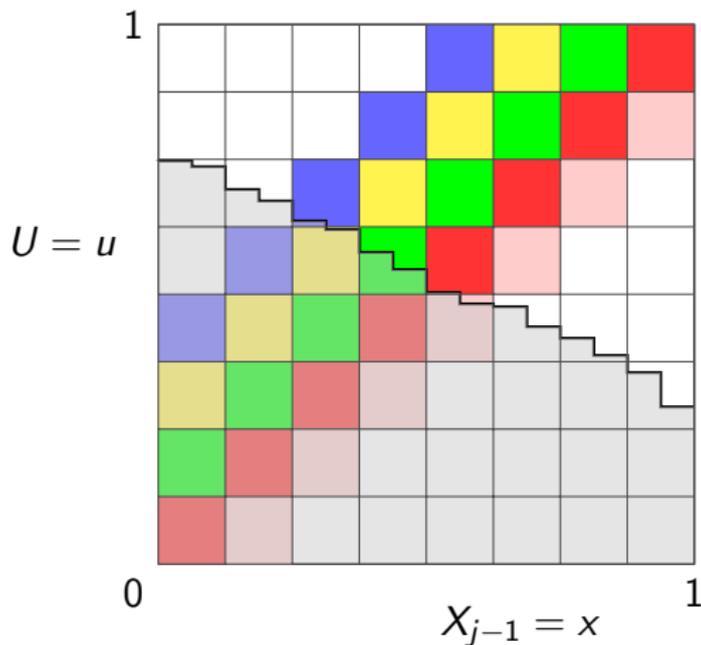
$\Delta_j = \sup_{0 \leq z \leq 1} |\hat{F}_j(z) - F_j(z)|.$



$2\sqrt{n} - 1$  diagonal strings of squares.

The boundary crosses at most one square in each string.

At most  $2\sqrt{n} - 1$  squares out of  $n$  may contribute to  $|\hat{F}_j(z) - \tilde{F}_j(z)|$ .



$\sqrt{n} \times \sqrt{n}$  squares:  $2\sqrt{n} - 1$  diagonal strings of squares.  
 The boundary crosses at most one square in each string.  
 So at most  $2\sqrt{n} - 1$  squares may contribute to the error (or variance), and

$$\text{Var}[\hat{F}_j(z) - \tilde{F}_j(z)] \leq \frac{(2\sqrt{n} - 1)}{4n^2} \leq \frac{n^{-3/2}}{2}.$$

## Variance bound for stratified sampling

**Assumption 3.** Assump. 2 (one point per square) + second coordinate of each point is uniformly dist. in square, and these are **independent** or have **negative covariance**.

**Proposition.** Under Assump. 3,

$$\text{Var}[\bar{Y}_{n,j}] \leq \left( \frac{1}{4} \sum_{k=1}^j (\Lambda_k + 1) \prod_{i=k+1}^j \Lambda_i^2 \right) V^2(g_j) n^{-3/2}.$$

**Corollary.** If all  $\Lambda_j \leq \rho < 1$ , then

$$\text{Var}[\bar{Y}_{n,j}] \leq \frac{1 + \rho}{4(1 - \rho^2)} V^2(g_j) n^{-3/2}.$$

Works also with RQMC if we can show that for any pair of small squares, the **indicators** that the two RQMC point of those squares are in the histogram **do not have a positive covariance**.

## In $\ell = 1$ and $d > 1$

RQMC points are now in  $d + 1$  dimensions. Unit cube partitioned in  $n = k^{d+1}$  subcubes.

**Assumption 4.** Assump. 2 (one point per square) + the randomized parts of the points are pairwise **independent** in their squares.

**Proposition.** Under Assump. 4, if  $\varphi_j$  is monotone non-decreasing,

$$\text{Var}[\bar{Y}_{n,j}] = \mathcal{O}(n^{-1-1/(d+1)}).$$

Consider **diagonal string of squares** from  $(0, \dots, 0)$  to  $(1, \dots, 1)$  and all parallel diagonal strings. There are less than  $(d + 1)k^d$  of those, and the histogram boundary can cross at most one square in each. Then

$$\text{Var}[\hat{F}_j(z) - \tilde{F}_j(z)] < \frac{(d + 1)k^d}{4n^2} = \frac{d + 1}{4} n^{-1-1/(d+1)}.$$

## Conclusion

Empirically, the variance converges as  $\mathcal{O}(n^{-2})$  for some examples, even for a large number of steps.

We have convergence proofs for special cases, but not yet  $\mathcal{O}(n^{-2})$ .