Quasi-Random Sampling in Computational Finance

Christiane Lemieux

Deaprtment of Statistics and Actuarial Science

University of Waterloo

Recent Advances in Mathematical Finance and Insurance
Fields-CFI Conference
September 27, 2023

Problem setup

- In many applications, in particular in finance and insurance, quantities of interest can be written as $\mathbb{E}[\Psi_0(\mathbf{X})] = \int \Psi_0(\mathbf{x}) dH(\mathbf{x})$, where

  $\mathbf{X} = (X_1, \ldots, X_d) \in \mathbb{R}^d$ is a random vector on a probability space $(\Omega, \mathcal{F}, \mathbb{P})$ with distribution function $H: \mathbb{R}^d \to [0, 1]$

  $\Psi_0 : \mathbb{R}^d \to \mathbb{R}$ is a measurable function.

- MC and QMC methods have been used widely on such problems, with d sometimes very large. In most cases, the dependence among the X_j's has been modeled by a multivariate normal distribution (MVN) *(Paskov and Traub (1995), Caflisch and Moskowitz (1995), Acworth, Broadie and Glasserman (1997), Wang and Sloan (2005), Lin and Wang (2008), etc).*

Goals

- QMC methods have most been used on problems involving stochastic models that can be formulated in terms of *independent* random variables.

$$\text{ex: } \mathbf{X} \sim MVN \longleftarrow \mathbf{Z} \sim \text{iid } N(0,1) \longleftarrow \mathbf{U} \sim U([0,1)^d]$$

- Much less has been done with QMC methods on more complex stochastic problems modelling dependence through *copulas*.
- In this talk we're interested in setups where
  - we want to go beyond MVN (e.g., copulas), and/or
  - we have to (or choose to) use something other than inversion

  and we want to study how we can use QMC methods successfully for such problems.

Agenda

1. Overview of Quasi-Random Sampling
2. Combination of QMC with Copula Sampling
   - QMC-based simulation of copula models
   - Numerical examples
3. Combination of QMC with Sampling Methods other than Inversion
   - Methods for the black-box setting
   - Methods in the acceptance-rejection setting
   - Numerical examples

Talk based on:

- "Quasi-random numbers for copula models" by Cambou, Hofert, Lemieux, published in Statistics and Computing, 2017
- "Quasi-random sampling with black box or acceptance-rejection inputs" by Hintz, Hofert and Lemieux published in Advances in Modeling and Simulation – A Festschrift for Pierre L'Ecuyer (Springer, 2022)
- "On the dependence structure and quality of scrambled (t, m, s)-nets, by Wiart, Lemieux and Dong, published in Monte Carlo Methods and Applications, 2021.
- "Dependence Properties of Scrambled Halton Sequences", by Dong and Lemieux, published in Mathematics in Computers and Simulation, 2022

1. Overview of Quasi-Random Sampling

- Main application is high-dimensional integration, where the goal is to evaluate

$$\mu = \int_{[0,1)^d} f(\mathbf{u}) d\mathbf{u}.$$

- Original problem of estimating $\mathbb{E}(\Psi_0(\mathbf{X}))$ can be written this way by viewing $f = \Psi_0 \circ g$ where g transforms $\mathbf{U}$ into $\mathbf{X}$, e.g., for $\mathbf{X} \sim \mathrm{MVN}(\boldsymbol{\mu}, \Sigma)$, take $\mathbf{X} = A(\Phi^{-1}(U_1), \ldots, \Phi^{-1}(U_d))^\top + \boldsymbol{\mu}$, where A is such that $AA^\top = \Sigma$
- With MC, we choose an iid sample $P_n = \{\mathbf{u}_1, \ldots, \mathbf{u}_n\}$ and form the estimator

$$\hat{\mu} = \frac{1}{n} \sum_{i=1}^{n} f(\mathbf{u}_i), \qquad \mathbf{u}_i \sim U(0,1)^d$$

- **Idea of QMC**: replace the random sample used in MC by a point set P_n that is more uniform $\Rightarrow$ low-discrepancy point set

Low-discrepancy point sets

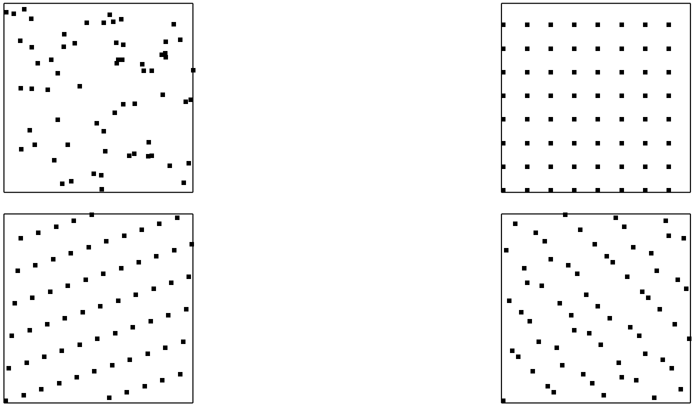


Figure: Four different point sets with $n = 64$: pseudorandom (top left), rectangular grid (top right), Korobov lattice (bottom left), and Sobol' (bottom right).

Theoretical properties of QMC

If we use a low-discrepancy point set $P_n = \{\mathbf{u}_1, \dots, \mathbf{u}_n\}$ to form an approximation

$$Q_n = \frac{1}{n}\sum_{i=1}^{n} f(\mathbf{u}_i) \qquad \text{for} \qquad \mu = \int_{[0,1)^d} f(\mathbf{u})d\mathbf{u}$$

then can show that for "well-behaved functions", we have

$$E_n = |\mu - Q_n| \leqslant V(f)D^*(P_n),$$

where $V(f)$ is the variation of f in the sense of Hardy and Krause, and $D^*(P_n)$ is the *star discrepancy of* P_n.

Discrepancy

Consider "boxes" with one corner at the origin and other corner **v** in $[0,1)^d$. Compare **fraction** of points from P_n falling in the box with **volume** of the box. More precisely,

$$D^*(P_n) = \sup_{\mathbf{v}\in[0,1)^d} |E_n(\mathbf{v})/n - J(\mathbf{v})|$$

where $E_n(\mathbf{v})$ counts **how many points** of P_n are in the box $B(\mathbf{v}) = [0, v_1) \times \ldots \times [0, v_d)$ and $J(\mathbf{v}) = v_1 \ldots v_d$ is the **volume** of the box.

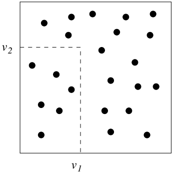


Figure: The dotted lines show a box with $v_1 = 0.4$ and $v_2 = 0.7$. We see that $E_n(\mathbf{v}) = 6$ out of $n = 23$ points fall in the box, thus producing a difference $|6/23 - v_1 v_2| = 0.019$.

Discrepancy

Consider "boxes" with one corner at the origin and other corner **v** in $[0,1)^d$. Compare **fraction** of points from P_n falling in the box with **volume** of the box. More precisely,

$$D^*(P_n) = \sup_{\mathbf{v} \in [0,1)^d} |E_n(\mathbf{v})/n - J(\mathbf{v})|$$

where $E_n(\mathbf{v})$ counts **how many points** of P_n are in the box $B(\mathbf{v}) = [0, v_1) \times \ldots \times [0, v_d)$ and $J(\mathbf{v}) = v_1 \ldots v_d$ is the **volume** of the box.

Low-discrepancy sequence $\Rightarrow D^*(P_n) \in O(n^{-1} \log^d n) \Rightarrow E_n \in O(n^{-1} \log^d n)$: for fixed d, better than MC's probabilistic error $O(1/\sqrt{n})$, but ...

$$\frac{(\log n)^d/n}{1/\sqrt{n}} \to 0 \text{ as } n \to \infty \text{ but, e.g., } \frac{1}{\sqrt{n}} \leqslant \frac{(\log n)^d}{n} \quad \text{if } d = 10 \text{ and } n < 10^{28}$$

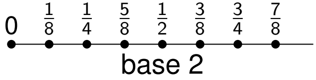
Low-discrepancy sequences: a first example

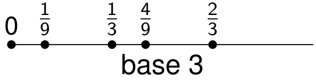
In one dimension, we can construct a sequence of points $u_0, u_1, \ldots$ with a low discrepancy as follows:

1. Choose a base b
2. To define u_i:
   - expand i in base b, i.e., write $i = \sum_{l=0}^{\infty} a_l b^l$:
   - apply *radical-inverse function*:
     $u_i = S_b(i) := \sum_{l=0}^{\infty} a_l b^{-l-1}$,

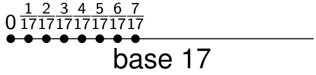
This yields the *van der Corput sequence in base* b*, denoted* S_b *(goes back to 1935)*

van der Corput Sequences





As the base increases, the “space-filling” properties of the sequence deteriorate...



Extending the van der Corput sequence to $d > 1$

How do we do this? First approach:

- use a different base for each dimension (Halton sequence, 1960).
- That is, let S_b denote the van der Corput sequence in base b, and $S_b(n)$ be the nth term of this sequence.
- The Halton sequence in d dimensions is given by $(S_{b_1}, \dots, S_{b_d})$ where the b_j's are pairwise co-primes.
- Typically, take b_j to be the jth prime number.

Advantages: simple to understand and implement.

Disadvantages: doesn't work so well in medium to high dimensions (say above 40 or 50)

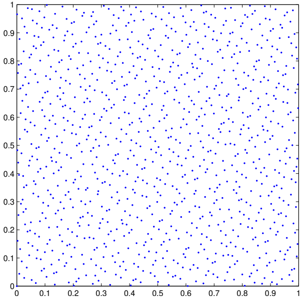
Halton sequence

$\mathbf{u}_1 = (0, 0, 0, \ldots)$ $\quad$ $\mathbf{u}_2 = (1/2, 1/3, 1/5, \ldots)$

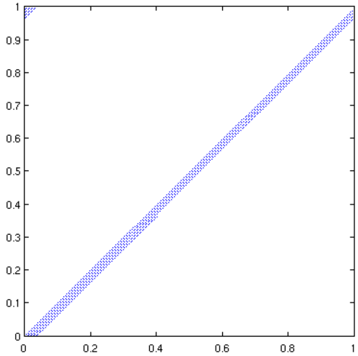
$\mathbf{u}_3 = (1/4, 2/3, 2/5, \ldots)$ $\quad$ $\mathbf{u}_4 = (3/4, 1/9, 3/5, \ldots)$

$\mathbf{u}_5 = (1/8, 4/9, 4/5, \ldots)$

First two dimensions:

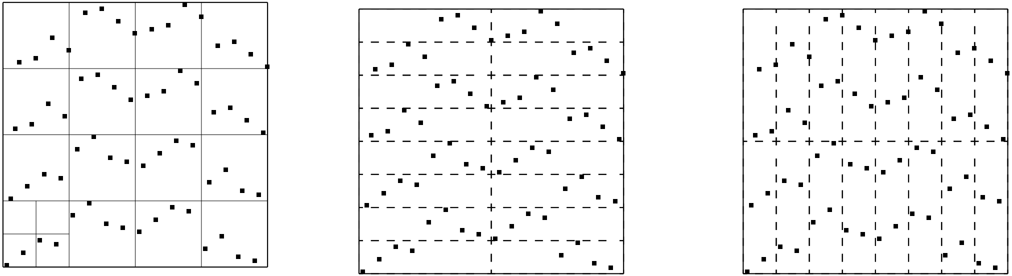


1000 first points of the Halton sequence defined with p_{49} and p_{50} (49th and 50th prime numbers ($p_{49} = 227$ and $p_{50} = 229$))



Second approach to extend VDC to $d > 1$

- If we want to use the same base b in each dimension, can "individualize" the different coordinates by applying a well-chosen **linear transformation** to the digits in the base b expansion of i.
- First construction based on this idea is the Sobol' sequence (1967).
- More general definition is that of **digital sequences**.



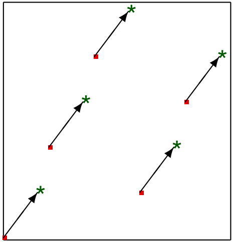
Randomized quasi-Monte Carlo

Problem: Bounds for deterministic QMC are not very helpful to provide actual error estimates

Idea of RQMC: randomize low-discrepancy point set P_n so that

- each $\mathbf{u}_i \sim \mathcal{U}([0,1)^d)$
- low discrepancy of P_n is preserved after the randomization

Example: Cranley-Patterson rotation: $\tilde{\mathbf{u}}_i = (\mathbf{u}_i + \mathbf{r}) \bmod 1$, where $\mathbf{r} \sim \mathcal{U}([0,1)^d)$.



Scrambling

- Method to randomize digital nets that introduces more randomness than a simple (digital) shift
- **(Nested uniform) scrambling** proposed by Owen in 1995: in base b, in each dimension, first randomly permute intervals of size $1/b$; then within each of them, split in b and randomly permute intervals of size b^{-2}, etc.
- **Random linear scrambling with digital shift:** less computationally expensive than nested scrambling
- Have shown two forms of scrambling gives an estimator with the same variance (Wiart, Lemieux, Dong, 2021)
- For smooth enough functions, Owen showed variance of scrambled digital sequences is in $O(n^{-3}\log^d n)$.

Variance Estimation with RQMC

- Define rQMC point set $\tilde{P}_n^b = \{\tilde{\mathbf{u}}_{i,b}, 0 \leqslant i < n\}$ and get unbiased estimator

$$\hat{\mu}_{\mathrm{rQMC},b} = \frac{1}{n}\sum_{i=0}^{n-1} f(\tilde{\mathbf{u}}_{i,b}).$$

- Using independent $\tilde{P}_n^b$, $b = 1, \ldots, B$, can use $\{\hat{\mu}_{\mathrm{rQMC},b},\ 1 \leqslant b \leqslant B\}$ to estimate the variance of $\hat{\mu}_{\mathrm{rQMC},b}$ by

$$\hat{\sigma}^2_{\mathrm{rQMC}} = \frac{1}{B-1}\sum_{b=1}^{B} (\hat{\mu}_{\mathrm{rQMC},b} - \bar{\mu}_{\mathrm{rQMC}})^2.$$

- For fixed computing budget, how to choose B vs n? Take B at least 20-25 to get reasonable estimates.

2. Combination with Copula Sampling

Using *Sklar's Theorem*, we write the joint cdf H of **X** as

$$H(\boldsymbol{x}) = C(F_1(x_1), \ldots, F_d(x_d)), \quad \boldsymbol{x} \in \mathbb{R}^d,$$

where $F_j(x) = P(X_j \leqslant x)$, are the marginal distribution functions of H and C is a copula.
Means we can write

$$\mathbb{E}[\Psi_0(\mathbf{X})] = \mathbb{E}_C[\Psi(\mathbf{U})] = \int_{[0,1)^d} \Psi(\mathbf{u}) \mathrm{d}C(\mathbf{u}),$$

where $\mathbf{U} = (U_1, \ldots, U_d) \sim C$, $\Psi : [0,1]^d \to \mathbb{R}$ is defined as

$$\Psi(u_1, \ldots, u_d) = \Psi_0(F_1^{-1}(u_1), \ldots, F_d^{-1}(u_d)),$$

and $F_j^{-1}(p) = \inf\{x \in \mathbb{R} : F_j(x) \geqslant p\}$, $j \in \{1, \ldots, d\}$ are the marginal quantile functions.

Hence to estimate $\mathbb{E}_C[\Psi(\mathbf{U})]$, proceed as follows:

1) sample $\mathbf{U}_i \sim C$;
2) transform each U_{ij} according to $X_{ij} = F_j^{-1}(U_{ij}), j = 1, \ldots, d$, and use

$$\frac{1}{n}\sum_{i=1}^{n}\Psi_0(\mathbf{X}_i) = \frac{1}{n}\sum_{i=1}^{n}\Psi(\mathbf{U}_i)$$

$\rightarrow$ How should we sample a copula if we use (r)QMC?

- To sample a d-dimensional copula C, use transformation $\phi_C : [0,1)^k \to [0,1)^d$, $k \geqslant d$, such that

$$\mathbf{U} = \phi_C(\mathbf{U}') \sim C \text{ when } \mathbf{U}' \sim \mathrm{U}(0,1)^k.$$

- In other words, ϕ_C transforms $\mathrm{U}(0,1)^k$ random variables to random variables with distribution function C.

$$\mathbf{U}' \sim \mathrm{U}([0,1)^k) \longrightarrow \mathbf{U} \sim C \longrightarrow \mathbf{X} \sim H$$

- With QMC-based sampling, prefer $k = d$, i.e., $\phi_C : [0,1)^d \to [0,1)^d$. Would also like $\phi_C = (\phi_{C,1}, \dots, \phi_{C,d})$ such that *each* $\phi_{C,j}$ *is monotone in each argument, and smooth*.
- Achieved by the conditional distribution method (CDM), which is the only known algorithm for sampling an arbitrary copula without further knowledge about its structure.

Conditional Distribution Method

For $j \in \{2, \dots, d\}$, let

$$C(u_j \mid u_1, \dots, u_{j-1}) = \mathbb{P}(U_j \leqslant u_j \mid U_1 = u_1, \dots, U_{j-1} = u_{j-1})$$

denote the *conditional copula of* U_j *at* u_j *given* $U_1 = u_1, \dots, U_{j-1} = u_{j-1}$. If $C^{-1}(u_j \mid u_1, \dots, u_{j-1})$ denotes the corresponding quantile function, the CDM is given as follows (Schmitz, 2003):

Theorem (Conditional distribution method)

Let C *be a* d*-dimensional copula,* $\mathbf{U}' \sim U[0,1]^d$*, and* ϕ_C^{CDM} *be given by*

$$\begin{aligned} U_1 &= U_1', \\ U_2 &= C^{-1}(U_2' \mid U_1), \\ &\vdots \\ U_d &= C^{-1}(U_d' \mid U_1, \dots, U_{d-1}). \end{aligned}$$

CDM for Archimedean copulas

Assuming the generator ψ is sufficiently often differentiable, the conditional Archimedean copulas follow from above theorem and are given by

$$C(u_j \,|\, u_1, \ldots, u_{j-1}) = \frac{\psi^{(j-1)}(W_j)}{\psi^{(j-1)}(W_{j-1})},$$

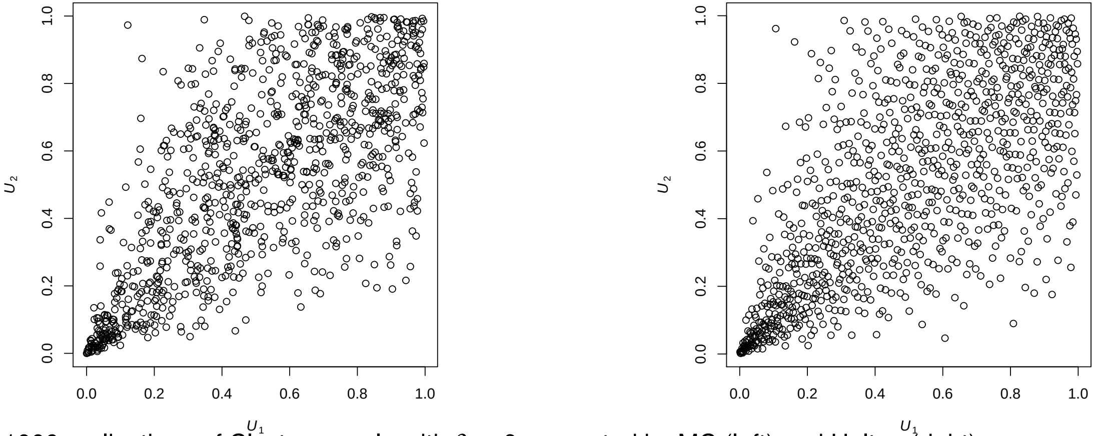
where $W_j = \sum_{k=1}^{j} \psi^{-1}(u_k)$ and thus

$$C^{-1}(u_j \,|\, u_1, \ldots, u_{j-1}) = \psi\left(\psi^{(j-1)^{-1}}\left(u_j \psi^{(j-1)}(W_{j-1})\right) - W_{j-1}\right).$$

The generator derivatives $\psi^{(j-1)}$ and their inverses $\psi^{(j-1)^{-1}}$ can be challenging to compute.

CDM for Clayton copula

$$C^{-1}(u_j \mid u_1, \dots, u_{j-1}) = \left(1 + \left(1 - (j-1) + \sum_{k=1}^{j-1} u_k^{-\theta}\right)\left(u_j^{-\frac{1}{j-1+1/\theta}} - 1\right)\right)^{-\frac{1}{\theta}}.$$



1000 realizations of Clayton copula with $\theta = 2$ generated by MC (left) and Halton (right).

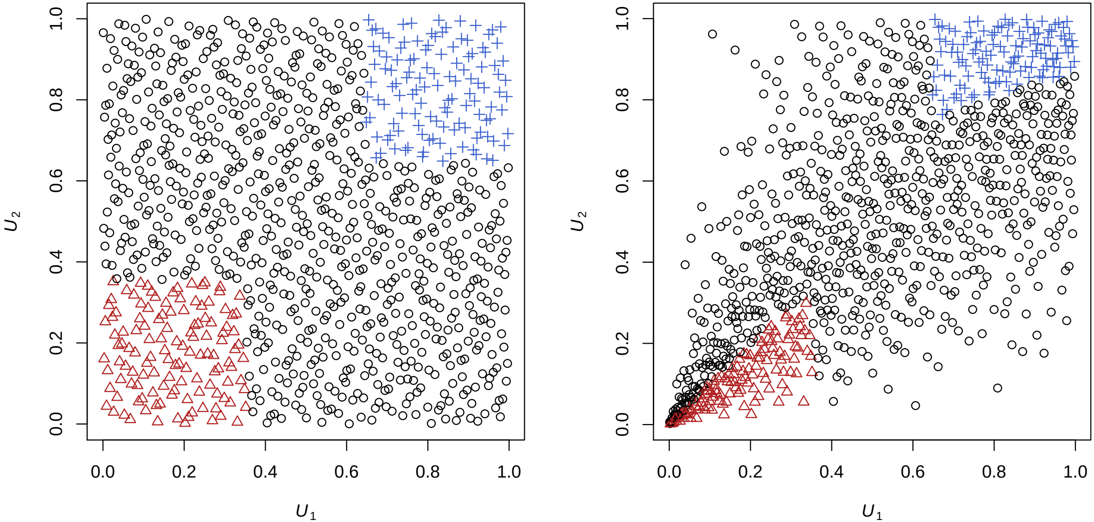
Sampling based on stochastic representation

- Approaches other than the CDM are often used to sample copulas.
- Based on transformations $\phi_C : [0,1)^k \to [0,1)^d$ with $k > d$ making use of well-chosen *stochastic representation*
- For example, the Marshall–Olkin algorithm uses a *conditional independence approach*: based on the fact that for an Archimedean copula C with completely monotone generator ψ,

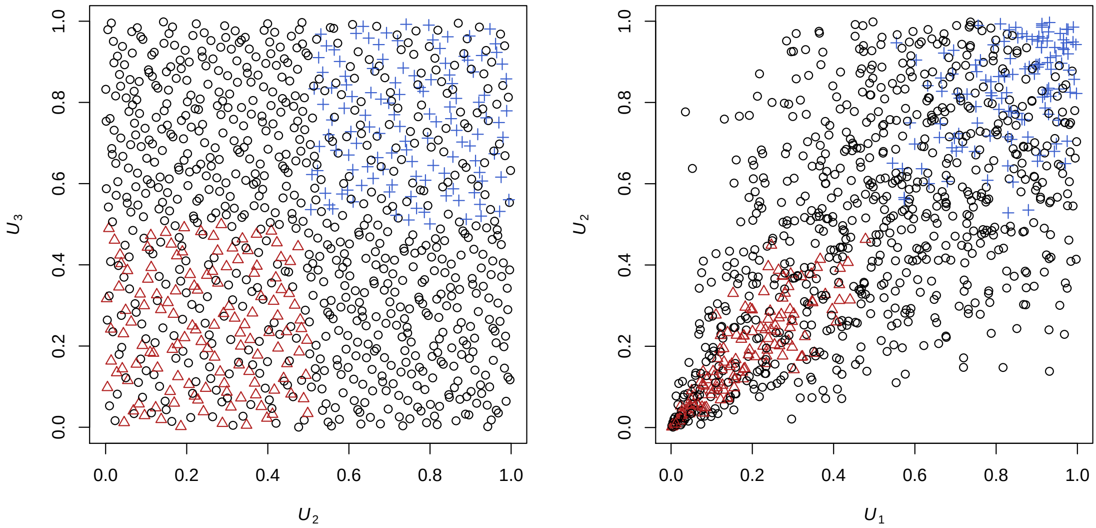
$$\mathbf{U} = (\psi(E_1/W), \ldots, \psi(E_d/W)) \sim C,$$

where $W \sim F = \mathcal{LS}^{-1}[\psi]$ (i.e., F is the distribution whose Laplace transform is ψ), independent of $E_1, \ldots, E_d \sim \text{Exp}(1)$ (iid).
Here $k = d + 1$ if we use inversion to generate E_j's and W.

1000 realizations of the 1st two components of a 3d-Halton sequence with marked points ($\triangle$ and $+$) in the respective regions $[0, \sqrt{1/8}]^2$ and $[1 - \sqrt{1/8}, 1]^2$ (left): corresponding ϕ_C^{CDM}-transformed points (right) to a Clayton copula with $\theta = 2$



1000 realizations of 2nd & 3rd components of a 3d-Halton sequence with marked points ($\triangle$ and $+$) corresponding to respective regions $[0, 0.5]^3$ and $[0.5, 1]^3$ (left): corresponding ϕ_C^{MO}-transformed points (right) to a Clayton copula with $\theta = 2$.



Error behavior

Can work with bound from Hlawka and Mück (1972): $P_n = \{\mathbf{v}_1, \ldots, \mathbf{v}_n\}$, $\mathbf{u}_i = \phi_C(\mathbf{v}_i)$,

$$\left| \frac{1}{n} \sum_{i=1}^{n} \Psi(\mathbf{u}_i) - \mathbb{E}[\Psi(\mathbf{U})] \right| \leqslant D^*(P_n) \|\Psi \circ \phi_C\|_{d,1},$$

where

$$\|\Psi \circ \phi_C\|_{d,1} = \sum_{l=1}^{s} \sum_{\boldsymbol{\alpha}} \int_{[0,1)^l} \left| \frac{\partial^l \Psi \circ \phi_C(v_{\alpha_1}, \ldots, v_{\alpha_l}, \mathbf{1})}{\partial v_{\alpha_1} \cdots \partial v_{\alpha_l}} \right| dv_{\alpha_1} \ldots dv_{\alpha_l}$$

and the second sum is taken over all nonempty subsets $\boldsymbol{\alpha} = \{\alpha_1, \ldots, \alpha_l\} \subseteq \{1, \ldots, d\}$.

With CDM, if the $C(u_j | u_1, \ldots, u_{j-1})$ and Ψ all have finite mixed partial derivatives, then can show $\|\Psi \circ \phi_C\|_{d,1} < \infty$

Conditions to have bounded variation with CDM

Proposition: Assume that Ψ has continuous mixed partial derivatives up to total order d and there exist $m, M, K > 0$ such that for all $\mathbf{u} \in (0,1)^d$, $c(\mathbf{u}) \geqslant m > 0$, $C_{i|1\ldots i-1} = C(u_i \,|\, u_1, \ldots, u_{i-1})$ and

$$\left| \frac{\partial^k C_{i|1\ldots i-1}}{\partial u_{\alpha_1} \cdots \partial u_{\alpha_k}} \right| \leqslant M, \;\; \alpha_1, \ldots, \alpha_k \in \{1, \ldots, i\}, \tag{1}$$

for each $1 \leqslant k \leqslant i \leqslant d$, $1 \leqslant k \leqslant l \leqslant d$ and $\{\alpha_1, \ldots, \alpha_l\} \subseteq \{1, \ldots, d\}$, we have

$$\left| \frac{\partial^k \Psi(u_1, \ldots, u_d)}{\partial u_{\beta_1} \cdots \partial u_{\beta_k}} \right| \leqslant K, \quad \beta_j \in \{\alpha_1, \ldots, \alpha_l\}, \; 1 \leqslant j \leqslant k. \tag{2}$$

Then there exists a constant $C^{(d)}$ s.t. for $\mathbf{u}_i = \phi_C^{\text{CDM}}(\mathbf{v}_i)$, $i = 1, \ldots, n$,

$$\left| \frac{1}{n} \sum_{i=1}^{n} \Psi(\mathbf{u}_i) - \mathbb{E}[\Psi(\mathbf{u})] \right| \leqslant D^*(P_n) K C^{(d)} (M^d/m)^{2d-1}.$$

Results for MO

- Using Thm 2.1 from Constantine and Savits (1996) (multivariate Faa di Bruno), get

$$\frac{\partial^l \Psi \circ \phi_C(v_{\alpha_1}, \ldots, v_{\alpha_l}, \mathbf{1})}{\partial v_{\alpha_1} \cdots \partial v_{\alpha_l}} = \sum_{1 \leqslant |\boldsymbol{\beta}| \leqslant l} \frac{\partial^{|\boldsymbol{\beta}|}\Psi}{\partial^{\beta_1} u_1 \ldots \partial^{\beta_d} u_d} \sum_{s=1}^{l} \sum_{\gamma, \mathbf{k}} c_\gamma \prod_{j=1}^{s} \frac{\partial^{|\gamma_j|} \phi_{C,k_j}(v_{\alpha_1}, \ldots, v_{\alpha_l}, \mathbf{1})}{\partial^{\gamma_{j,1}} v_{\alpha_1} \ldots \partial^{\gamma_{j,l}} v_{\alpha_l}} \tag{3}$$

  where $\boldsymbol{\beta} \in \mathbb{N}_0^d$ and $|\boldsymbol{\beta}| = \sum_{j=1}^{d} \beta_j$.
- Then use properties of MO algorithm to analyze ϕ_{C,k_j}: made easier by fact that each only involves two v_ℓ at a time

Error behaviour for MO for continuous W

Proposition: Let $P_n = \{\mathbf{v}_i, i = 1, \ldots, n\} \subseteq [0,1)^{d+1}$ and $\mathbf{u}_i = \phi_C^{\text{MO}}(\mathbf{v}_i)$. Assume $W \sim F$ is continuously distributed and that:

1. the point set P_n excludes the origin and there exists some $p \geqslant 1$ such that $\min_{1 \leqslant i \leqslant n} v_{i,1} \geqslant 1/pn$;
2. the function Ψ satisfies $|\Psi(\mathbf{u})| < \infty$ for all $\mathbf{u} \in [0,1)^{d+1}$ and

$$\frac{\partial^{|\boldsymbol{\beta}|}\Psi}{\partial^{\beta_1}u_1 \ldots \partial^{\beta_d}u_d} < \infty \text{ for all } \boldsymbol{\beta} = (\beta_1, \ldots, \beta_d), \tag{4}$$

   with $\beta_l \in \{0, \ldots, d\}$ and $|\boldsymbol{\beta}| \leqslant d$;
3. the generator $\psi(\cdot)$ is such that $\psi'(t) + t\psi''(t)$ has at most one zero t^* in $(0, \infty)$ and it satisfies $-t^*\psi'(t^*) < \infty$; and
4. $F^{-1}(1 - 1/pn)$ is in $O(n^a)$ for some constant $a > 0$;

then
$$\left|\frac{1}{n}\sum_{i=1}^{n} \Psi(\mathbf{u}_i) - \mathbb{E}[\Psi(\mathbf{u})]\right| \leqslant C^{(d)}(\log n)D^*(P_n).$$

Error behaviour for MO algorithm for discrete W

Proposition Let $P_n = \{\mathbf{v}_i, i = 1, \dots, n\} \subseteq [0,1)^{d+1}$ and $\mathbf{u}_i = \phi_C^{\text{MO}}(\mathbf{v}_i)$. Assume $W \sim F$ is discrete. If (4) holds and

1. there exists some $p \geqslant 1$ such that the point set P_n satisfies $\max_{1 \leqslant i \leqslant n} v_{i,1} \leqslant 1 - 1/pn$;
2. there exist constants $c > 0$ and $q \in (0,1)$ such that $1 - F(l) \leqslant cq^l$ for $l \geqslant 1$;

then there exists a constant $C^{(d)}$ such that

$$\left| \frac{1}{n} \sum_{i=1}^{n} \Psi(\mathbf{u}_i) - \mathbb{E}[\Psi(\mathbf{u})] \right| \leqslant C^{(d)} (\log n) D^*(P_n).$$

Numerical Examples

- Asset prices X_j for jth asset at time T, for $j = 1, \ldots, d$.
- Same marginal distribution for each j but dependence across assets induced by copula.
- Consider functionals of the aggregate sum $S = X_1 + \ldots + X_d$:
- **Basket call option**: $\Psi_0(S) = \max(0, S/d - K)$.
- **Value-at-Risk**: $\mathrm{VaR}_{0.99}(S) = F_S^{-1}(0.99) = \inf\{x \in \mathbb{R} : F_S(x) \geqslant 0.99\}$
- **Expected shortfall**: $\mathrm{ES}_{0.99}(S) = \frac{1}{1-0.99} \int_{0.99}^{1} F_S^{-1}(u)du$

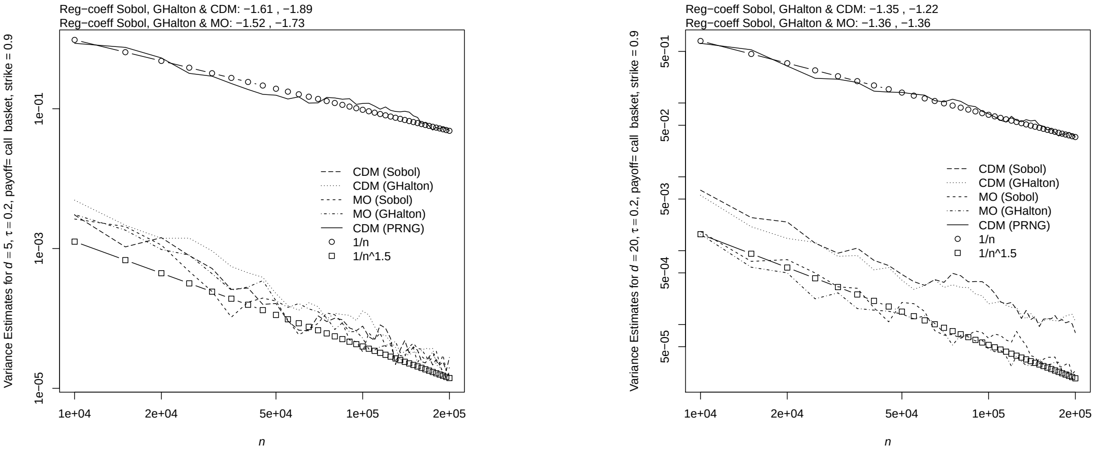


Figure: Variance estimates for the functional Basket Call with lognormal margins based on $\mathrm{B} = 25$ repetitions for a Clayton copula with parameter such that Kendall's tau equals 0.2 for $\mathrm{d} = 5$ (left) and $\mathrm{d} = 20$ (right).

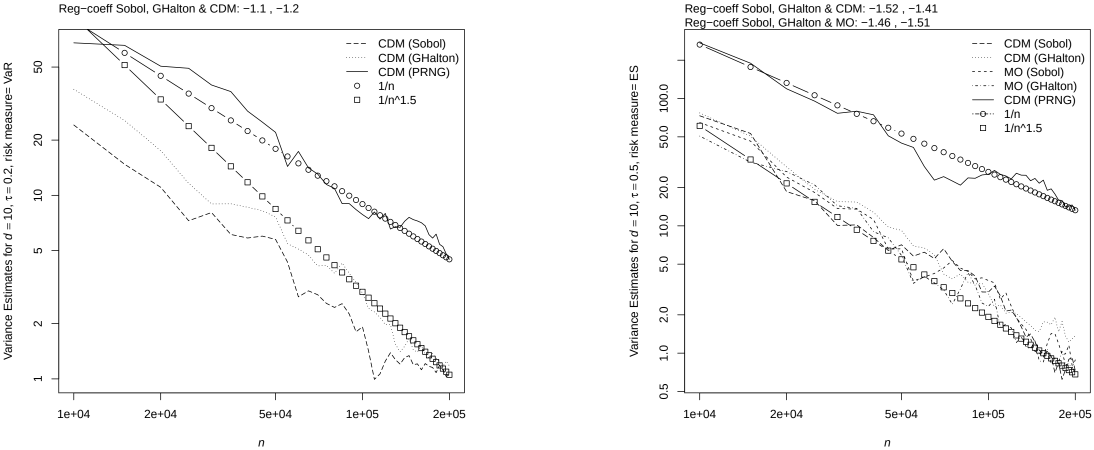


Figure: Variance estimates for the functional $\text{Var}_{0.99}$ with lognormal margins and exchangeable $\mathtt{t}$ copula (left); and for $\text{ES}_{0.99}$ with Pareto margins for a Clayton copula (right) for $\mathrm{d} = 10$

3- Combination with Sampling Methods other than Inversion

Assume $\mathbf{X} = (\mathbf{Y}, W)$, i.e., consider the problem of estimating the quantity

$$\mu = \mathbb{E}(\Psi_0(\mathbf{Y}, W)) \tag{5}$$

where and $\mathbf{Y} \sim F_{\mathbf{Y}}$ is a d-dimensional random vector independent of $W \sim F_W$.

Goal: successfully combine rQMC with following classes of sampling methods for W:

1. Black-box case. Here, we assume that we have a (random) function $R_W : \mathbb{N} \to \mathbb{R}^n$ such that if $R_W(n) = \mathbf{W}$ for $\mathbf{W} = (W_1, \ldots, W_n)$ then $W_i \overset{\text{ind.}}{\sim} F_W$ for $i = 1, \ldots, n$. As such, we have a “black box” function that returns samples from F_W of any size.
2. AR algorithms for W, where the proposal (or envelope) distribution and the acceptance decision can be sampled by inversion of uniforms.

No inversion, no monotonicity? Or can we get it back?

RQMC as Dependence Sampling

With RQMC sampling:

$\tilde{u}_{1,1}, \tilde{u}_{1,2}, \ldots, \tilde{u}_{1,d} \sim U(0,1)^d \Rightarrow$ **same as MC**

$\tilde{u}_{2,1}, \tilde{u}_{2,2}, \ldots, \tilde{u}_{2,d} \sim U(0,1)^d \Rightarrow$ **same as MC**

$\vdots \qquad \vdots \qquad \qquad \vdots$

$\tilde{u}_{n,1}, \tilde{u}_{n,2}, \ldots, \tilde{u}_{n,d} \sim U(0,1)^d \Rightarrow$ **same as MC**

dependence across points
$\Rightarrow$ **not like MC**

$\Rightarrow$ each $\mathbf{X}_i$ generated from $\tilde{\mathbf{u}}_i$ has the **same distribution** as when using MC simulation

Negative Dependence

- Methods likes antithetic variates and Latin Hypercube Sampling have a form of negative dependence that implies they can do no worse than MC for functions monotone in each coordinate
- Relies on fact that for any i, ℓ, $P(U_{ij} \leqslant u, U_{\ell,j} \leqslant v) \leqslant uv$, for each $j = 1, \ldots, d$, $0 \leqslant u, v \leqslant 1$ (coordinate-wide negative dependence)
- For digital scrambled $(0, m, d)-$nets, can show that $P(\mathbf{U}_i \leqslant \mathbf{u}, \mathbf{U}_\ell \leqslant \mathbf{v}) \leqslant \prod_{j=1}^{s} u_j v_j$ (negative lower orthant dependence)
- Also true for scrambled Halton sequences
- Sufficient condition on f to ensure scrambled $(0, m, d)-$net is no worse than MC is much stronger (quasi-monotonicity)

Based on Wiart, Lemieux, Dong (2021), and Dong and Lemieux (2022)

Methods for the black-box setting

Classical MC estimator $\hat{\mu}_n^{\text{mc}}$ based on n samples for μ can be written as

$$\hat{\mu}_n^{\text{mc}} = \frac{1}{n} \sum_{i=1}^{n} \Psi_0(T_{\mathbf{Y}}(\mathbf{U}_i), W_i), \tag{6}$$

where $\mathbf{U}_i \overset{\text{ind.}}{\sim} U(0,1)^{d+k}$ is independent of $W_1, \ldots, W_n \sim F_W$ obtained by calling $R_W(n)$.

Methods for the black-box setting

Typical way to apply RQMC: we first rewrite $\mu = \mathbb{E}(\Psi_0(\mathbf{Y}, W))$ as

$$\mu = \int_{(0,1)^{d+1}} \Psi_0(T_{\mathbf{Y}}(u_{1:d}), Q(u_{d+1}))\, d\mathbf{u}, \tag{7}$$

where $\mathbf{u} = (u_{1:d}, u_{d+1})$ with $u_{1:d} = (u_1, \ldots, u_d)$, and $Q : [0, 1] \to \mathbb{R}$ is the quantile function $F_W^{\leftarrow}$

If Q was known could use inversion to sample W_i and thus could use RQMC sampling with $(d+1)-$dimensional point set $\tilde{P}_{b,n} = \{\mathbf{u}_{b,1}, \ldots, \mathbf{u}_{b,n}\} \subseteq [0, 1)^{d+1}$, where $\mathbf{u}_{b,i} = (u_{b,i,1}, \ldots, u_{b,i,d+1})$ for $b = 1, \ldots, B$ and then use

$$\hat{\mu}^{\text{rqmc}}_{b,n} = \frac{1}{n}\sum_{i=1}^{n} \Psi_0(T_{\mathbf{Y}}(\mathbf{u}_{b,i,1:d}), Q(u_{b,i,d+1})), \quad b = 1, \ldots, B, \tag{8}$$

Methods based on the empirical quantile function

- Since Q is not known we'll construct approximation using values from random sample $\{W_{b,1}, \ldots, W_{b,n}\}$ of F_W obtained by calling $R_W(Bn)$.
- How? use the sampled $W_{b,i}$ to construct empirical quantile function $\hat{Q}_{n,b}, b = 1, \ldots, B$
- Then use coordinate $d+1$ of RQMC point set to sample from W by inverting $\hat{Q}_{n,b}$, i.e., return $\hat{Q}_{n,b}(u_{b,i,d+1}) = W_{b,(\lceil n u_{b,i,d+1} \rceil)}$, where $W_{b,(1)} \leqslant \cdots \leqslant W_{b,(n)}$, i.e., if $\frac{j-1}{n} < u_{b,i,d+1} \leqslant \frac{j}{n}$ then return $W_{b,(j)}$.
- Let

$$\hat{\mu}_{b,n}^{\text{b-eqf}} = \frac{1}{n} \sum_{i=1}^{n} \Psi_0(T_{\mathbf{Y}}(\mathbf{u}_{b,i,1:d}), W_{b,(\lceil n u_{b,i,d+1} \rceil)}), \quad b = 1, \ldots, B,$$

  i.e., in each of the B randomizations, estimate Q by its empirical quantile function $\hat{Q}_{n,b}$ obtained from n calls to the black-box function $R_W(n)$.

Rank-based estimator

- Let $r^n(u_{b,i,d+1})$ be the rank of $u_{b,i,d+1}$ among $u_{b,1,d+1}, \dots, u_{b,n,d+1}$.
  Rank-based estimator is

$$\hat{\mu}_{b,n}^{\text{b-rk}} = \frac{1}{n} \sum_{i=1}^{n} \Psi_0(T_{\mathbf{Y}}(\mathbf{u}_{b,i,1:d}), W_{b,(r^n(u_{b,i,d+1}))}), \quad b = 1, \dots, B. \tag{9}$$

- $\hat{\mu}_{b,n}^{\text{b-rk}}$ uses $W_{b,i}$ is used **exactly once**
- If $\tilde{P}_{b,n,d+1}$ is properly stratified, then $\hat{\mu}_{b,n}^{\text{b-rk}}$ and $\hat{\mu}_{b,n}^{\text{b-eqf}}$ coincide
- Could also use the Bn samples $W_{b,i}$ to construct ONE empirical quantile function

Proposition: Let $b \in \{1, \dots, B\}$ and let $\tilde{P}_{b,n,d+1}$ be properly stratified. Then $\hat{\mu}_{b,n}^{\text{b-rk}}$ (and therefore $\hat{\mu}_{b,n}^{\text{b-eqf}}$) is unbiased for μ.

Methods for the Acceptance/Rejection (AR) setting

We assume W can be sampled using AR and explore how we can apply RQMC in this setting.

AR: (assume q_W is pdf of W) choose proposal pdf p_W and find constant $c > 0$ s.t. $q_W(w)/p_W(w) \leqslant c$ for all w; then sample $w \sim p_W$ and accept with probability $q_W(w)/cp_W(w)$

$\rightarrow$ need at least 2 $\mathcal{U}(0,1)$ for each sampled W

- When using AR, there is no a-priori bound on how many uniforms are needed: If T_{AR} denotes the AR transformation, we can write

$$\mu = \mathbb{E}(f(\mathbf{U})) = \mathbb{E}(\Psi_0(T_{\mathbf{Y}}(\mathbf{U}_{1:d}), T_{AR}(\mathbf{U}_{(d+1):\infty}))$$

  with $\mathbf{U} \sim U(0,1)^\infty$
- The integrand f is a non-monotone and discontinuous function of some of its input uniforms, a result from the acceptance decision. This can diminish the variance reduction effect of RQMC over MC.
- In previous work combining AR and (R)QMC, formulation as infinite-dimensional integral has been avoided by **rejecting sample points** where first sampled W was not accepted: **AR-n** method

Schematic description of AR-n

i	Sample **Y**				F^{-1}	AR
1	$u_{1,1}$	$u_{1,2}$	$\dots$	$u_{1,d}$	$u_{1,d+1}$	$u_{1,d+2}$
2	$u_{2,1}$	$u_{2,2}$	$\dots$	$u_{2,d}$	$u_{2,d+1}$	$u_{2,d+2}$
3	$u_{3,1}$	$u_{3,2}$	$\dots$	$u_{3,d}$	$u_{3,d+1}$	$u_{3,d+2}$
4	$u_{4,1}$	$u_{4,2}$	$\dots$	$u_{4,d}$	$u_{4,d+1}$	$u_{4,d+2}$
5	$u_{5,1}$	$u_{5,2}$	$\dots$	$u_{5,d}$	$u_{5,d+1}$	$u_{5,d+2}$
6	$u_{6,1}$	$u_{6,2}$	$\dots$	$u_{6,d}$	$u_{6,d+1}$	$u_{6,d+2}$

Schematic description of AR-d

i	Sample $\mathbf{Y}$				F^{-1}	AR	F^{-1}	AR	F^{-1}	AR	...
1	$u_{1,1}$	$u_{1,2}$	...	$u_{1,d}$	$u_{1,d+1}$	$u_{1,d+2}$	$u_{1,d+3}$	$u_{1,d+4}$	$u_{1,d+5}$	$u_{1,d+6}$	
2	$u_{2,1}$	$u_{2,2}$	...	$u_{2,d}$	$u_{2,d+1}$	$u_{2,d+2}$					
3	$u_{3,1}$	$u_{3,2}$	...	$u_{3,d}$	$u_{3,d+1}$	$u_{3,d+2}$					
4	$u_{4,1}$	$u_{4,2}$	...	$u_{4,d}$	$u_{4,d+1}$	$u_{4,d+2}$	$u_{4,d+3}$	$u_{4,d+4}$			

- A potential advantage of AR-d over AR-n for numerical integration is that it only uses the first n points of the LDS rather than a subset of the first $N > n$ points in the sequence.
- **Proposition**: Both methods produce an unbiased estimator
- **Idea**: to make the AR-d output monotone in coordinate $d + 1$ of the underlying LDS, use the rank transformations from our black-box setting: i.e., we re-order the outputs $W_1, \ldots, W_n$ so that their order matches the ordering of $u_{1,d+1}, \ldots, u_{n,d+1}$.

Example: Basket option pricing

- **Goal:** estimate value of a Basket call option with strike K, given by

$$\mu_{\text{bskt}} = e^{-r}\mathbb{E}\left(\max\left\{\frac{1}{d}\sum_{j=1}^{d} S_j - K, 0\right\}\right);$$

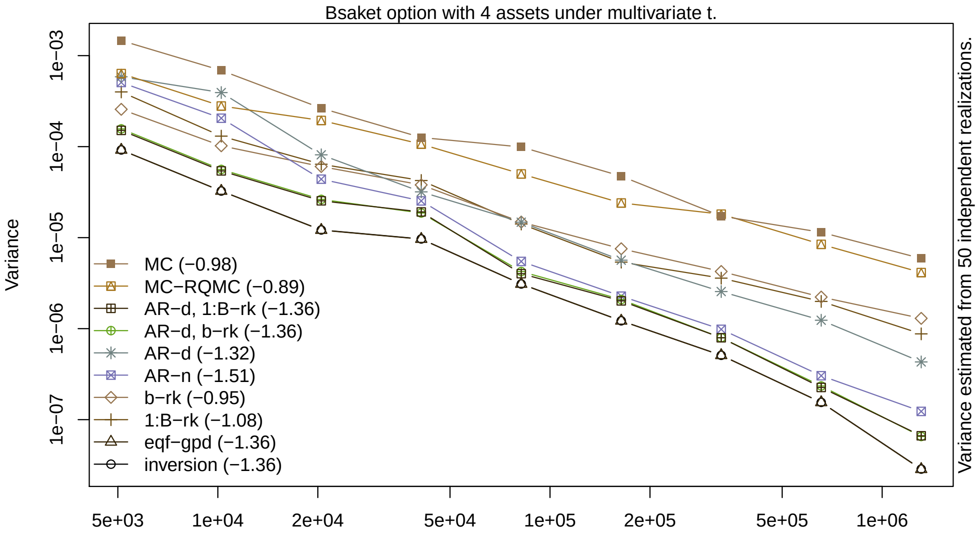
- First, assume that the dependence of the log-normal assets S_j, $j = 1, \ldots, d$, is modelled via a t-copula $\Rightarrow S_j$ have stochastic representation

$$S_j = F_{\text{LN}}^{-1}(U_j), \quad U_j = F_{t_\nu}(X_j), \quad j = 1, \ldots, d, \quad \mathbf{X} \sim t_d(\nu, \mathbf{0}, \Sigma);$$

where $X = \mu + \sqrt{W}Y$ with $W \sim \text{IG}(\nu/2, \nu/2)$ (inverse gamma) independent of $Y \sim N_d(0, \Sigma)$ and where Σ is a correlation matrix.
- RQMC methods will be based on digitally shifted Sobol' sequence.

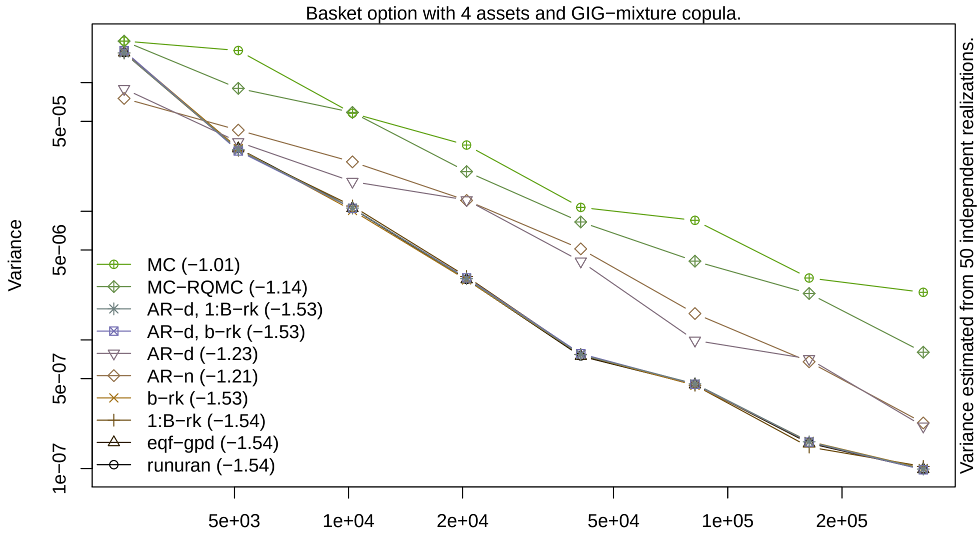
- `MC`: Use MC for W and **Y**;
- `MC-RQMC`: use MC for W and RQMC for **Y**,
- `AR-d`: sample W based on AR whilst moving along the coordinates of a point
- `AR-n`: sample W based on AR whilst moving along the index of the point in the LDS
- `AR-d, b-rk`: Use AR-d in each repetition b and additionally reorder the n samples $W_{1,b}, \ldots, W_{n,b}$ according to $u_{1,b,d+1}, \ldots, u_{n,b,d+1}$ for $b = 1, \ldots, B$.
- `b-rk`: Treat R_W as black-box and compute $\hat{\mu}_{b,n}^{\text{b-rk}}$
- `inversion`: inversion-based estimator (if available)
- `AR, 1:B-rk`: Use AR-d and sort all the sample $W_{1,1}, \ldots, W_{n,B}$ according to $u_{1,1,d+1}, \ldots, u_{n,B,d+1}$.
- `1:B-rk`: Treat R_W as black-box and compute $\hat{\mu}_{b,n}^{\text{1:B-rk}}$.
- `eqf-gpd`: First, build gpd based estimate $\hat{Q}$ using samples obtained from the black box R_W, then treat it as true Q and proceed with inversion.

Results with t copula ($Y \sim N_d(0, \Sigma)$ and $W \sim IG(\nu/2, \nu/2)$)



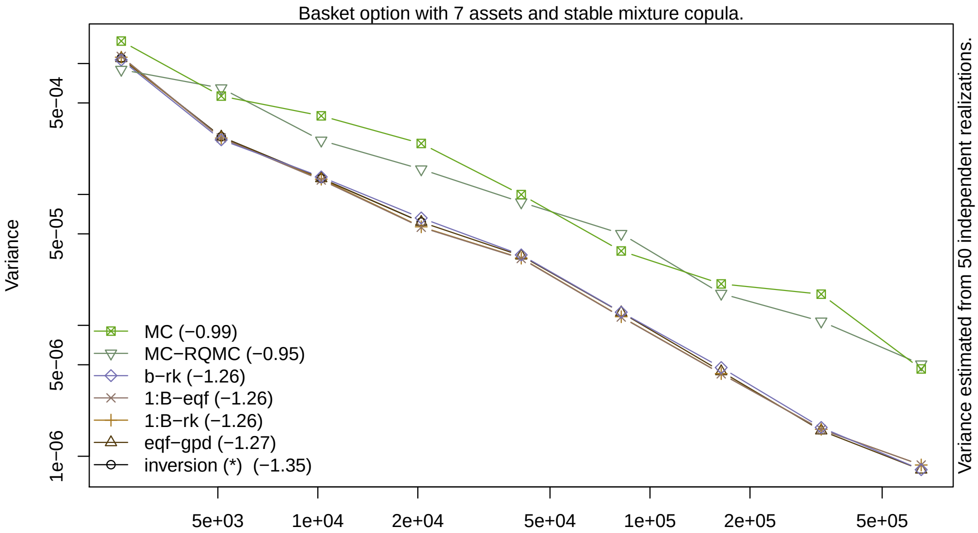
Results with GIG copula

W is now a generalized inverse gamma distribution (corresponding Q not known)



Example with stable mixture copula

W now has a stable distribution: no quantile fct and no known AR algorithm



Run time information

	inversion	eqf-gpd	1:B-rk	b-rk	MC-RQMC	MC
CPU	23.8	4.4	0.4	0.4	0.4	0.4
REff	0.06	0.33	3.64	3.59	1.26	1.00

Average run times in seconds (top) and estimated efficiencies (bottom) when computing various estimators with total sample size $n = 20 \times 2^{10}$ to estimate μ_{bskt} under a 7-dimensional stable mixture copula with $\alpha = 0.9$, $\beta = 1$ and $\gamma = 1$.

Conclusion

- Have shown how to apply (r)QMC successfully on problems going beyond the usual/easy setup of independence and inversion.
- Monotonicity seems important to preserve when working with QMC
- Cases where we have more than one variable to sample by methods other than inversion will need additional work, especially if not independent