QMC for unbounded integrands

Art B. Owen
Stanford University
owen@stat.stanford.edu

Setting for QMC

$$I = \int_{(0,1)^d} f(x)dx$$

$$\hat{I} = \frac{1}{n}\sum_{i=1}^{n} f(x_i)$$

MC	x_i random	and independent
QMC	x_i deterministic	and low discrepancy
RQMC	x_i random	and low discrepancy

Koksma-Hlawka inequality

$$
\begin{aligned}
|\hat{I} - I| &\le D_n^*(x_1, \dots, x_n) \times \|f\|_{\text{HK}} \\
D_n^*(\cdot) &= \text{star discrepancy} \\
\|\cdot\|_{\text{HK}} &= \text{total variation, Hardy \& Krause}
\end{aligned}
$$

Rates and bounds

$D_n^*(\cdot) = O(n^{-1}\log(n)^{d-1})$ attainable

Then $|\hat{I} - I| = O(n^{-1+\epsilon})$

Beating MC $\text{RMSE} = O(n^{-1/2})$

Unless:

$$\|f\|_{\text{HK}} = \infty \quad \Longrightarrow \quad |\hat{I} - I| \le \infty.$$

Infinite HK variation

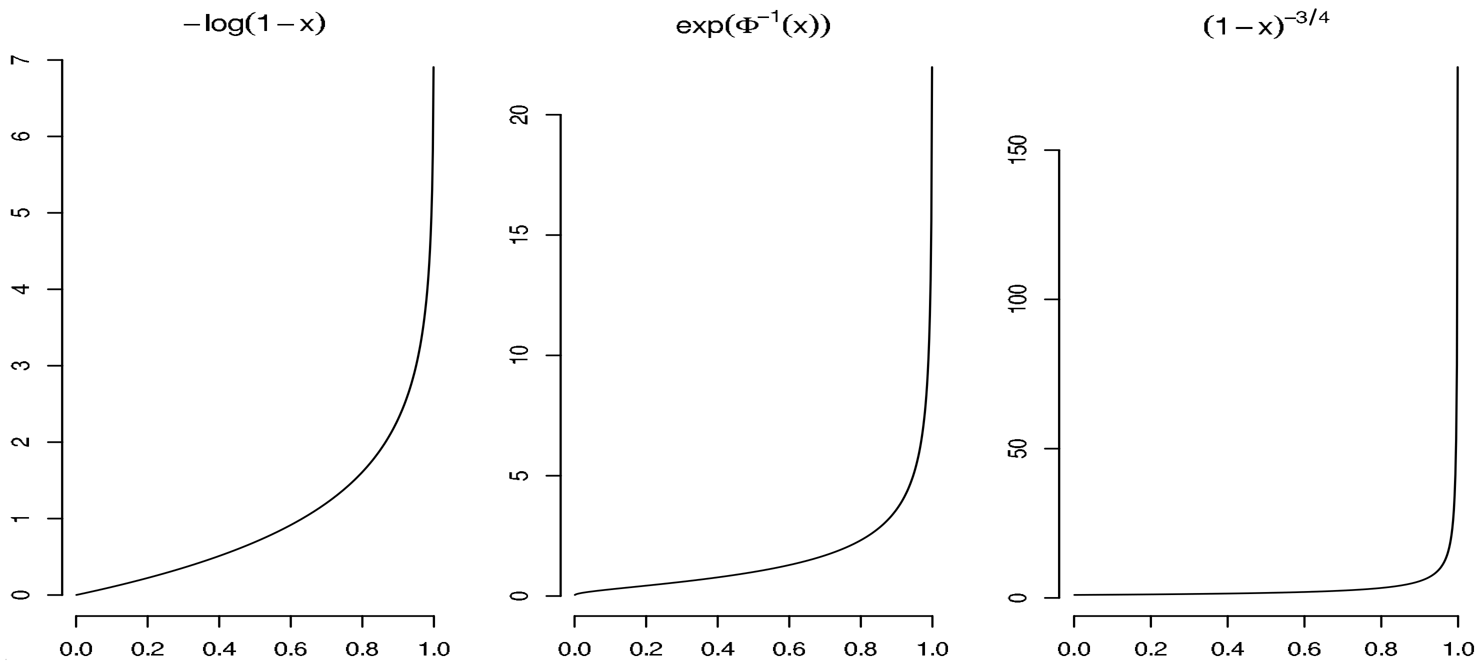
Examples

- Indicator (characteristic) functions of sets $d \geq 2$
- Piecewise functions, e.g. $\max(g_1(x), g_2(x))$ $d \geq 3$
- unbounded functions

Sources of unbounded functions

- Feynmann diagrams (Kinoshita) in physics
- Unbounded domains mapped to $(0,1)^d$
  e.g. $f(\Phi^{-1}(x))$ for Gaussian CDF $\Phi(\cdot)$
- Importance sampling
  e.g. $\int f(x)q(x)dx = \int f(x)\frac{q(x)}{p(x)}p(x)dx$

Example unbounded integrands



Classical techniques

Avoid the singularity

Midpoint rules $1/n, \ldots, (n-1)/n$

Other "open" rules Davis & Rabinowitz

Monte Carlo

Subtract the singularity

$\widetilde{f}(x) = f(x) - s(x)$

$\widetilde{f}(x)$ bounded and $\int s(x)dx$ known

Use $\int s(x)dx$ plus quadrature on $\widetilde{f}$

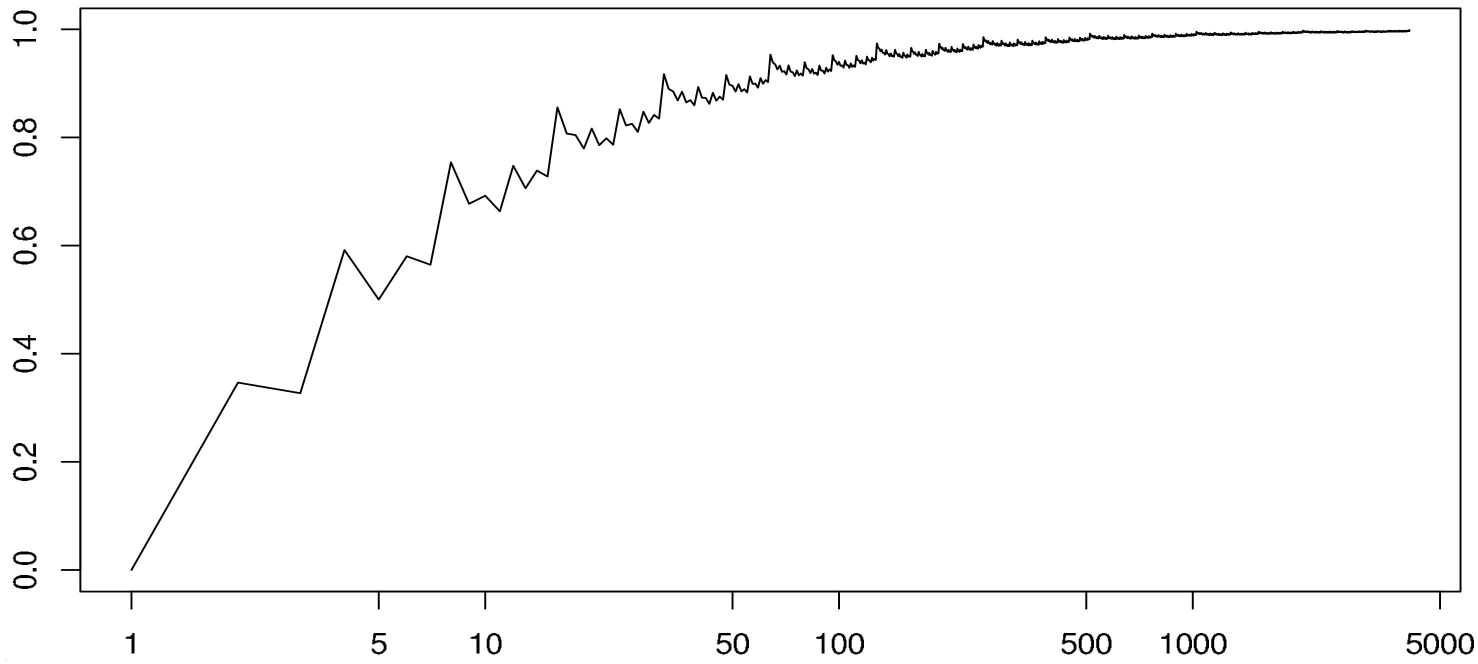
Avoidance is usually simpler

van der Corput points

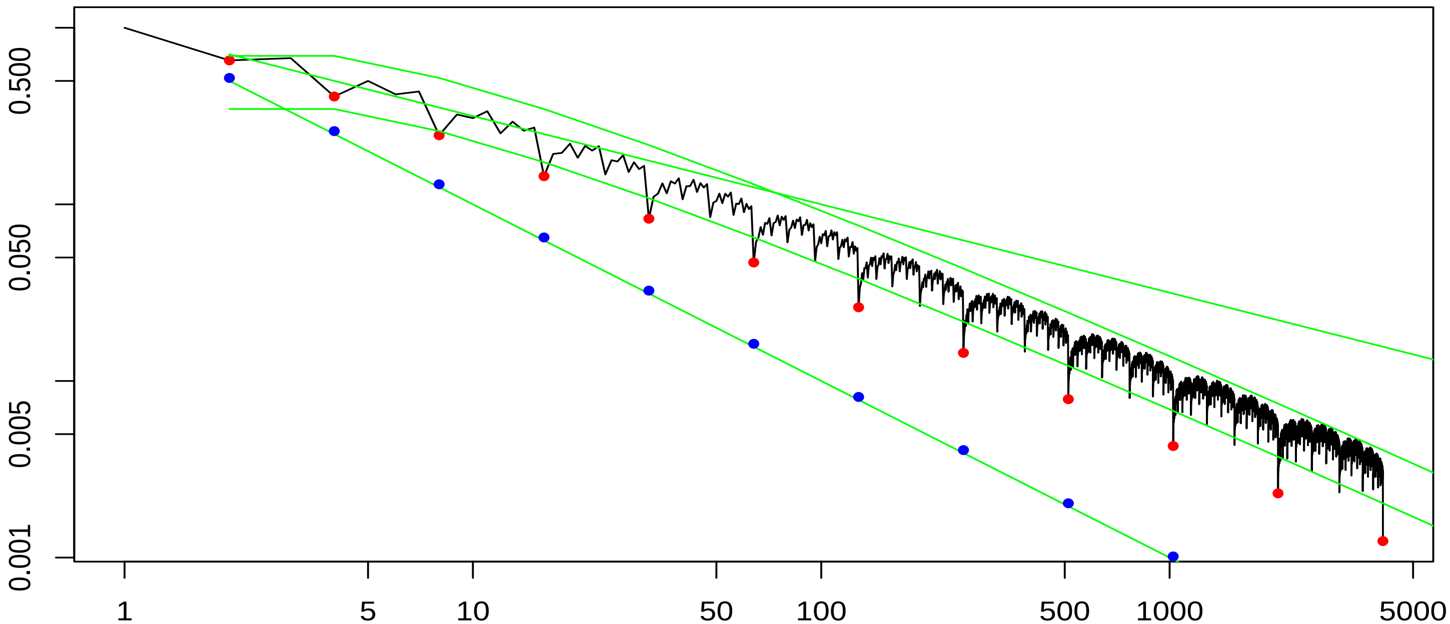
$$x_0 = 0 \quad x_1 = 1/2 \quad x_2 = 1/4 \quad x_3 = 3/4 \quad x_4 = 1/8 \quad x_5 = 5/8 \quad \cdots$$

Avoid singularity at $x = 1$

For $f(x) = -\log(1-x)$, looks like $\hat{I}_n \to I = 1$



Error vs n



— $n^{-1/2}$ n^{-1} $\log(n)/n$ $\log(n)^2/n$

● marks $n = 2^m$

● RQMC RMSE

van der Corput and $-\log(1-x)$

For $n = 2^m$

x_i are	$0/n \quad 1/n \quad 2/n \cdots (n-1)/n$				Left endpoint rule
$\int_0^{1-1/n} f(x)dx$	$\le$	$\hat{I}_n$	$\le$	$\int_0^1 f(x)dx$	by monotonicity
$1 - 1/n - \log(n)/n$	$\le$	$\hat{I}_n$	$\le$	1	calculus
so $\lvert\hat{I}_n - I\rvert = O(\log(n)/n) = O(n^{-1+\epsilon})$					for $n = 2^m$

For general n

Consider $n = 100 = 1100100_{(2)}$

$x_1, \ldots, x_{64}$	a left endpoint rule
Next 32 points	shifted left endpoint rule
Final 4 points	shifted left endpoint rule
Generally:	one rule for every binary 1 in n
At most:	$\log_2(n+1)$ shifted rules

Ultimately:

$$1 \quad \geq \quad \hat{I}_n \quad \geq \quad 1 - \log_2(n+1)\left(\frac{1}{n} + \frac{\log(n)}{n}\right)$$

$$|\hat{I}_n - I| \quad = \quad O(\log(n)^2/n) \quad = \quad O(n^{-1+\epsilon})$$

Randomized van der Corput

For $\mathbf{n = 2^m}$

A scrambled $(0, m, 1)$-net

Don't get RMSE $= O(n^{-3/2}) \quad \cdots \quad$ because $\int f'(x)^2 dx = \infty$

$\max_i f(x_i)$ has variance 1 (indep of n)

k'th largest $f(x_i)$ has variance $1/k^2 + O(1/k^3)$ (indep of n)

$E((\hat{I}_n - I)^2) \doteq 1.0803/n^2$ RMSE $\doteq 1.0394/n$

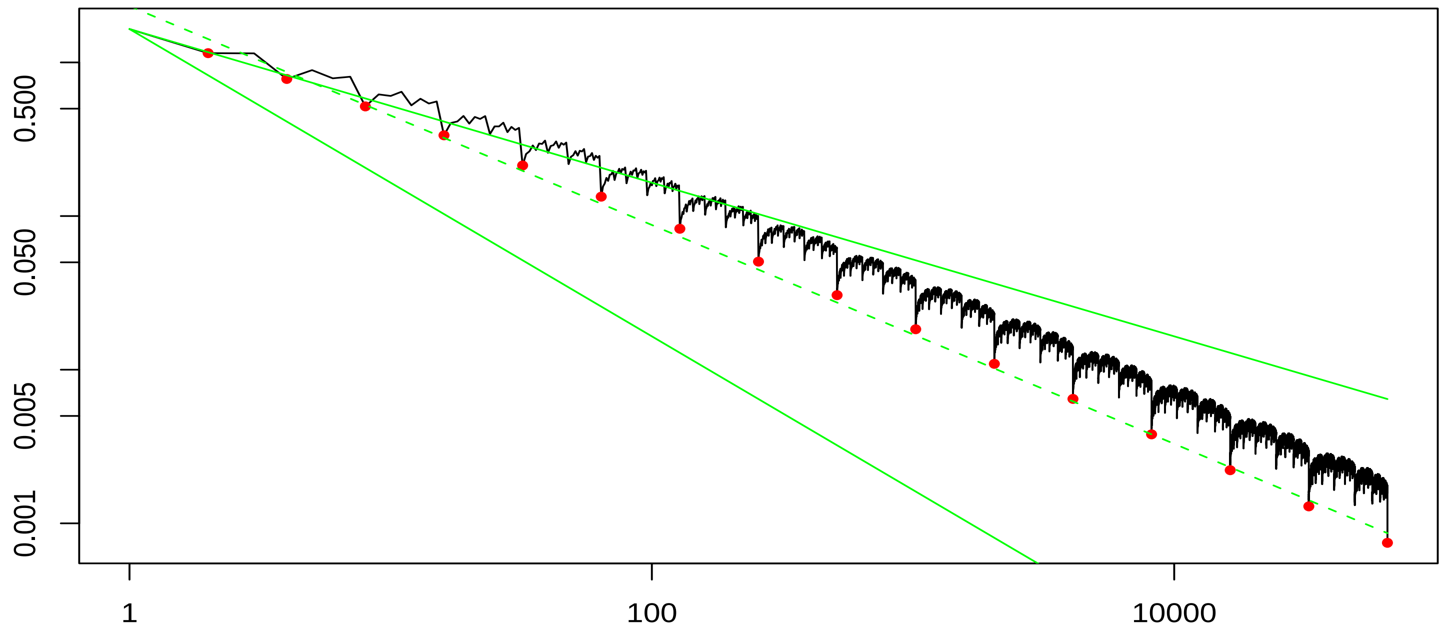
For general $\mathbf{n}$

RMSE $= O(\log(n)/n)$

scrambling only improves by $O(\log(n))$

Log normal mean

$$\varphi(x) = \frac{e^{-x^2/2}}{\sqrt{2\pi}} \qquad \Phi(x) = \int_{-\infty}^{x} \varphi(z)dz \qquad f(x) = e^{\Phi^{-1}(x)} \qquad I = e^{1/2} \doteq 1.654$$



Also $O(n^{-1+\epsilon})$ (Eventually) Dashed: $\propto n^{-0.7122}$

Reference lines: $n^{-1/2}$ n^{-1} $\log(n)/n$ ● marks $n = 2^m$

Sobol' (1973)

Remarkable paper Covers $d = 1$ case and $d \geq 1$

For $d = 1$ & singularity at 0

Define $c_n = \min_{1 \leq i \leq n} x_i$

$$D_n^* \times \int_{c_n}^1 |f'(x)|dx \to 0 \implies |\hat{I}_n - I| \to 0$$

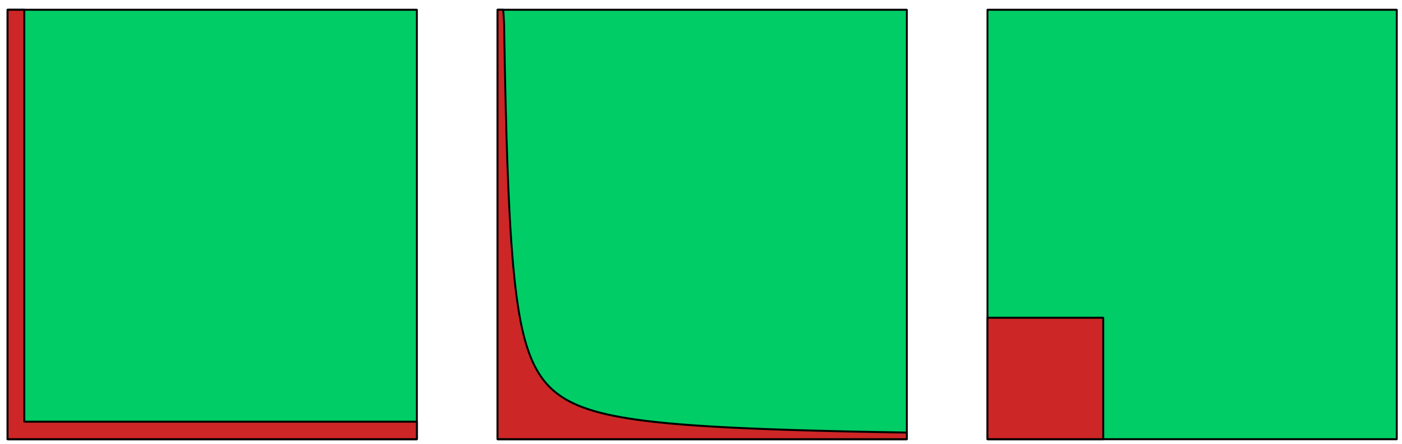
For van der Corput (starting at $x_1 = 1/2$): $\frac{1/2}{n+1} \leq c_n \leq \frac{2}{n+1}$

For $f(x) = x^{-\beta}$ $|\hat{I} - I| = O(n^{\beta-1}\log(n))$

For $f(x) = x^{-1}\log(x)^{-\gamma}$ $|\hat{I} - I| = O(\log(n)^{1-\gamma})$

Avoiding the origin in $(0,1)^d$

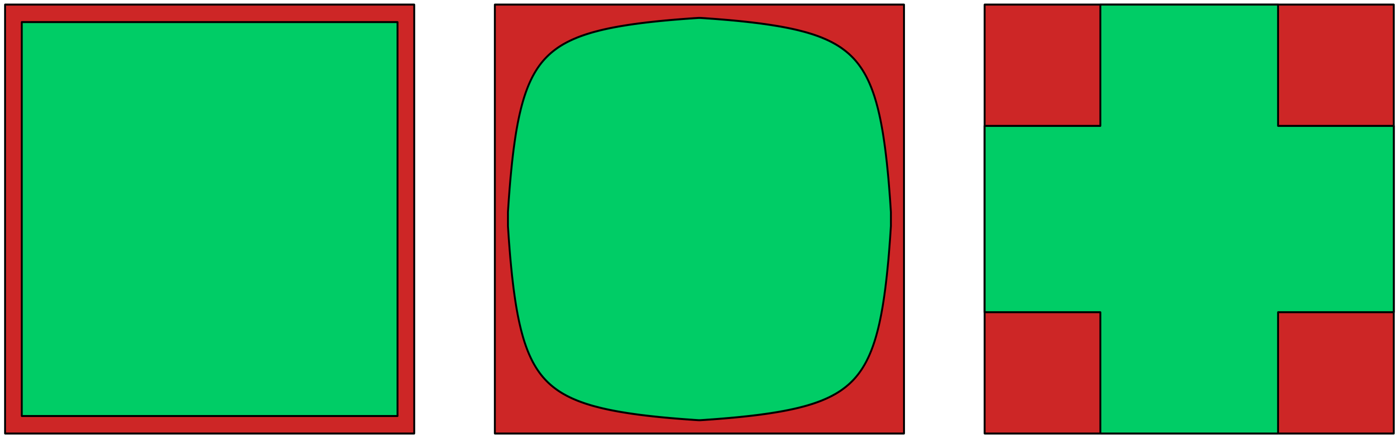
L $\quad$ H $\quad$ C



$K_{\text{orig}}^{\text{min}}(\epsilon)$ $\quad$ $K_{\text{orig}}^{\text{prod}}(\epsilon)$ $\quad$ $K_{\text{orig}}^{\text{max}}(\epsilon)$

$$\min_i \min_j x_{ij} \geq \epsilon \qquad \min_i \prod_j x_{ij} \geq \epsilon \qquad \min_i \max_j x_{ij} \geq \epsilon$$

Avoiding all corners in $(0,1)^d$



$K_{\text{corn}}^{\text{min}}(\epsilon)$ $K_{\text{corn}}^{\text{prod}}(\epsilon)$ $K_{\text{corn}}^{\text{max}}(\epsilon)$

$$\min_i \min_j z_{ij} \ge \epsilon \qquad \min_i \prod_j z_{ij} \ge \epsilon \qquad \min_i \max_j z_{ij} \ge \epsilon$$

Where $z_{ij} = \min(x_{ij}, 1 - x_{ij})$

Sobol' (1973) $d \geq 1$

Uses hyperbolic origin avoidance (typo makes it look like L avoidance)

$$K_n = K_{\text{orig}}^{\text{prod}}(\epsilon_n) \qquad \epsilon_n = \min_{1\leq i\leq n} \prod_{j=1}^{d} x_{ij}$$

Theorem 2

If

$$D_n^* \times \int_{K_n} \left|\frac{\partial^d f(x)}{\partial x}\right| dx \ \rightarrow\ 0,$$

and

$$\int_{(0,1)^d} x_1 x_2 \cdots x_d \left|\frac{\partial^d f(x)}{\partial x}\right| dx \ <\ \infty,$$

and similarly for $2^d - 1$ coordinate projections,

then $|\hat{I} - I| \rightarrow 0$

Sobol' (1973) continued

If $x_1, \dots, x_{2^m}$ a (τ, m, d)–net (from LP$_\tau$-sequence)

Then $x_i \in K_{\text{orig}}^{\text{prod}}(2^{-d-\tau}/n)$

For $$f(x_0) = x_{01}^{-\beta_1} x_{02}^{-\beta_2} \dots x_{0d}^{-\beta_d} \quad x_0 \in (0,1)^d$$

$$|\hat{I} - I| \to 0$$

if all $\beta_j < 1$

After Sobol' (1973)

Hickernell & Sloan & Wasilkowski

Thorough study of tractability over unbounded domains

Singularities removed by transformation

See also Mathé & Wei and Genz & Monahan

Klinger article and dissertation

Among $x_1, \dots, x_n$ two Halton points cannot be in the same small box

Same for digital nets with $t = 0$

After deleting point at origin, remainder are in $K_{\text{orig}}^{\max}$

Kronecker points $x_{ij} = \sqrt{p_j} i \mod 1$ are in $K_{\text{orig}}^{\text{prod}}(\gamma n^{-1-\delta})$

Hartinger & Kainhofer & Tichy

Generalize to $\int h(x) f(x) dx$ with f singular at origin and h non-uniform

Sample with low h discrepancy, adapting Hlawka & Mück

De Doncker & Guan

Accelerate convergence on singularities

Three epsilon approach

First extend $f \cdots$ details to follow

$$f_\epsilon(x) = \widetilde{f}(x) = \begin{cases} f(x), & x \in K(\epsilon) \\ ???, & x \notin K(\epsilon) \end{cases}$$

Then

$$\begin{aligned} |\hat{I} - I| &\leq |\hat{I} - \hat{I}_\epsilon| + |\hat{I}_\epsilon - I_\epsilon| + |I_\epsilon - I| \\ I_\epsilon &= \int f_\epsilon(x)dx \\ \hat{I}_\epsilon &= \frac{1}{n}\sum_{i=1}^n f_\epsilon(x_i) \end{aligned}$$

Analysis

Suppose $x_1, \ldots, x_n \in K(\epsilon_n)$ with $\epsilon = \epsilon_n \to 0$

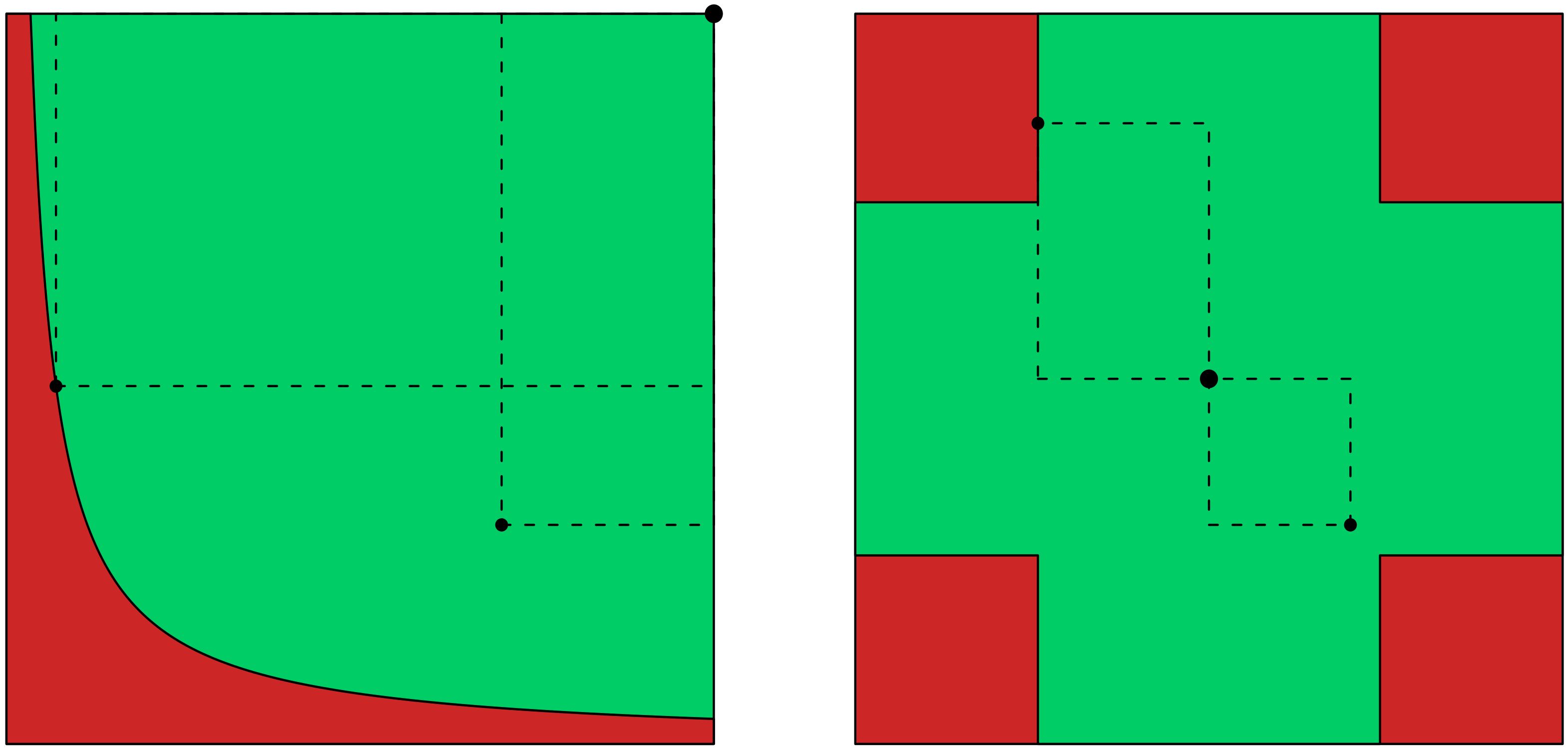
$$\begin{aligned}
\hat{I} - \hat{I}_\epsilon &= \frac{1}{n}\sum_{i=1}^{n} f(x_i) - f_\epsilon(x_i) \quad = \quad 0 \\
|I - I_\epsilon| &\leq \int_{K^c} |f(x) - f_\epsilon(x)|dx \\
|\hat{I}_\epsilon - I_\epsilon| &\leq D_n^* \times \|f_\epsilon\|_{\text{HK}}
\end{aligned}$$

So that:

$$\begin{aligned}
|\hat{I} - I| &\leq \int_{K^c} |f(x) - f_\epsilon(x)|dx + D_n^* \times \|f_\epsilon\|_{\text{HK}} \\
&= \text{Truncation error} \quad + \quad \text{Quadrature error}
\end{aligned}$$

We need $f_\epsilon \approx f$ with small variation

Extendable regions



Region K with anchor $c \in K$

And $x \in K \implies \text{Rect}[x, c] \subseteq K$

Sobol' extension

Assume $\partial^d f/\partial x$ exists (extendable function)

For $d=1$ $\quad f(z) = f(c) + \int_c^z f'(x)dx$ with $\int\limits_c^z = -\int\limits_z^c$

Then $\quad \widetilde{f}(z) = f(c) + \int_c^z f'(x)\, 1_{x\in K}\, dx$

For $d=2$
$$\widetilde{f}(z) = f(c) + \int_{c_1}^{z_1} \frac{\partial f(x_1,c_2)}{\partial x_1} 1_{(x_1,c_2)\in K}\, dx_1 + \int_{c_2}^{z_2} \frac{\partial f(c_1,x_2)}{\partial x_2} 1_{(c_1,x_2)}$$
$$+ \int_{c_1}^{z_1}\int_{c_2}^{z_2} \frac{\partial^2 f(x_1,x_2)}{\partial x_1 \partial x_2} 1_{x\in K}\, dx_2\, dx_1$$

For $d\geq 1$
$$\widetilde{f}(z) = \sum_{u\subseteq\{1,\dots,d\}} \int_{[c_{-u},z_{-u}]} \frac{\partial^{|u|} f(x_{-u}:c_u)}{\partial x_u} 1_{x_{-u}:c_u\in K}\, dx_{-u}$$

Vitali variation of $\widetilde{f}$

$$\frac{\partial^d f(x)}{\partial x} = 0, \quad x \notin K, \quad \text{and,}$$

$$\|\widetilde{f}\|_{\text{Vitali}} \le \int_K \left| \frac{\partial^d f(x)}{\partial x} \right| dx$$

Could hardly be lower!

Sobol' (1973) must have known (proof not published)

Hardy & Krause variation of $\widetilde{f}$

If $c = (1, \dots, 1)$ use $2^d - 1$ "slices" $f(x_u : c_{-u})$

Else chop $(0,1)^d$ into to 2^d subcubes

Halton sequences

Radical inversion

$$\begin{aligned} i &= \cdots a_3 a_2 a_2 a_1 \quad \text{base } b \\ \phi_b(i) &= 0.a_1 a_2 a_3 \cdots \quad \text{base } b \end{aligned}$$

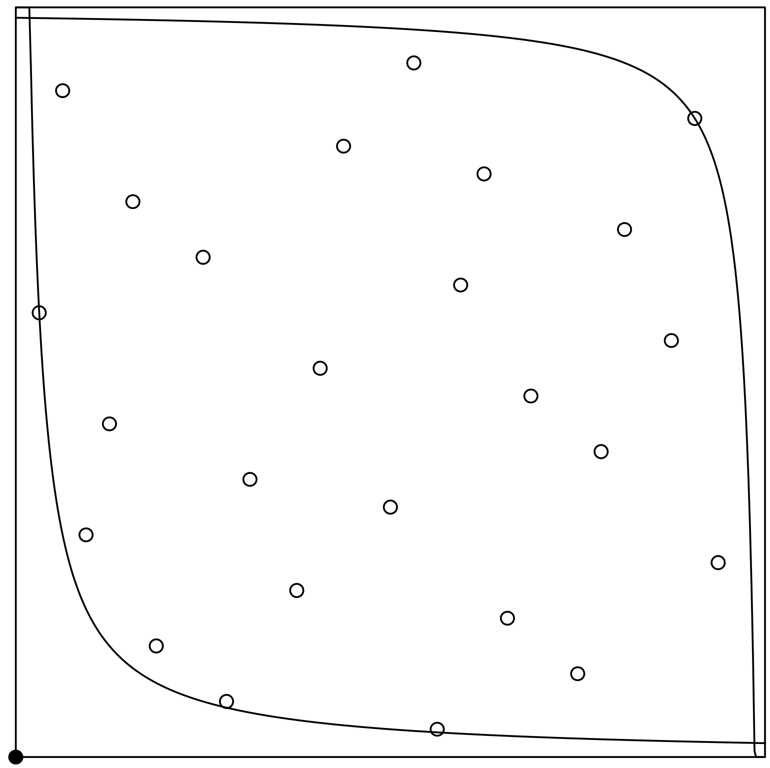
van der Corput

$$x_i = \phi_2(i)$$

Halton

$$\begin{aligned} x_i &= (\phi_{p_1}(i), \phi_{p_2}(i), \ldots, \phi_{p_d}(i)) \\ p_j & \quad \text{relatively prime, usually first } d \text{ primes} \end{aligned}$$

First few Halton points



Appear to avoid origin, if we start at x_1 not $x_0 = (0, 0)$

But there are other large gaps too

Theorem

Let $x_n = (\phi_{p_1}(n), \ldots, \phi_{p_d}(n))$

Where p_j are distinct primes

Then:

$$\prod_{j=1}^{d} x_{nj} \geq \frac{A}{n}$$

$$\prod_{j=1}^{d} (1 - x_{nj}) \geq \frac{A}{n+1}$$

$$\prod_{j=1}^{d} \min(x_{nj}, (1 - x_{nj})) \geq \frac{A}{n(n+1)} \qquad (*)$$

$$\text{where } A = \prod_{j=1}^{d} p_j^{-1}$$

(*) Thanks to Reinhhold Kainhofer for improving the constant

Idea of proof

$$\phi_b(i) \approx 0 \quad \Longleftrightarrow \quad \phi_b(i) \text{ has many leading } 0\text{'s} \quad (\text{base } b)$$
$$\Longleftrightarrow \quad i \text{ has many trailing } 0\text{'s}$$

But i cannot be a power of two primes (if $i > 1$)

Similarly

$$\phi_b(i) \approx 1 \quad \Longleftrightarrow \quad \phi_b(i) \text{ has many leading } b-1\text{'s} \quad (\text{base } b)$$
$$\Longleftrightarrow \quad i \text{ has many trailing } b-1\text{'s}$$
$$\Longleftrightarrow \quad i+1 \text{ has many trailing } 0\text{'s}$$

$i + 1$ cannot be a power of two primes either

Proof

List primes $P_1 < P_2 < P_3 < \cdots$ put $p_j = P_{r(j)}$

First part

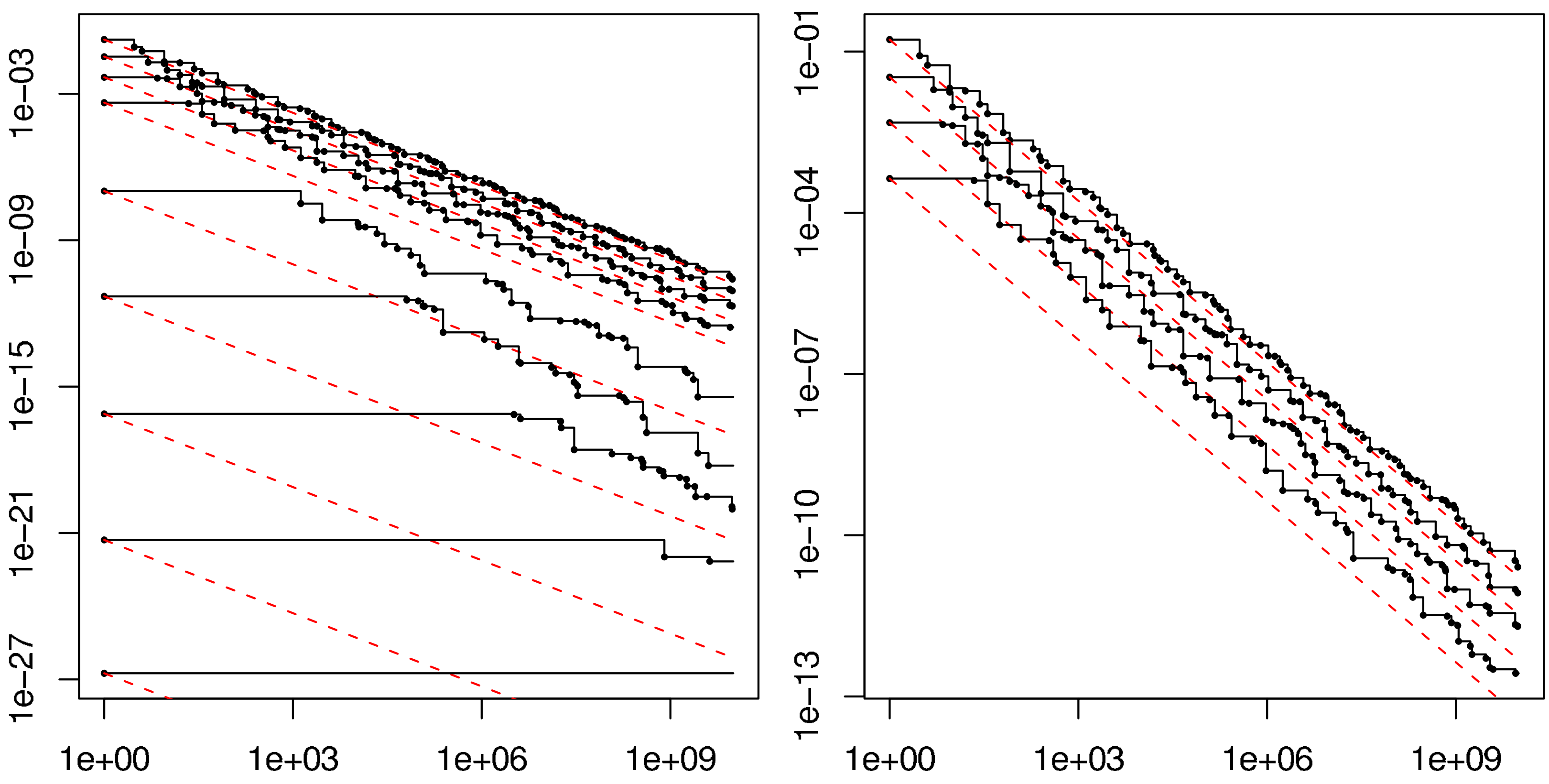
$$\begin{aligned} n &= \prod_r P_r^{a_r} \qquad a_r \text{ trailing zeros in base } P_r \\ \phi_{P_r}(n) &\geq P_r^{-a_r-1} \end{aligned}$$

$$\prod_{j=1}^d x_{nj} \geq \prod_{j=1}^d p_j^{-a_{r(j)}-1} \geq \prod_{j=1}^d p_j^{-1} \prod_r P_r^{-a_r} = \frac{A}{n}$$

For the rest

$$n+1 = \prod_r P_r^{b_r} \quad \text{and} \quad 1 - \phi_{P_r}(n) \geq P_r^{-b_r-1} \quad \cdots$$

Hyperbolic origin avoidance



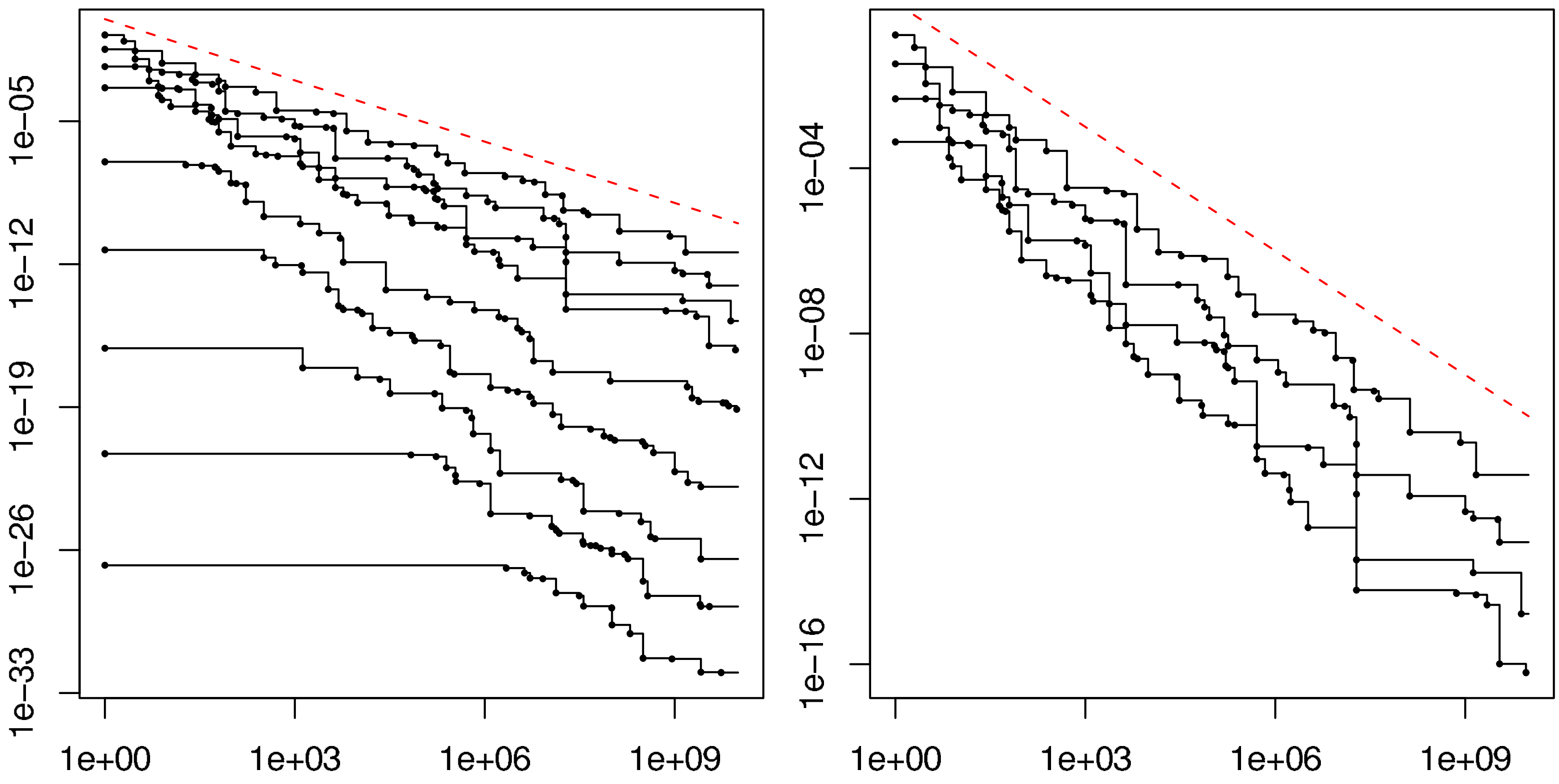
$$\min_{1\le i\le n}\prod_{j=1}^{d} x_{ij}$$ vs n

$$1 \le n \le 10^{10}$$

$$d = 2, 3, 4, 5, 8, 11, 14, 17, 20$$

Reference lines are lower bounds

Hyperbolic corner avoidance



$$\min_{1\le i\le n}\prod_{j=1}^{d}\min(x_{ij},1-x_{ij})\quad \text{vs}\quad n$$

$$d=2,3,4,5,8,11,14,17,20 \qquad - - - \quad n^{-1}$$

Halton corner avoidance

$$D_{\text{corn}}^{n,d} = \min_{1\le i\le n} \prod_{j=1}^{d} \min(x_{ij}, 1-x_{ij})$$

$$D_{\text{corn}}^{n,d} = O(n^{-r_d}) \quad 1 \le r_d \le 2$$

Apparent rate of $D_{\text{orig}}^{n,d}$

$d \in 2:5$		1.10	1.24	1.34	1.39
$d \in 6:10$	1.60	1.57	1.57	1.59	1.63
$d \in 11:15$	1.58	1.51	1.69	1.66	1.75
$d \in 16:20$	1.47	1.62	1.52	1.79	1.82

Growth conditions

Singularity at origin

$$|f(x)| \leq B\prod_{j=1}^{d} x_j^{-A_j} \quad \text{and,}$$

$$|\partial^u f(x)| \leq B\left(\prod_{j\in u} x_j^{-A_j-1}\right)\left(\prod_{j\notin u} x_j^{-A_j}\right)$$

Singularity at corners

$$|f(x)| \leq B\prod_{j=1}^{d} \min(x_j, 1-x_j)^{-A_j} \quad \text{and,}$$

$$|\partial^u f(x)| \leq B\left(\prod_{j\in u} \min(x_j, 1-x_j)^{-A_j-1}\right)\left(\prod_{j\notin u} \min(x_j, 1-x_j)^{-A_j}\right)$$

For $0 < A_j < 1$ and $0 < B < \infty$

Avoiding L shaped regions

Theorem

If $\quad x_1, \ldots, x_n \in K^{\text{min}}_{\text{orig}}(Cn^{-r}) \qquad$ respectively $K^{\text{min}}_{\text{corn}}(Cn^{-r})$

and $\quad f$ obeys origin (corner) growth conditions,

then $\quad |\hat{I} - I| \le C_1 D_n^* \, n^{r\sum_{j=1}^d A_j} + C_2 n^{r(\max_j A_j - 1)}$

Corollary: when $D_n^* = O(n^{-1+\epsilon})$,

$$|\hat{I} - I| = O(n^{-1+r\sum_j A_j})$$

How to avoid L shaped regions

Halton points do (after x_0) with $r = 1$

Linear adjustment $\cdots$

$$\widetilde{x}_i = \epsilon_n + (1 - \epsilon_n)x_i \qquad \text{to avoid origin}$$
$$\widetilde{x}_i = \epsilon_n + (1 - 2\epsilon_n)x_i \qquad \text{to avoid corners}$$

$\cdots$ preserves low discrepancy

$$D_n^*(\widetilde{x}_1, \ldots, \widetilde{x}_n) = D_n^*(x_1, \ldots, x_n) + O((1+2\epsilon_n)^d - 1)$$ Hlawka & Mück

$$\epsilon_n = Cn^{-1} \quad \text{ok}$$

Comparison to Monte Carlo

Suppose $f(x) = \prod_{j=1}^{d} x_j^{-A_j}$ $\quad \max_j A_j < 1$

QMC with n^{-r} origin avoidance has

$$|\hat{I} - I| = O(n^{-1+r\sum_j A_j})$$

MC accuracy

$$\max_j A_j < 1/2 \implies \int f(x)^2 dx < \infty \implies \text{MC has RMSE } O(n^{-1/2})$$

QMC wins if $\sum_{j=1}^{d} A_j < (2r)^{-1}$ can lose otherwise

Extreme example $\sum_j A_j > 1$

Low discrepancy x_i and $\widetilde{x}_i = \frac{1}{n} + (1 - \frac{2}{n})x_i$

$$x_1 = (0, 0, \dots, 0) \implies \widetilde{x}_1 = (1/n, \dots, 1/n)$$

$$\hat{I} \geq n^{\sum A_j - 1} \to \infty$$

Avoiding hyperbolic regions

Theorem

If $x_1, \dots, x_n \in K^{\text{prod}}_{\text{orig}}(Cn^{-r})$ resp. $K^{\text{prod}}_{\text{corn}}(Cn^{-r})$

and f obeys origin (corner) growth conditions,

then $|\hat{I} - I| \le C_1 D_n^* \, n^{\eta + r \max_j A_j} + C_2 n^{r(\max_j A_j - 1)}$

Any $\eta > 0$ and $\eta = 0$ if $\max A_j$ unique

Corollary: for $D_n^* = O(n^{-1+\epsilon/2})$,

$$|\hat{I} - I| = O(n^{-1+\epsilon + r \max_j A_j})$$

How to avoid hyperbolic regions

Halton points do (after x_0) with $1 \le r \le 2$

Linear adjustment fails

Eg $x_1 = (0, 0, \ldots, 0)$ has to move $O(n^{-r/d})$ to enter $K_{\text{orig}}^{\text{prod}}(n^{-r})$

An $n^{-1/d}$ move affects discrepancy

Proving hyperbolic avoidance

Requires details of x_i construction

axis oriented discrepancy does not suffice

not even isotropic discrepancy ($n^{-1/d}$ again)

Comparison to Monte Carlo

If	x_i in $K^{\text{corn}}_{\text{prod}}(n^{-r})$ with $D^*_n = O(n^{-1+\epsilon})$
and	f obeys growth conditions with $\max_j A_j < 1/2$
then QMC has	$\lvert\hat{I} - I\rvert = O(n^{-1+\epsilon+r\max_j A_j})$
and beats MC	when $\max_j A_j < (2r)^{-1}$

Recall $\sum_j A_j < (2r)^{-1}$ **(for L avoidance)**

H avoidance closer to contours of f

Randomized quasi-Monte Carlo

Each $x_i \sim U(0,1)^d$, $\quad D_n^* \leq Cn^{-1+\eta}$, $\quad \forall \eta > 0$

$$\text{let} \quad f_\epsilon(x) = \begin{cases} f(x), & x \in K(\epsilon_n) \\ \widetilde{f}(x), & x \notin K(\epsilon_n) \end{cases}$$

as before

$$|\hat{I} - I| \leq |\hat{I} - \hat{I}_\epsilon| + |\hat{I}_\epsilon - I_\epsilon| + |I_\epsilon - I|$$

but $\Pr(x_i \notin K) = \text{vol}(K^c) > 0$

$$\begin{aligned} E(|\hat{I} - \hat{I}_\epsilon|) &\leq \frac{1}{n}\sum_{i=1}^{n} |E(f(x_i) - f_\epsilon(x_i))| = |I - I_\epsilon| \\ E(|\hat{I} - I|) &\leq |\hat{I}_\epsilon - I_\epsilon| + 2|I_\epsilon - I| \\ &\leq D_n^* \times \|f_\epsilon\|_{\text{HK}} + 2|I_\epsilon - I| \end{aligned}$$

RQMC

Doubles truncation error

But allows more K choices, e.g. level sets of f (if extendable)

RQMC $E(|\hat{I} - I|) \le \inf_K \Big(D_n^* \times \|f_\epsilon\|_{\mathsf{HK}} + 2|I_\epsilon - I| \Big)$

QMC $|\hat{I} - I| \le \inf_K \Big(D_n^* \times \|f_\epsilon\|_{\mathsf{HK}} + |I_\epsilon - I| \Big)$

For RQMC $K(\epsilon_n)$ must be extendable

For QMC $K(\epsilon_n)$ must be extendable *and* avoided by x_i

Consequences

$E(|\hat{I} - I|) = O(n^{-1+\epsilon+\max_j A_j})$

Then RQMC beats $O(n^{-1/2})$ when $\max_j A_j < 1/2$

Can have $x_1 = (0, \ldots, 0)$ (prerandomization) e.g. lattices

Asian call and put option integrands

Obey growth conditions for any $A_j > 0$ good for (R)QMC

Like power singularity with $A_j = 0+$

but cusp $\Longrightarrow \|f\|_{\mathrm{HK}} = \infty$ even for bounded put

Summary

For unbounded f

1. QMC can have $|\hat{I} - I| = O(n^{-1+\epsilon}) \quad \forall \epsilon > 0$
2. or $\hat{I} \to \infty$
3. requires discrepancy **plus** avoidance

Acknowledgements

I. Sobol'

R. Kainhofer & B. Klinger

H. Niederreiter & D. Talay

M. Bossy & M. Deaconu & E. Tanré

M. Simonetti