Progress in high-dimensional numerical integration and its application to stochastic optimization

W. Römisch

Humboldt-University Berlin
Department of Mathematics

www.math.hu-berlin.de/~romisch

DFG Research Center MATHEON
Mathematics for key technologies

Short course, Rutgers University, May 21–23, 2013

Part II: Quasi-Monte Carlo methods and their recent developments

Contents:

Introduction to Quasi-Monte Carlo methods

We consider the approximate computation of

$$I_d(f) = \int_{[0,1]^d} f(\xi) d\xi$$

by a QMC algorithm

$$Q_{n,d}(f) = \frac{1}{n} \sum_{i=1}^{n} f(\xi^i)$$

with (non-random) points ξ^i, $i = 1, \ldots, n$, from $[0,1]^d$.

We assume that f belongs to a linear normed space $\mathbb{F}_d$ of functions on $[0,1]^d$ with norm $\|\cdot\|_d$ and unit ball $\mathbb{B}_d = \{f \in \mathbb{F}_d : \|f\|_d \leq 1\}$ such that I_d and $Q_{n,d}$ are linear bounded functionals on $\mathbb{F}_d$.

Worst-case (absolute) error of $Q_{n,d}$ over $\mathbb{B}_d$:

$$e(Q_{n,d}) = \sup_{f \in \mathbb{B}_d} \left| I_d(f) - Q_{n,d}(f) \right|$$

An approximation criterion may be based on the relative error and a given tolerance $\varepsilon > 0$, namely, in finding the smallest number $n_{\min}(\varepsilon, Q_{n,d}) \in \mathbb{N}$ such that

$$e(Q_{n,d}) \leq \varepsilon e(Q_{0,d}) = \varepsilon \|I_d\| \quad \text{for all } n \geq n_{\min}(\varepsilon, Q_{n,d}),$$

holds, where $Q_{0,d}(f) = 0$ and, hence, $e(Q_{0,d}) = \|I_d\|$.

The behavior of the error $e(Q_{n,d})$ with respect to $n \in \mathbb{N}$ and of $n_{\min}(\varepsilon, Q_{n,d})$ with respect to ε is of considerable interest. In both cases the dependence on the dimension d is often crucial, too.
The behavior of both quantities depends heavily on the normed space F_d.

It is desirable that an estimate of the form

$$n_{\min}(\varepsilon, Q_{n,d}) \leq C\, d^q \varepsilon^{-p} \qquad \text{('polynomial tractability')}$$

is valid for some constants $q \geq 0$, $C, p > 0$ and for every $\varepsilon \in (0, 1)$. Of course, $q = 0$ is highly desirable for high-dimensional problems.

Example 1:

Consider the Banach space $F_d = \mathrm{Lip}([0,1]^d)$ of Lipschitz continuous functions equipped with the norm

$$\|f\|_d = |f(0)| + \sup_{\xi \neq \tilde{\xi}} \frac{|f(\xi) - f(\tilde{\xi})|}{\|\xi - \tilde{\xi}\|}.$$

The best possible convergence rate is $e(Q_{n,d}) = O(n^{-\frac{1}{d}})$. (Bakhvalov 59).

The unit ball in $\mathrm{Lip}([0,1]^d)$ is too large !

Example 2:

Consider the Banach space $\mathbb{F}_d = C^r([0,1]^d)$ ($r \in \mathbb{N}$) of r times continuously differentiable functions with the norm

$$\|f\|_d = \max_{|\alpha| \leq r} \|f^{(\alpha)}\|_\infty,$$

where $\alpha = (\alpha_1, \ldots, \alpha_d) \in \mathbb{N}_0^d$ is used to denote by $f^{(\alpha)}$ a partial derivative of f of order $|\alpha| = \sum_{i=1}^d \alpha_i$, i.e.,

$$f^{(\alpha)}(\xi) = \frac{\partial^{|\alpha|} f}{\partial \xi_1^{\alpha_1} \cdots \partial \xi_d^{\alpha_d}}(\xi)\,.$$

The best possible convergence rate then is $e(Q_{n,d}) = O(n^{-\frac{r}{d}})$ (Novak 88).

Classical convergence results:

Theorem: (Proinov 88)
If the real function f is continuous on $[0,1]^d$, then there exists $C>0$ such that

$$|Q_{n,d}(f) - I_d(f)| \leq C\omega_f\big(D_n^*(\xi^1,\dots,\xi^n)^{\frac{1}{d}}\big),$$

where $\omega_f(\delta) = \sup\{|f(\xi)-f(\tilde{\xi})| : \|\xi-\tilde{\xi})\| \leq \delta,\ \xi,\tilde{\xi}\in[0,1]^d\}$ is the modulus of continuity of f and

$$D_n^*(\xi^1,\dots,\xi^n) := \sup_{x\in[0,1]^d} |\mathrm{disc}(x)|, \quad \mathrm{disc}(x) = \lambda^d([0,x)) - \frac{1}{n}\sum_{i=1}^n \mathbb{1}_{[0,x)}(\xi^i),$$

is the star-discrepancy of $\xi^1,\dots,\xi^n$ (λ^d denotes Lebesgue's measure on $\mathbb{R}^d$).

Theorem: (Koksma-Hlawka 61)
If $V_{\mathrm{HK}}(f)$ is the variation of f in the sense of Hardy and Krause, it holds

$$|I_d(f) - Q_{n,d}(f)| \leq V_{\mathrm{HK}}(f) D_n^*(\xi^1,\dots,\xi^n)$$

for any $n\in\mathbb{N}$ and any $\xi^1,\dots,\xi^n\in[0,1]^d$.

Example 3:

Consider the linear normed space $F_d = \mathrm{BV}_{\mathrm{HK}}([0,1]^d)$ of functions having bounded variation in the sense of Hardy and Krause equipped with the norm

$$\|f\|_d = |f(1,\dots,1)| + V_{\mathrm{HK}}(f)\,.$$

The Koksma-Hlawka theorem then implies

$$e(Q_{n,d}) \le D_n^*(\xi^1,\dots,\xi^n)$$

However, the variation in the sense of Hardy and Krause is a difficult quantity and it is not clear which functions belong to $\mathbb{F}_d$.

Variation of a function in the sense of Hardy and Krause (Owen 05)

Let $D = \{1, \ldots, d\}$ and we consider subsets u of D with cardinality $|u|$. By $-u$ we mean $-u = D \setminus u$. The expression ξ^u denotes the $|u|$-tuple of the components ξ_j, $j \in u$, of $\xi \in \mathbb{R}^d$. For example, we write $f(\xi) = f(\xi^u, \xi^{-u})$.

We consider the d-fold alternating sum of f over a d-dimensional interval $[a, b]$

$$\triangle(f; a, b) = \sum_{u \subseteq D} (-1)^{|u|} f(a^u, b^{-u}) \quad \text{and} \quad \triangle_u(f; a, b) = \sum_{v \subseteq u} (-1)^{|v|} f(a^v, b^{-v}) \quad (u \subseteq D).$$

The variation of f over a finite grid G in $[a, b)$ is (with $g^+ \in [a, b]$ denoting a successor to g)

$$V_G(f) = \sum_{g \in G} |\triangle(f; g, g^+)| .$$

If $\mathcal{G}$ denotes the set of all finite grids in $[a, b)$, the variation of f on $[a, b]$ in Vitali's sense is

$$V_{[a,b]}(f) = \sup_{G \in \mathcal{G}} V_G(f) .$$

The variation of f on $[a, b]$ in the sense of Hardy and Krause is

$$V_{\mathrm{HK}}(f; a, b) = \sum_{u \subset D} V_{[a^{-u}, b^{-u}]}(f(\xi^{-u}, b^u)) .$$

Bounded variation on $[a, b]$ in the sense of Hardy and Krause then means $V_{\mathrm{HK}}(f; a, b) < \infty$.

A first QMC construction

Radical inverse function:

For $i \in \mathbb{N}_0$, $b \in \mathbb{N}$, $b \geq 2$, the radical inverse function $\phi_b(i)$ is defined as follows:

$$\text{if } i = \sum_{k=1}^{\infty} i_k b^{k-1} \text{ with } i_k \in \{0, 1, \ldots, b-1\}, \text{ then } \phi_b(i) := \sum_{k=1}^{\infty} \frac{i_k}{b^k}.$$

Van der Corput sequence:

The sequence $(\phi_b(n))_{n \in \mathbb{N}_0}$ in $[0, 1)$ is called van der Corput sequence in base b.

Halton sequence:

Let p_i, $i = 1, \ldots, d$, be the first d prime numbers. The Halton sequence in d dimensions is given by

$$\xi^{i+1} = (\phi_{p_1}(i), \ldots, \phi_{p_d}(i)) \in [0, 1)^d \quad (i \in \mathbb{N}_0).$$

Theorem: The Halton sequence in d dimensions satisfies the estimate

$$D_n^*(\xi^1, \ldots, \xi^n) \leq C(d) \frac{(\log n)^d}{n}$$

for some constant $C(d)$ depending on d and all $n \in \mathbb{N}$.

It is known that the constant $C(d)$ gets very large even for moderately large d and that the right-hand side of the estimate increases with increasing n for all $n < \exp d$.

The case of kernel reproducing Hilbert spaces (Aronszajn 50)

We assume that $\mathbb{F}_d$ is a kernel reproducing Hilbert space with inner product $\langle \cdot, \cdot \rangle$ and kernel $K : [0,1]^d \times [0,1]^d \to \mathbb{R}$, i.e.,

$$K(\cdot, y) \in \mathbb{F}_d \text{ and } \langle f(\cdot), K(\cdot, y) \rangle = f(y) \quad (\forall y \in [0,1]^d, f \in \mathbb{F}_d).$$

If I_d is a linear bounded functional on $\mathbb{F}_d$, the quadrature error $e_n(Q_{n,d})$ allows the representation

$$e(Q_{n,d}) = \sup_{f \in \mathbb{B}_d} \left| I_d(f) - Q_{n,d}(f) \right| = \sup_{f \in \mathbb{B}_d} |\langle f, h_n \rangle| = \|h_n\|_d$$

according to Riesz' theorem for linear bounded functionals.

The representer $h_n \in \mathbb{F}_d$ of the quadrature error is of the form

$$h_n(x) = \int_{[0,1]^d} K(x,y)dy - \tfrac{1}{n} \sum_{i=1}^{n} K(x, \xi^i) \quad (\forall x \in [0,1]^d),$$

and it holds

$$e^2(Q_{n,d}) = \int_{[0,1]^{2d}} K(x,y)dx\, dy - \tfrac{2}{n} \sum_{i=1}^{n} \int_{[0,1]^d} K(\xi^i, y)dy + \tfrac{1}{n^2} \sum_{i,j=1}^{n} K(\xi^i, \xi^j)$$

(Hickernell 98)

Example: Weighted tensor product Sobolev spaces

$$\mathbb{F}_d = \mathcal{W}^{(1,\dots,1)}_{2,\gamma,\mathrm{mix}}([0,1]^d) = \bigotimes_{j=1}^{d} W^1_{2,\gamma_j}([0,1])$$

equipped with the weighted norm $\|f\|^2_\gamma = \langle f, f\rangle_\gamma$ and inner product

$$\langle f, g\rangle_\gamma = \sum_{u\subseteq\{1,\dots,d\}} \gamma_u^{-1} \int_{[0,1]^{|u|}} \frac{\partial^{|u|} f}{\partial x_u}(x_u, 1)\frac{\partial^{|u|} g}{\partial x_u}(x_u, 1) dx_u\,,$$

where $\gamma_1 \geq \gamma_2 \geq \cdots \geq \gamma_d > 0$, $\gamma_u = \prod_{j\in u} \gamma_j$, is a kernel reproducing Hilbert space with the kernel

$$K_{d,\gamma}(x, y) = \prod_{j=1}^{d} (1 + \gamma_j \mu(x_j, y_j)) \quad (x, y \in [0, 1]^d),$$

where

$$\mu(t, s) = \begin{cases} \min\{|t-1|, |s-1|\} & , (t-1)(s-1) > 0, \\ 0 & , \text{else.} \end{cases}$$

Note that $f \in \mathbb{F}_d$ iff $\frac{\partial^{|u|} f}{\partial x_u}(\cdot, 1) \in L_2([0,1]^{|u|})$ for all $u \subseteq D$.

Theorem: (Sloan-Woźniakowski 98)
Let $\mathbb{F}_d = \mathcal{W}^{(1,\ldots,1)}_{2,\gamma,\mathrm{mix}}([0,1]^d)$. Then the worst-case error

$$e^2(Q_{n,d}) = \sup_{\|f\|_\gamma \le 1} |I_d(f) - Q_{n,d}(f)| = \sum_{\emptyset \neq u \subseteq D} \prod_{j \in u} \gamma_j \int_{[0,1]^{|u|}} \mathrm{disc}^2(x_u, 1) dx_u$$

is the so-called weighted L_2-discrepancy of $\xi^1, \ldots, \xi^n$.

Note that any $f \in \mathbb{F}_d$ is of bounded variation $V_{\mathrm{HK}}(f)$ in the sense of Hardy and Krause and it holds

$$V_{\mathrm{HK}}(f) = \sum_{\emptyset \neq u \subseteq D} \int_{[0,1]^{|u|}} \left| \frac{\partial^{|u|} f}{\partial x_u}(x_u, 1) \right| dx_u \,.$$

Extended (weighted) Koksma-Hlawka inequality:

$$|I_d(f) - Q_{n,d}(f)| \leq \|\mathrm{disc}(\cdot)\|_{\gamma,p,p'} \|f\|_{\gamma,q,q'} \,,$$

where $1 \leq p, p', q, q' \leq \infty$, $\frac{1}{p} + \frac{1}{q} = 1$, $\frac{1}{p'} + \frac{1}{q'} = 1$, and

$$\|\mathrm{disc}(\cdot)\|_{p,p'} = \Big(\sum_{u \subseteq D} \gamma_u \Big(\int_{[0,1]^{|u|}} |\mathrm{disc}(x_u, 1)|^{p'} \, dx_u \Big)^{\frac{p}{p'}} \Big)^{\frac{1}{p}}$$

and

$$\|f\|_{q,q'} = \Big(\sum_{u \subseteq D} \gamma_u^{-1} \Big(\int_{[0,1]^{|u|}} \Big| \frac{\partial^{|u|} f}{\partial x_u}(x_u, 1) \Big|^{q'} dx_u \Big)^{\frac{q}{q'}} \Big)^{\frac{1}{q}}$$

with the obvious modifications if one or more of p, p', q, q' are infinite.
In particular, the classical Koksma-Hlawka inequality essentially corresponds to $p = p' = \infty$ if f belongs to the tensor product Sobolev space $\mathcal{W}_{2,\mathrm{mix}}^{(1,\ldots,1)}([0,1]^d)$ which is defined next.

Starting point is the Hlawka-Zaremba identity

$$\frac{1}{n} \sum_{i=1}^{n} f(\xi^i) - \int_{[0,1]^d} f(x) dx = \sum_{u \subseteq D} (-1)^{|u|} \int_{[0,1]^{|u|}} \frac{\partial^{|u|} f}{\partial x_u}(x_u, 1) \mathrm{disc}(x_u, 1) dx_u \,.$$

First general QMC construction: Digital nets (Sobol 69, Niederreiter 87)

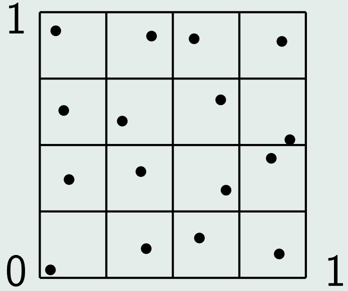
Elementary subintervals E in base $b \in \mathbb{N}$, $b \geq 2$:

$$E = \prod_{j=1}^{d} \left[\frac{a_j}{b^{d_j}}, \frac{a_j + 1}{b^{d_j}} \right),$$

where $a_i, d_i \in \mathbb{Z}_+, 0 \leq a_i < b^{d_i}, i = 1, \ldots, d$.

Let $m, t \in \mathbb{Z}_+$, $m \geq t$. A set of b^m points in $[0,1)^d$ is a (t, m, d)-net in base b if every elementary subinterval E in base b with $\lambda^d(E) = b^{t-m}$ contains b^t points.

Illustration of a $(0, 4, 2)$-net with $b = 2$



A sequence (ξ^i) in $[0,1)^d$ is a (t, d)-sequence in base b if, for all integers $k \in \mathbb{Z}_+$ and $m > t$, the set

$$\{\xi^i : kb^m \leq i < (k+1)b^m\}$$

is a (t, m, d)-net in base b.

For fixed b and m the (t, m, d)-net condition gets stronger if the *quality parameter* t gets smaller. The quantity $m-t$ is called the *strength* of the (t, m, d)-nets.

Theorem: There exist (t, d)-sequences (ξ^i) in $[0, 1]^d$ such that

$$D_n^*(\xi^1, \ldots, \xi^n) = O(n^{-1}(\log n)^{d-1}) \leq C(\delta, d)\, n^{-1+\delta} \quad (\forall \delta > 0).$$

Note that, in general, the constant $C(\delta, d)$ depends indeed upon δ and the dimension d. However, the constants for (t, d)-sequences are essentially smaller compared to the Halton sequence.

Specific sequences:

The Sobol' sequence (Sobol' 67) is a (t, d)-sequence in base $b = 2$, where t is a non-decreasing function of d;
the Faure sequence (Faure 82) is a $(0, d)$-sequence with $d \leq b$;
the Niederreiter squences (Niederreiter 88) include both Sobol' and Faure constructions as special cases; and the Niederreiter-Xing sequences.

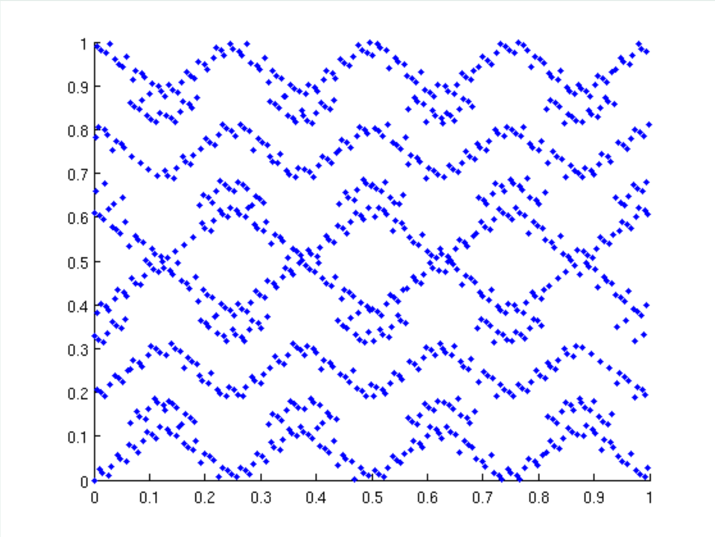
(Dick-Pillichshammer 10, Dick-Kuo-Sloan 14).

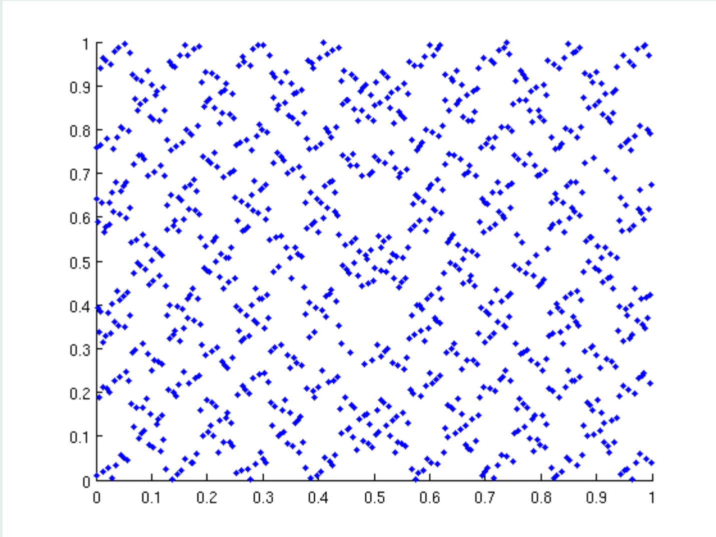
Recent development:

Scrambling of (t, m, d)-nets and (t, d)-sequences

Idea: Random permutation of the digits in each component (Owen 95).

Scrambled nets and sequences combine favorable properties of MC and QMC and improve their convergence properties (in a probabilistic sense).





left: $1\,000$ Niederreiter-points for $d = 40$, projection $(16, 18)$.
right: $1\,000$ Scrambled-Niederreiter-points for $d = 40$, projection $(16, 18)$.

Second general QMC construction: Lattices (Korobov 59, Sloan-Joe 94)

A *lattice* in $\mathbb{R}^d$ is a discrete subset of $\mathbb{R}^d$ which is closed under addition and subtraction. An *integration lattice* in $\mathbb{R}^d$ is a lattice which contains $\mathbb{Z}^d$ as a subset. A lattice rule is an equal-weight quadrature rule whose quadrature points are those points of an integration lattice that lie in $[0,1)^d$.
Every lattice rule can be written as a *multiple sum* involving one or more generating vectors.

Rank-1 lattice rule:
An n-point rank-1 lattice rule in d dimensions, also called the method of good lattice points, is a QMC method with quadrature points

$$\left\{\xi^i = \left\{\tfrac{i-1}{n} g\right\} : i = 1, \ldots, n\right\},$$

where $g \in \mathbb{Z}^d$ is the generating vector. The braces indicate that the fractional part is taken for each component, i.e., $\{z\} = z - \lfloor z \rfloor \in [0,1)$ for each $z \in \mathbb{R}_+$.
The components of g can be restricted to $\{0, 1, \ldots, n-1\}$ or even to

$$\mathbb{G}_n = \{z \in \mathbb{Z} : 1 \leq z \leq n-1 \text{ and } \gcd(z,n) = 1\}.$$

The number of elements in $\mathbb{G}_n$ is $\varphi(n) = |\mathbb{G}_n|$, the Euler totient function.

Example: (Korobov construction)

Given $a \in \mathbb{N}$, $1 \leq a \leq n-1$, with $\gcd(a,n) = 1$ we define

$$g = g(a) = (1, a, a^2, \ldots, a^{d-1}) \mod n\,.$$

Example: (Component-by-component (CBC) construction)

Given n, construct a generating vector $g = (g_1, \ldots, g_d)$ as follows:

1. Set $g_1 = 1$.

i. With $g_1, \ldots, g_{i-1}$ held fixed, choose $g_i \in \mathbb{G}_n$ to minimize a desired error criterion in i dimensions.

Theorem:

Let $Q_{n,d}$ denote a rank-1 lattice rule with generating vector g and the integrand f have an absolutely convergent complex Fourier series. Then

$$|I_d(f) - Q_{n,d}(f)| \leq c \sum_{\substack{h \in \mathbb{Z}^d \setminus \{0\} \\ h \cdot g \equiv 0 (\mathrm{mod}\ n)}} \frac{1}{(\bar{h}_1 \cdots \bar{h}_d)^\alpha},$$

where $f \in E_\alpha(c) = \{f : |\hat{f}(h)| \leq \frac{c}{(\bar{h}_1 \cdots \bar{h}_d)^\alpha}, h \in \mathbb{Z}^d\}$ with $c > 0$, $\alpha > 1$, $\bar{h} = \max\{1, |h|\}$ and $\hat{f}(h)$, $h \in \mathbb{Z}^d$, denoting the Fourier coefficients of f.

(Sloan-Joe 94)

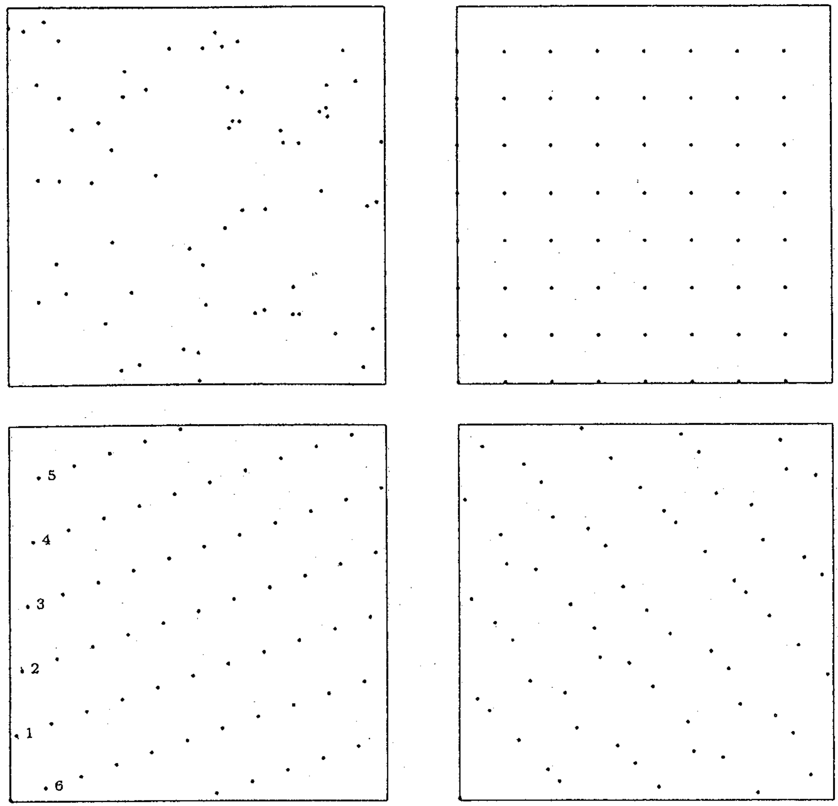


Fig. 5.3 Four different point sets with $n = 64$: random (top left), rectangular grid (top right), Korobov lattice (bottom left), and Sobol' (bottom right).

Recent development: Randomly shifted lattice rules:

If $\triangle$ is a sample from the uniform distribution in $[0,1]^d$. put

$$Q_{n,d}(\triangle; f) = \frac{1}{n}\sum_{i=1}^{n} f\Big(\{\tfrac{i-1}{n}g + \triangle\}\Big).$$

If $f \in \mathbb{F}_d$ one obtains

$$|I_d(f) - Q_{n,d}(\triangle; f)| \le e(Q_{n,d}(\triangle; \cdot))\|f\|_d$$

Hence, it follows

$$\mathbb{E}[|I_d(f) - Q_{n,d}(\triangle; f)|^2] \le \int_{[0,1]^d} e^2(Q_{n,d}(\triangle; \cdot))d\triangle\, \|f\|_d^2$$

Theorem:

If $\mathbb{F}_d$ is a kernel reproducing Hilbert space with kernel K, it holds

$$\int_{[0,1]^d} e^2(Q_{n,d}(\triangle; \cdot))d\triangle = -\int_{[0,1]^d}\int_{[0,1]^d} K(x,y)dxdy + \frac{1}{n^2}\sum_{i,j=1}^{n-1} K^{\mathrm{sh}}(\xi^i, \xi^j)\,,$$

where $\xi^i = \frac{i-1}{n}$, $i = 1, \ldots, n$, and K^{sh} is the shift-average kernel

$$K^{\mathrm{sh}}(x,y) = \int_{[0,1]^d} K(\{x+\triangle\}, \{y+\triangle\})d\triangle\,.$$

The kernel K^{sh} is shift-invariant and it can be shown

$$(\hat{e}(Q_{n,d}(\triangle;\cdot)))^2 := \int_{[0,1]^d} e^2(Q_{n,d}(\triangle;\cdot))d\triangle = e^2(Q_{n,d}),$$

where $Q_{n,d}(f) = \frac{1}{n}\sum_{i=1}^{n} f\left(\{\frac{i-1}{n}g\}\right)$ and the worst-case error $e(Q_{n,d})$ is taken in the reproducing kernel Hilbert space with kernel K^{sh}.

Theorem:

Let n be prime, $\mathbb{F}_d = \mathcal{W}_{2,\gamma,\mathrm{mix}}^{(1,\ldots,1)}([0,1]^d)$ and $g \in \mathbb{Z}^d$ be CBC constructed. Then there exists for any $\delta \in (0, \frac{1}{2}]$ a constant $C(\delta) > 0$ such that the mean quadrature error attains the optimal convergence rate

$$\hat{e}(Q_{n,d}(\triangle;\cdot)) \leq C(\delta)n^{-1+\delta},$$

where the constant $C(\delta)$ grows when δ decreases, but does not depend on the dimension d if the sequence (γ_j) satisfies the condition

$$\sum_{j=1}^{\infty} \gamma_j^{\frac{1}{2(1-\delta)}} < \infty \qquad (\text{e.g. } \gamma_j = \tfrac{1}{j^2}).$$

(Sloan/Wožniakowski 98, Sloan/Kuo/Joe 02, Kuo 03)

Conclusions

- Classical Quasi-Monte Carlo methods converge faster than Monte Carlo schemes, but the convergence rate becomes effective only for $n \geq e^d$.
- QMC methods can be constructed via integration lattices or via (t, m, d)-nets.
- Scrambled (t, m, d)-nets combine favorable properties of MC and QMC and have improved convergence properties.
- Recently developed randomly shifted lattice rules **lift the curse of dimensionality** and converge significantly faster than Monte Carlo.
- This presentation didn't cover the more recent development of digitally shifted polynomial lattice rules which allow for higher order convergence rates and error estimates of the form

$$\hat{e}(Q_{n,d}) \leq C(\delta) n^{-r+\delta},$$

if f belongs to $\mathcal{W}^{(r,\dots,r)}_{2,\gamma,\mathrm{mix}}([0,1]^d)$ and $\delta \in (0, \frac{1}{2}]$ (Dick-Pillichshammer 10).

References

I. M. Sobol': *Multidimensional Quadrature Formulas and Haar Functions*, Nauka, Moscow, 1969 (in Russian).

H. Niederreiter: *Random Number Generation and Quasi-Monte Carlo Methods*, SIAM, Philadelphia, 1992.

I. H. Sloan, S. Joe: *Lattice Methods for Multiple Integration*, Clarendon Press, Oxford 1994.

F. Y. Kuo, I. H. Sloan: Lifting the curse of dimensionality, *Notices of the AMS* 52 (2005), 1320–1328.

C. Lemieux: *Monte Carlo and Quasi-Monte Carlo Sampling*, Springer, 2009.

J. Dick, F. Pillichshammer: *Digital Nets and Sequences: Discrepancy Theory and Quasi-Monte Carlo Integration*, Cambridge University Press, 2010.

F. Y. Kuo, Ch. Schwab, I. H. Sloan: Quasi-Monte Carlo methods for high-dimensional integration: the standard (weighted Hilbert space) setting and beyond, ANZIAM Journal 53 (2011), 1–37.

J. Dick, F. Y. Kuo, I. H. Sloan: High-dimensional integration – the Quasi-Monte Carlo way, *Acta Numerica* (2014), 1–153.

I. H. Sloan, H. Woźniakowski: When are Quasi Monte Carlo algorithms efficient for high-dimensional integration, *Journal of Complexity* 14 (1998), 1–33.

I. H. Sloan, F. Y. Kuo and S. Joe: Constructing randomly shifted lattice rules in weighted Sobolev spaces, *SIAM Journal Numerical Analysis* 40 (2002), 1650–1665.

F. Y. Kuo: Component-by-component constructions achieve the optimal rate of convergence in weighted Korobov and Sobolev spaces, *Journal of Complexity* 19 (2003), 301-320.

A. B. Owen: Multidimensional variation for Quasi-Monte Carlo, in J. Fan, G. Li (Eds.), International Conference on Statistics, World Scientific Publ., 2005, 49–74.

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