Tutorial on quasi-Monte Carlo methods

Josef Dick

School of Mathematics and Statistics, UNSW, Sydney, Australia
josef.dick@unsw.edu.au
Roughly speaking:

- Markov chain Monte Carlo and quasi-Monte Carlo are for different types of problems;

- If you have a problem where Monte Carlo does not work, then chances are quasi-Monte Carlo will not work as well;

- If Monte Carlo works, but you want a faster method ⇒ try (randomized) quasi-Monte Carlo (some tweaking might be necessary).

- Quasi-Monte Carlo is an "experimental design" approach to Monte Carlo simulation;

In this talk we shall discuss how quasi-Monte Carlo can be faster than Monte Carlo under certain assumptions.
DISCREPANCY, REPRODUCING KERNEL HILBERT SPACES AND WORST-CASE ERROR

QUASI-MONTE CARLO POINT SETS

RANDOMIZATIONS

WEIGHTED FUNCTION SPACES AND TRACTABILITY
The task is to approximate an integral

\[ I_s(f) = \int_{[0,1]^s} f(z) \, dz \]

for some integrand \( f \) by some quadrature rule

\[ Q_{N,s}(f) = \frac{1}{N} \sum_{n=0}^{N-1} f(x_n) \]

at some sample points \( x_0, \ldots, x_{N-1} \in [0,1]^s \).
\[ \int_{[0,1]^s} f(x) \, dx \approx \frac{1}{N} \sum_{n=0}^{N-1} f(x_n) \]

In other words:

Area under curve = Volume of cube \times average function value.

• If \( x_0, \ldots, x_{N-1} \in [0, 1]^s \) are chosen randomly ⇒ Monte Carlo

• If \( x_0, \ldots, x_{N-1} \in [0, 1]^s \) chosen deterministically ⇒ Quasi-Monte Carlo
Smooth integrands

- Integrand $f : [0, 1] \to \mathbb{R}$; say continuously differentiable;
- We want to study the integration error:

$$
\int_0^1 f(x) \, dx - \frac{1}{N} \sum_{n=0}^{N-1} f(x_n).
$$

- Representation:

$$
f(x) = f(1) - \int_x^1 f'(t) \, dt = f(1) - \int_0^1 1_{[0,t]}(x)f'(t) \, dt,
$$

where

$$
1_{[0,t]}(x) = \begin{cases} 
1 & \text{if } x \in [0, t], \\
0 & \text{otherwise}.
\end{cases}
$$
• Substitute:

\[
\int_0^1 f(x) \, dx = \int_0^1 \left( f(1) - \int_0^1 1_{[0,t]}(x)f'(t) \, dt \right) \, dx
\]

\[
= f(1) - \int_0^1 \int_0^1 1_{[0,t]}(x)f'(t) \, dx \, dt,
\]

\[
\frac{1}{N} \sum_{n=0}^{N-1} f(x_n) = \frac{1}{N} \sum_{n=0}^{N-1} \left( f(1) - \int_0^1 1_{[0,t]}(x)f'(t) \, dt \right)
\]

\[
= f(1) - \int_0^1 \frac{1}{N} \sum_{n=0}^{N-1} 1_{[0,t]}(x_n)f'(t) \, dt.
\]

• Integration error:

\[
\int_0^1 f(x) \, dx - \frac{1}{N} \sum_{n=0}^{N-1} f(x_n) = \int_0^1 \left( \frac{1}{N} \sum_{n=0}^{N-1} 1_{[0,t]}(x_n) - \int_0^1 1_{[0,t]}(x) \, dx \right) f'(t) \, dt.
\]
Local discrepancy

Integration error:

\[
\int_0^1 f(x) \, dx - \frac{1}{N} \sum_{n=0}^{N-1} f(x_n) = \int_0^1 \left( \frac{1}{N} \sum_{n=0}^{N-1} 1_{[0,t]}(x_n) - t \right) f'(t) \, dt.
\]

Let \( P = \{x_0, \ldots, x_{N-1}\} \subset [0, 1] \). Define the **local discrepancy** by

\[
\Delta_P(t) = \frac{1}{N} \sum_{n=0}^{N-1} 1_{[0,t]}(x_n) - t, \quad t \in [0, 1].
\]

Then

\[
\int_0^1 f(x) \, dx - \frac{1}{N} \sum_{n=0}^{N-1} f(x_n) = \int_0^1 \Delta_P(t)f'(t) \, dt.
\]
Koksma-Hlawka inequality

\[
\left| \int_0^1 f(x) \, dx - \frac{1}{N} \sum_{n=0}^{N-1} f(x_n) \right| = \left| \int_0^1 \Delta_P(t)f'(t) \, dt \right|
\leq \left( \int_0^1 |\Delta_P(t)|^p \, dt \right)^{1/p} \left( \int_0^1 |f'(t)|^q \, dt \right)^{1/q}
= \|\Delta_P\|_{L_p} \|f'\|_{L_q}, \quad \frac{1}{p} + \frac{1}{q} = 1,
\]

where \(\|g\|_{L_p} = \left( \int |g|^p \right)^{1/p} \).
Interpretation of discrepancy

Let $P = \{x_0, \ldots, x_{N-1}\} \subset [0, 1]$. Recall the definition of the local discrepancy:

\[
\Delta_P(t) = \frac{1}{N} \sum_{n=0}^{N-1} 1_{[0,t]}(x_n) - t, \quad t \in [0, 1].
\]

Local discrepancy measures difference between uniform distribution and empirical distribution of quadrature points $P = \{x_0, \ldots, x_{N-1}\}$.

This is the Kolmogorov-Smirnov test for the difference between the empirical distribution of $\{x_0, \ldots, x_{N-1}\}$ and the uniform distribution.
Function space

Representation

\[ f(x) = f(1) - \int_x^1 f'(t) \, dt. \]

Define inner product:

\[ \langle f, g \rangle = f(1)g(1) + \int_0^1 f'(t)g'(t) \, dt. \]

and norm

\[ \|f\| = \sqrt{|f(1)|^2 + \int_0^1 |f'(t)|^2 \, dt}. \]

Function space:

\[ \mathcal{H} = \{ f : [0, 1] \to \mathbb{R} : f \text{ absolutely continuous and } \|f\| < \infty \}. \]
The **worst-case error** is defined by

\[
e(\mathcal{H}, P) = \sup_{f \in \mathcal{H}, \|f\| \leq 1} \left| \int_0^1 f(x) \, dx - \frac{1}{N} \sum_{n=0}^{N-1} f(x_n) \right|
\]
Recall:

\[ f(y) = f(1) \cdot 1 + \int_0^1 f'(x)(-1_{[0,x]}(y)) \, dx \]

and

\[ \langle f, g \rangle = f(1)g(1) + \int_0^1 f'(x)g'(x) \, dx. \]

**Goal:** Find set of functions \( g_y \in \mathcal{H} \) for each \( y \in [0, 1] \) such that

\[ \langle f, g_y \rangle = f(y) \quad \text{for all } f \in \mathcal{H}. \]

**Conclusions:**

- \( g_y(1) = 1 \) for all \( y \in [0, 1] \);
- \( g'_y(x) = -1_{[0,x]}(y) = \begin{cases} -1 & \text{if } y \leq x, \\ 0 & \text{otherwise}; \end{cases} \)
- Make \( g \) continuous such that \( g \in \mathcal{H} \);
It follows that

$$g_y(x) = 1 + \min\{1 - x, 1 - y\}.$$  

The function

$$K(x, y) := g_y(x) = 1 + \min\{1 - x, 1 - y\}, \quad x, y \in [0, 1]$$

is called **reproducing kernel**.

The space $\mathcal{H}$ is called a reproducing kernel Hilbert space (with reproducing kernel $K$).
Numerical integration in reproducing kernel Hilbert spaces

Function representation:

\[ \int_0^1 f(z) \, dz = \int_0^1 \langle f, K(\cdot, z) \rangle \, dz = \left\langle f, \int_0^1 K(\cdot, z) \, dz \right\rangle \]

\[ \frac{1}{N} \sum_{n=0}^{N-1} f(x_n) = \frac{1}{N} \sum_{n=0}^{N-1} \langle f, K(\cdot, x_n) \rangle = \left\langle f, \frac{1}{N} \sum_{n=0}^{N-1} K(\cdot, x_n) \right\rangle \]

Integration error

\[ \int_0^1 f(z) \, dz - \frac{1}{N} \sum_{n=0}^{N-1} f(x_n) = \left\langle f, \int_0^1 K(\cdot, z) \, dz - \frac{1}{N} \sum_{n=0}^{N-1} K(\cdot, x_n) \right\rangle \]

\[ = \langle f, h \rangle, \]

where

\[ h(x) = \int_0^1 K(x, z) \, dz - \frac{1}{N} \sum_{n=0}^{N-1} K(x, x_n). \]
Thus

\[ e(\mathcal{H}, P) = \sup_{f \in \mathcal{H}, \|f\| \leq 1} \left| \int_0^1 f(z) \, dz - \frac{1}{N} \sum_{n=0}^{N-1} f(x_n) \right| \]

\[ = \sup_{f \in \mathcal{H}, \|f\| = 1} |\langle f, h \rangle| \]

\[ = \sup_{f \in \mathcal{H}, f \neq 0} \left| \left\langle \frac{f}{\|f\|}, h \right\rangle \right| \]

\[ = \frac{\langle h, h \rangle}{\|h\|} = \|h\|, \]

since the supremum is attained when choosing \( f = \frac{h}{\|h\|} \in \mathcal{H} \).
Worst-case error in reproducing kernel Hilbert spaces

\[ e^2(\mathcal{H}, P) = \int_0^1 \int_0^1 K(x, y) \, dx \, dy - \frac{2}{N} \sum_{n=0}^{N-1} \int_0^1 K(x, x_n) \, dx \]

\[ + \frac{1}{N^2} \sum_{n,m=0}^{N-1} K(x_n, x_m) \]
Numerical integration in higher dimensions

- Tensor product space $\mathcal{H}_s = \mathcal{H} \times \cdots \times \mathcal{H}$.
- Reproducing kernel

$$K(\mathbf{x}, \mathbf{y}) = \prod_{i=1}^{s} [1 + \min\{1 - x_i, 1 - y_i\}],$$

where $\mathbf{x} = (x_1, \ldots, x_s), \mathbf{y} = (y_1, \ldots, y_s) \in [0, 1]^s$.
- Functions $f \in \mathcal{H}_s$ have partial mixed derivatives up to order 1 in each variable

$$\frac{\partial^{|u|} f}{\partial \mathbf{x}_u} \in L_2([0, 1]^s),$$

where $u \subseteq \{1, \ldots, s\}, \mathbf{x}_u = (x_i)_{i \in u}$ and $|u|$ denotes the cardinality of $u$ and where $\frac{\partial^{|u|} f}{\partial \mathbf{x}_u}(\mathbf{x}_u, \mathbf{1}) = 0$ for all $u \subset \{1, \ldots, s\}$. 
Worst-case error

Again

\[ e^2(\mathcal{H}, P) = \int_{[0,1]^s} \int_{[0,1]^s} K(x, y) \, dx \, dy - \frac{2}{N} \sum_{n=0}^{N-1} \int_{[0,1]^s} K(x, x_n) \, dx \]

\[ + \frac{1}{N^2} \sum_{n,m=0}^{N-1} K(x_n, x_m) \]

and

\[ e^2(\mathcal{H}, P) = \int_{[0,1]^s} \Delta_P(x) \frac{\partial^s f}{\partial x} (x) \, dx, \]

where

\[ \Delta_P(x) = \frac{1}{N} \sum_{n=0}^{N-1} 1_{[0,x]}(x_n) - \prod_{i=1}^{s} x_i. \]
Point set $P = \{x_0, \ldots, x_{N-1}\} \subset [0, 1]^s$, $t = (t_1, \ldots, t_s) \in [0, 1]^s$.

- Local discrepancy:

$$\Delta_P(t) = \frac{1}{N} \sum_{n=0}^{N-1} 1_{[0,t]}(x_n) - \prod_{i=1}^{s} t_i,$$

where $[0, t] = \prod_{i=1}^{s} [0, t_i]$. 
Koksma-Hlawka inequality

Let $f : [0, 1]^s \rightarrow \mathbb{R}$ with

$$
\|f\| = \left( \int_{[0,1]^s} \left| \frac{\partial^s}{\partial x} f(x) \right|^q \, dx \right)^{1/q},
$$

and where $\frac{\partial^{|u|}}{\partial x_u} f(x_u, 1) = 0$ for all $u \subset \{1, \ldots, s\}$.

Then

$$
\left| \int_{[0,1]^s} f(x) \, dx - \frac{1}{N} \sum_{n=0}^{N-1} f(x_n) \right| \leq \|f\| \|\Delta P\|_{L_p},
$$

where $\frac{1}{p} + \frac{1}{q} = 1$.

$\Rightarrow$ Construct points $P = \{x_0, \ldots, x_{N-1}\} \subset [0, 1]^s$ with small discrepancy

$$
L_p(P) = \|\Delta P\|_{L_p}.
$$

It is often useful to consider different reproducing kernels yielding different worst-case errors and discrepancies. We will see further examples later.
• **Lattice rules** (Bilyk, Brauchart, Cools, D., Hellekalek, Hickernell, Hlawka, Joe, Keller, Korobov, Kritzer, Kuo, Larcher, L’Ecuyer, Lemieux, Leobacher, Niederreiter, Nuyens, Pillichshammer, Sinescu, Sloan, Temlyakov, Wang, Woźniakowski, ...)

• **Digital nets and sequences** (Baldeaux, Bierbrauer, Brauchart, Chen, D., Edel, Faure, Hellekalek, Hofer, Keller, Kritzer, Kuo, Larcher, Leobacher, Niederreiter, Owen, Özbudak, Pillichshammer, Pirsic, Schmid, Skriganov, Sobol’, Wang, Xing, Yue, ...)

• **Hammersley-Halton sequences** (Atanassov, De Clerck, Faure, Halton, Hammersley, Kritzer, Larcher, Lemieux, Pillichshammer, Pirsic, White, ...)

• **Kronecker sequences** (Beck, Hellekalek, Larcher, Niederreiter, Schoissengeier, ...)
Let $N \in \mathbb{N}$, let

$$g = (g_1, \ldots, g_s) \in \{1, \ldots, N - 1\}^s.$$

Choose the quadrature points as

$$x_n = \left\{ \frac{ng}{N} \right\}, \quad \text{for } n = 0, \ldots, N - 1,$$

where $\{z\} = z - \lfloor z \rfloor$ for $z \in \mathbb{R}_0^+$. 
Fibonacci lattice rules

Lattice rule with $N = 55$ points and generating vector $g = (1, 34)$. 
Fibonacci lattice rules

Lattice rule with $N = 89$ points and generating vector $g = (1, 55)$. 

![Fibonacci lattice rule with 89 points and generating vector (1, 55).](image-url)
In comparison: Random point set

Random set of 64 points generated by Matlab using Mersenne Twister.
How to find generating vector $g$?

Reproducing kernel:

$$K(x, y) = \prod_{i=1}^{s} \left( 1 + 2\pi B_{\alpha} (\{x_i - y_i\}) \right),$$

where $\{x_i - y_i\} = (x_i - y_i) - \lfloor x_i - y_i \rfloor$ is the fractional part of $x_i - y_i$.

Reproducing kernel Hilbert space of Fourier series:

$$f(x) = \sum_{h \in \mathbb{Z}^s} \hat{f}(h)e^{2\pi i h \cdot x}.$$
Worst-case error:

\[ e_{\alpha}^2(g) = -1 + \frac{1}{N} \sum_{n=0}^{N-1} \prod_{i=1}^{s} \left( 1 + 2\pi B_{\alpha} \left( \{ \frac{ng_i}{N} \} \right) \right), \]

where \( B_{\alpha} \) is the Bernoulli polynomial of order \( \alpha \). For instance \( B_2(x) = x^2 - x + 1/6 \).
Component-by-component construction (Korobov, Sloan-Reztsov, Sloan-Kuo-Joe, Nuyens-Cools)

- Set $g_1 = 1$.
- For $d = 2, \ldots, s$ assume that we have found $g_2, \ldots, g_{d-1}$. Then find $g_d \in \{1, \ldots, N - 1\}$ which minimizes $e_{\alpha}(g_1, \ldots, g_{d-1}, g)$ as a function of $g$, i.e.

$$
g_d = \arg\min_{g \in \{1, \ldots, N - 1\}} e^2_{\alpha}(g_1, \ldots, g_{d-1}, g).
$$

Using fast Fourier transform (Nuyens, Cools), a good generating vector $g \in \{1, \ldots, N - 1\}^s$ can be found in $O(sN \log N)$ operations.
Hammersley-Halton sequence

Radical inverse function in base $b$:

- Let $n \in \mathbb{N}_0$ have base $b$ expansion
  \[ n = n_0 + n_1 b + \cdots + n_{a-1} b^{a-1}. \]
- Set
  \[ \varphi_b(n) = \frac{n_0}{b} + \frac{n_1}{b^2} + \cdots + \frac{n_{a-1}}{b^a} \in [0, 1]. \]

Hammersley-Halton sequence:

- Let $\mathbb{P} = \{p_1, p_2, \ldots, \}$ be the set of prime numbers in increasing order, i.e. $p_1 = 2, p_2 = 3, p_3 = 5, p_4 = 7, \ldots$.
- Define quadrature points $x_0, x_1, \ldots$ by
  \[ x_n = (\varphi_{p_1}(n), \varphi_{p_2}(n), \ldots, \varphi_{p_s}(n)) \quad \text{for } n = 0, 1, 2, \ldots. \]
Hammersley-Halton point set with 64 points.
• Choose prime number $b$ and finite field $\mathbb{Z}_b = \{0, 1, \ldots, b - 1\}$ of order $b$.

• Choose $C_1, \ldots, C_s \in \mathbb{Z}_b^{m \times m}$.

• Let $n = n_0 + n_1 b + \cdots + n_{m-1} b^{m-1}$ and set $
vec{n} = (n_0, \ldots, n_{m-1})^\top \in \mathbb{Z}_b^m$.

• Let $\vec{y}_{n,i} = C_i \nvec{n}$ for $1 \leq i \leq s$, $0 \leq n < b^m$.

• For $\vec{y}_{n,i} = (y_{n,i,1}, \ldots, y_{n,i,m})^\top \in \mathbb{Z}_b^m$ let

$$x_{n,i} = \frac{y_{n,i,1}}{b} + \cdots + \frac{y_{n,i,m}}{b^m}.$$ 

• Set $\vec{x}_n = (x_{n,1}, \ldots, x_{n,s})$ for $0 \leq n < b^m$. 
Let $m, s \geq 1$ and $b \geq 2$ be integers. A point set $P = \{x_0, \ldots, x_{b^m-1}\}$ is called a $(t, m, s)$-net in base $b$, if for all integers $d_1, \ldots, d_s \geq 0$ with

$$d_1 + \cdots + d_s = m - t$$

the number of points in the elementary intervals

$$\prod_{i=1}^{s} \left[ \frac{a_i}{b^{d_i}}, \frac{a_i + 1}{b^{d_i}} \right)$$

where $0 \leq a_i < b^{d_i}$,

is $b^t$. 
If $P = \{x_0, \ldots, x_{b^m-1}\} \subset [0, 1]^s$ is a $(t, m, s)$-net, then local discrepancy function

$$\Delta_P \left( \frac{a_1}{b^{d_1}}, \ldots, \frac{a_s}{b^{d_s}} \right) = 0$$

for all $0 \leq a_i < b^{d_i}$, $d_i \geq 0$, $d_1 + \cdots + d_s = m - t$. 
The first 64 points of a Sobol sequence which form a \((0, 6, 2)\)-net in base 2.
The first 64 points of a Niederreiter-Xing sequence.
The first 1024 points of a Niederreiter-Xing sequence.
Let $\alpha_1, \ldots, \alpha_s \in \mathbb{R}$ be linearly independent over $\mathbb{Q}$.

Let

$$x_n = (\{n\alpha_1\}, \ldots, \{n\alpha_s\}) \quad \text{for } n = 0, 1, 2, \ldots$$

For instance, one can choose $\alpha_i = \sqrt{p_i}$ where $p_i$ is the $i$th prime number.
Kronecker sequence

The first 64 points of a Kronecker sequence.
Discrepancy bounds

• It is known that for all the constructions above one has

\[ L_p(P) \leq C_s \frac{(\log N)^{c(s)}}{N}, \quad \text{for all } N \geq 1, 1 \leq p \leq \infty, \]

where \( C_s > 0 \) and \( c(s) \approx s \).

• Lower bound: For all point sets \( P \) consisting of \( N \) points we have

\[ L_p(P) \geq C'_s \frac{(\log N)^{(s-1)/2}}{N} \quad \text{for all } N \geq 1, 1 < p \leq \infty. \]

In comparison: random point set

\[ \mathbb{E}(L_2(P)) = O \left( \sqrt{\log \log N} \right). \]
For $1 < p < \infty$, the exact order of convergence is known to be of order

$$\frac{(\log N)^{(s-1)/2}}{N}.$$ 

Explicit constructions of such points were achieved by Chen & Skriganov for $p = 2$ and Skriganov for $1 < p < \infty$.

**Great open problem:** What is the correct order of convergence of

$$\min_{P \subseteq [0,1]^s, |P| = N} L_\infty(P)?$$
Randomized quasi-Monte Carlo

Introduce random element to deterministic point set.

Has several advantages:

- yields an unbiased estimator;
- there is a statistical error estimate;
- better rate of convergence of the random-case error compared to worst-case error;
- performs similar to Monte Carlo for $L_2$ functions but has better rate of convergence for smooth functions;
Lattice rules can be randomized using a random shift:

- **Lattice point set:**
  \[
  \left\{ \frac{ng}{N}, \quad n = 0, 1, \ldots, N - 1 \right\}.
  \]

- **Shifted lattice rule:** choose \( \sigma \in [0, 1]^s \) uniformly distributed; then the shifted lattice point set is given by
  \[
  \left\{ \frac{ng}{N} + \sigma, \quad n = 0, 1, \ldots, N - 1 \right\}.
  \]
Lattice point set
Shifted lattice point set
Owen’s scrambling

• Let $\mathbb{Z}_b = \{0, 1, \ldots, b - 1\}$ and let

\[ x = \frac{x_1}{b} + \frac{x_2}{b^2} + \cdots, \quad x_1, x_2, \ldots \in \mathbb{Z}_b. \]

• Randomly choose permutations $\sigma, \sigma_{x_1}, \sigma_{x_1,x_2}, \ldots : \mathbb{Z}_b \to \mathbb{Z}_b$.

• Let

\[ y_1 = \sigma(x_1) \]
\[ y_2 = \sigma_{x_1}(x_2) \]
\[ y_3 = \sigma_{x_1,x_2}(x_3) \]
\[ \ldots \]

• Set

\[ y = \frac{y_1}{b} + \frac{y_2}{b^2} + \cdots. \]
The first 1024 points of a Sobol sequence (left) and a scrambled Sobol sequence (right).
Let
\[ e(R) = \int_{[0,1]^s} f(x) \, dx - \frac{1}{N} \sum_{n=0}^{b^m-1} f(y_n). \]

Then
- \[ \mathbb{E}(e(R)) = 0 \]
- \[ \text{Var}(e(R)) = \mathcal{O}\left(\frac{(\log N)^s}{N^{2\alpha+1}}\right) \]

for integrands with smoothness \(0 \leq \alpha \leq 1\).
Although the convergence rate is better for quasi-Monte Carlo, there is a stronger dependence on the dimension:

- Monte Carlo: error $= \mathcal{O}(N^{-1/2})$;
- Quasi-Monte Carlo: error $= \mathcal{O}\left(\frac{(\log N)^{s-1}}{N}\right)$;

Notice that $g(N) := N^{-1}(\log N)^{s-1}$ is an increasing function for $N \leq e^{s-1}$.

So if $s$ is large, say $s \approx 30$, then $N^{-1}(\log N)^{29}$ increases for $N \leq 10^{12}$. 
For integration error in $\mathcal{H}$ with reproducing kernel

$$K(x, y) = \prod_{i=1}^{s} \min\{1 - x_i, 1 - y_i\}$$

it is known that

$$e^2(\mathcal{H}, P) \geq \left(1 - N \left(\frac{8}{9}\right)^s\right) e^2(\mathcal{H}, \emptyset).$$

$\Rightarrow$ Error can only decrease if

$$N > \text{constant} \cdot \left(\frac{9}{8}\right)^s.$$
Study weighted function spaces: Introduce \( \gamma_1, \gamma_2, \ldots, > 0 \) and define

\[
K(x, y) = \prod_{i=1}^{s} (1 + \gamma_i \min \{1 - x_i, 1 - y_i\}).
\]

Then if

\[
\sum_{i=1}^{\infty} \gamma_i < \infty
\]

we have

\[
e(H_{s, \gamma}, P) \leq CN^{-\delta},
\]

where \( C > 0 \) is independent of the dimension \( s \) and \( \delta < 1 \).
Tractability

- Minimal error over all methods using $N$ points:

\[ e^*_N(\mathcal{H}) = \inf_{P:|P|=N} e(\mathcal{H}, P); \]

- Inverse of the error:

\[ N^*(s, \varepsilon) = \min\{N \in \mathbb{N} : e^*_N(\mathcal{H}) \leq \varepsilon e^*_0(\mathcal{H})\} \]

for $\varepsilon > 0$;

- Strong tractability:

\[ N^*(s, \varepsilon) \leq C\varepsilon^{-\beta} \]

for some constant $C > 0$;

Sloan and Woźniakowski: For the function space considered above, we get strong tractability if

\[ \sum_{i=1}^{\infty} \gamma_i < \infty. \]
Recall the local discrepancy function
\[ \Delta_P(t) = \frac{1}{N} \sum_{n=0}^{N-1} 1_{[0,t)}(x_n) - t_1 \cdots t_s, \quad \text{where } P = \{x_0, \ldots, x_{N-1}\}. \]

The star-discrepancy is given by
\[ D^*_N(P) = \sup_{t \in [0,1]^s} |\Delta_P(t)|. \]

Result by Heinrich, Novak, Woźniakowski, Wasilkowski:

Minimal star discrepancy
\[ D^*(s, N) = \inf_P D^*_N(P) \leq C \sqrt{\frac{s}{N}} \quad \text{for all } s, N \in \mathbb{N}. \]

This implies that
\[ N^*(s, \varepsilon) \leq Cs\varepsilon^{-2}. \]
• Quasi-Monte Carlo methods which achieve convergence rates of order

\[ N^{-\alpha} (\log N)^{\alpha s}, \quad \alpha > 1 \]

for sufficiently smooth integrands;

• Construction of point sets whose star-discrepancy is tractable;

• Completely uniformly distributed point sets to speed up convergence in Markov chain Monte Carlo algorithms;

• Infinite dimensional integration: Integrands have infinitely many variables;

• Connections to codes, orthogonal arrays and experimental designs;

• Quasi-Monte Carlo for function approximation;

• Choosing the weights in applications;

• Quasi-Monte Carlo on the sphere;
Thank You!

Special thanks to
- Dirk Nuyens for creating many of the pictures in this talk;
- Friedrich Pillichshammer for letting me use a picture of his daughters;
- Fred Hickernell, Peter Kritzer, Art Owen, and Friedrich Pillichshammer for helpful comments on the slides;
- All colleagues who worked on quasi-Monte Carlo methods over the years.