Application of Global Sensitivity Analysis and Quasi Monte Carlo methods in finance

Sergei Kucherenko

Imperial College London, UK
s.kucherenko@imperial.ac.uk
Outline

Application of MC methods to high dimensional path dependent integrals

Comparison of MC and Quasi MC methods. Properties of Low Discrepancy Sequences

Why Quasi MC remains superior over MC in high dimensions and for discontinuous payoffs although classical theory does not support this effect?

Global Sensitivity Analysis and Sobol’ Sensitivity Indices

Effective dimensions and how to reduce them (Brownian bridge, PCA, control variate techniques)

Practical examples (Asian options; Cox, Ingersoll and Ross interest rate model; martingale control variate approximations)

Derivative based Global Sensitivity Measures and “Global Greeks”
MC methods in finance

Many problem in finance can be formulated as high-dimensional integrals

$$I[f] = \int_{H^n} f(\vec{x}) d\vec{x}$$

Option pricing: the Wiener integral over paths of Brownian motion in the space of functionals $F[x(t)]:$

$$I = \int_C F[x(t)] d_W x,$$

$x(t)$ – continuous in $0 \leq t \leq T$, $x(0) = x_0$

$I = E(F[W(t)])$, $W(t)$ – random Wiener processes (a Brownian motion)

**Monte Carlo approach:** to construct many random paths $W(t)$, evaluate functional and average results
Why integrals are high dimensional?

Main factors to consider:

1. The number of risk factors (underlyings) involved: e.g., valuation of basket option, yield-curve sensitivities of an interest rate option, VaR.
2. The number of time steps required (for each factor): e.g., valuation of Asian options, stochastic volatility (path dependence).
3. The number of options (basket of options)

When all factors are present, the dimension is the product of all factors.
Monte Carlo integration methods

Deterministic integration methods in high dimensions are not practical because of the "Curse of Dimensionality"

\[ I[f] = E[f(\bar{x})] \]

Monte Carlo: \( I_N[f] = \frac{1}{N} \sum_{i=1}^{N} f(\bar{z}_i) \)

\( \{\bar{z}_i\} \) – is a sequence of random points in \( H^n \)

Error: \( \varepsilon = |I[f] - I_N[f]| \)

\[ \varepsilon_N = \left( E(\varepsilon^2) \right)^{1/2} = \sigma(f) \frac{N}{N^{1/2}} \rightarrow \]

Convergence does not depend on dimensionality but it is slow
How to improve MC?

Slow convergence: \( \varepsilon_N = \frac{\sigma(f)}{N^{1/2}} \)

To improve MC convergence:

I. Decrease \( \sigma(f) \) by applying variance reduction techniques:
   - antithetic variables;
   - control variates;
   - stratified sampling;
   - importance sampling.

   but the rate of convergence remains \( \varepsilon_N \sim \frac{1}{N^{1/2}} \)

II. Use better (more uniformly distributed) sequences.
Quasi random sequences (Low discrepancy sequences)

Discrepancy is a measure of deviation from uniformity:

Definition: \( Q(\vec{y}) \in H^n \), \( Q(\vec{y}) = [0, y_1) \times [0, y_2) \times \ldots \times [0, y_n) \),

\( m(Q) \) – volume of \( Q \)

\[
D_N = \sup_{Q(\vec{y}) \in H^n} \left| \frac{N_{Q(\vec{y})}}{N} - m(Q) \right|
\]

Random sequences: \( D_N \rightarrow (\ln \ln N) / N^{1/2} \sim 1 / N^{1/2} \)

\( D_N \leq c(d) \frac{(\ln N)^n}{N} \) – Low discrepancy sequences (LDS)

Convergence: \( \varepsilon_{QMC} = |I[f] - I_N[f]| \leq V(f)D_N \),

\( \varepsilon_{QMC} = \frac{O(\ln N)^n}{N} \)

Assymptotically \( \varepsilon_{QMC} \sim O(1 / N) \rightarrow \) much higher than

\( \varepsilon_{MC} \sim O(1 / \sqrt{N}) \)
Sobol’ Sequences vrs Random numbers and regular grid

Unlike random numbers, successive Sobol’ points “know” about the position of previously sampled points and fill the gaps between them.
What is the optimal way to arrange N points in two dimensions?

Regular Grid

Sobol’ Sequence

Low dimensional projections of low discrepancy sequences are better distributed than higher dimensional projections.
Comparison of Sobol’ Sequences and random numbers

Normal probability plots

Histograms
Evaluation of quantiles I. Low quantile

Distribution \( f(x) = \sum_{i=1}^{n} x_i^2 \), dimension \( n = 5 \).

\( x_i \sim N(0,1) \) are independent standard normal variates

Low quantile (percentile for the cumulative distribution function) = 0.05

A superior convergence of the QMC method
Evaluation of quantiles II. High quantile

Distribution \( f(x) = \sum_{i=1}^{n} x_i^2 \), dimension \( n = 5 \).

\( x_i \sim N(0,1) \) are independent standard normal variates.

High quantile (percentile for the cumulative distribution function) = 0.95

QMC convergences faster than MC and LHS
How to construct a low discrepancy sequence?

Van der Corput sequence:
Number \(i\) written in base \(b\): \(i = (\cdots a_4 a_3 a_2 a_1 a_0)_b\)

In the decimal system: \(i = \sum_{j=1}^{m} a_j b^j\), \(0 \leq a_j \leq b - 1\)

Reverse the digits and add a radix point to obtain a number within the unit interval:
\(y = (0. a_4 a_3 a_2 a_1)_b\)

In the decimal system: \(h(i; b) = \sum_{j=1}^{m} a_j b^{-j-1}\)

Example:
\(i = 4, b = 2\)
\(4 = 1 \times 2^2 + 0 \times 2^1 + 0 \times 2^0 = (100)_2\)
\(0.001 \rightarrow 0 \times 2^{-1} + 0 \times 2^{-2} + 1 \times 2^{-3} = 1/8\)

Holton LDS:
\(h(i; b) = \sum_{j=1}^{m} a_j b^{-j-1}\)
\(\{(h(i; 2), h(i; 3), \ldots, h(i; b_n))\},\)
for each dimension a Van der Corput sequence but with a different prime number \(b_\star\)
How to construct Sobol’ Sequence?

A permutation of Van der Corput sequence in base 2

Number \( i \) written in base 2: \( i = (\cdots a_4 a_3 a_2 a_1 a_0)_2 \)

In the decimal system: \( i = \sum_{j=1}^{m} a_j 2^j \), \( 0 \leq a_j \leq 1 \)

\( y_i = (a_0 a_1 a_2 a_3 \ldots a_L)_2 \)

Construct vector \( g_i = C^j y_i \), \( C^j \) permutation matrix for dimension \( j \)

\( x_i \)-th element for dimension \( j \) is given by:

\[ x_i^j (i) = \sum_{l=1}^{L} g_{i,l} 2^{-l-1} \]

In practice:

1. Construct direction numbers:

\( \nu^j = (0, \nu_1^j, \nu_2^j, \nu_3^j \ldots \nu_b^j)_2 \)

The \( \nu^j \) numbers are given by the equation \( \nu^j = m^j / 2^j \),

where \( m^j < 2^j \) is an odd integer, \( \nu^j \) satisfy a recurrence relation

using the coefficients of a primitive polynomial in the Galois field G(2)
2. The Sobol' integer $x_n^j$ number:

$$x_n^j(i) = a_1v_1^j \oplus a_2v_2^j \oplus \ldots \oplus v_b^j$$

where $\oplus$ is an addition modulo 2 operator: $0 \oplus 0 = 0, 1 \oplus 1 = 0, 0 \oplus 1 = 1, 1 \oplus 0 = 1$.

$\oplus$ can also be seen as bit wise XOR.

3. Convert integer $x_n^j(i)$ to a uniform variate:

$$y_n^j(i) = x_n^j(i) / 2^b$$

**Generation of direction numbers:**

Primitive polynomial (irreducible polynomial with binary coefficients over the Galois filed G(2)):

$$P_l = x^q + c_1x^{q-1} + \ldots + c_{q-1}x + 1, \quad a_k \in \{0,1\}$$

Examples of primitive polynomials: $x + 1, \quad x^2 + x + 1, \quad x^3 + x + 1, \quad x^3 + x^2 + 1$.

A different primitive polynomial is used in each dimension:

$$m_i = 2c_1m_{i-1} \oplus 2^2c_2m_{i-2} \ldots \oplus 2^q m_{i-q} \oplus m_{i-q}$$

To fully define the direction numbers initial numbers $m_1, m_2, \ldots, m_n$ are required.
2D projections from adjacent dimensions for Sobol’ LDS

In Sobol's algorithm direction numbers is a key component to its efficiency

Ref: Peter Jackel, Monte Carlo Methods in Finance, John Wiley & Sons, 2002
Why Sobol’ sequences are so efficient and widely used in finance?

Convergence: \( \varepsilon = \frac{O(\ln N)^n}{N} \) – for all LDS

For Sobol' LDS: \( \varepsilon = \frac{O(\ln N)^{n-1}}{N} \), if \( N = 2^k, k \) – integer

Sobol' LDS:
1. Best uniformity of distribution as N goes to infinity.
2. Good distribution for fairly small initial sets.
3. A very fast computational algorithm.

Known LDS: Faure, Sobol’, Niederreiter

Many practical studies have proven that the Sobol’ LDS is superior to other LDS:

"Preponderance of the experimental evidence amassed to date points to Sobol' sequences as the most effective quasi-Monte Carlo method for application in financial engineering."

A low-discrepancy sequence is said to satisfy Property A if for any binary segment (not an arbitrary subset) of the $n$-dimensional sequence of length $2^n$ there is exactly one point in each $2^n$ hyper-octant that results from subdividing the unit hypercube along each of its length extensions into half.

A low-discrepancy sequence is said to satisfy Property A’ if for any binary segment (not an arbitrary subset) of the $n$-dimensional sequence of length $4^n$ there is exactly one point in each $4^n$ hyper-octant that results from subdividing the unit hypercube along each of its length extensions into four equal parts.
Distributions of 4 points in two dimensions

property A

<table>
<thead>
<tr>
<th>Method</th>
<th>Result</th>
</tr>
</thead>
<tbody>
<tr>
<td>MC</td>
<td>No</td>
</tr>
<tr>
<td>LHS</td>
<td>No</td>
</tr>
<tr>
<td>Sobol'</td>
<td>Yes</td>
</tr>
</tbody>
</table>
Distributions of 16 points in two dimensions

MC -> No
LHS -> No
Sobol' -> Yes

Property A’

No
No
Yes
Comparison of different Sobol’ sequence generators

<table>
<thead>
<tr>
<th>Generator</th>
<th>Maximum Dimension</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sobol-370</td>
<td>370</td>
</tr>
<tr>
<td>Sobol-8192</td>
<td>8192</td>
</tr>
<tr>
<td>Jackel</td>
<td>32</td>
</tr>
<tr>
<td>Lemieux</td>
<td>360</td>
</tr>
<tr>
<td>Kuo1</td>
<td>4925</td>
</tr>
<tr>
<td>Kuo2</td>
<td>3946</td>
</tr>
<tr>
<td>Kuo3</td>
<td>4686</td>
</tr>
<tr>
<td>JoeKuoD5</td>
<td>1999</td>
</tr>
<tr>
<td>JoeKuoD6</td>
<td>1799</td>
</tr>
<tr>
<td>JoeKuoD7</td>
<td>1899</td>
</tr>
</tbody>
</table>

The winner is SobolSeq generator: Sobol’ sequences satisfy two additional uniformity properties: Property A for all dimensions and Property A’ for adjacent dimensions. Maximum dimension: 16384

Free version www.broda.co.uk (including 64-bit version)
Comparison of different Sobol' sequence generators. Integration test

\[ I = \int_{[0,1]^s} \prod_{i=1}^{s} (1 + c_i(x_i - 0.5)) \, dx_i \]
Discrepancy
I. Low Dimensions

Discrepancy, n=5

Discrepancy, n=20
Discrepancy
II. High Dimensions

MC in high-dimensions has smaller discrepancy!
Are QMC efficient for high dimensional problems?

\[ \varepsilon_{QMC} = \frac{O(\ln N)^n}{N} \]

Assymptotically \( \varepsilon_{QMC} \sim O(1/N) \)

but \( \varepsilon_{QMC} \) increases with \( N \) until \( N \approx \exp(n) \)

\( n = 50, N \approx 5 \times 10^{21} \) – not achievable for practical applications

“For high-dimensional problems (n > 12),
QMC offers no practical advantage over Monte Carlo”

( Bratley, Fox, and Niederreiter (1992)) ?!
Option pricing. Discretization of the Wiener process

The asset follows geometrical Brownian motion:

\[ dS = \mu S \, dt + \sigma S \, dW, \quad dW = z(\sqrt{dt})^{1/2}, \quad z \sim N(0,1) \]

Using Ito's lemma

\[ S(t) = S_0 \exp[(\mu - \frac{1}{2} \sigma^2) t + \sigma W(t)], \quad W(t) - \text{Wiener path} \]

For time step \( \Delta t \)

\[ S(t + \Delta t) = S(t) \exp[(r - \frac{1}{2} \sigma^2) \Delta t + \sigma(W(t + \Delta t) - W(t))] \]

For the standard discretization algorithm

\[ W(t_{i+1}) = W(t_i) + \sqrt{\Delta t} z_{i+1}, \quad \Delta t = T / n, \quad 0 \leq i \leq n - 1 \]

A terminal asset value:

\[ S(T) = S_0 \exp[(r - \frac{1}{2} \sigma^2) T + \sigma \sqrt{\Delta t} (z_1 + z_2 + \ldots + z_n)] \]
Approximations of path dependent integrals with Brownian bridge scheme

Geometrical Brownian motion:
\[ dS = \mu S \, dt + \sigma S \, dW \]

SDE:
\[ dW = z \sqrt{dt}, \quad z \sim N(0,1) \]

Brownian bridge algorithm:
\[
\begin{align*}
W(T) &= W_0 + \sqrt{T} z_1, \\
W(T/2) &= \frac{1}{2} (W(T) + W_0) + \frac{1}{2} \sqrt{T} z_2, \\
W(T/4) &= \frac{1}{2} (W(T/2) + W_0) + \frac{1}{2} \sqrt{T/2} z_3, \\
W(3T/4) &= \frac{1}{2} (W(T/2) + W(T)) + \frac{1}{2} \sqrt{T/2} z_4, \\
\vdots \\
W((n-1)T/n) &= \frac{1}{2} (W(((n-2)T/n) + W(T)) + \frac{1}{2} \sqrt{2T/n} z_n.
\end{align*}
\]
The value of European style options

\[ C(K, T) = e^{-rT} E^Q \left[ P(S(t), K) \right] \]

The payoff function for an Asian call option

\[ P_A = \max(\bar{S} - K, 0), \]

For a geometric average Asian call: \( \bar{S} = \left( \prod_{i=1}^{n} S_i \right)^{1/n} \)

\[ C(K, T) = e^{-rT} \int_{H^n} \max[0, \left( \prod_{i=1}^{n} S_0 \exp\left( (r - \frac{\sigma^2}{2})t_i + \sigma \sqrt{\frac{T}{n}} \sum_{j=1}^{i} \Phi^{-1}(u_j) \right) \right]^{1/n} \]

\[ -K) du_1 ... du_n. \]

And there is a closed form solution
MC simulation of option pricing. Discretization

In a general case

\[ C_N(K,T) = e^{-rT} \left[ \frac{1}{N} \sum_{i=1}^{N} P\left( S_0, S_1^{(i)}, \cdots, S_T^{(i)}, K \right) \right] \]

For the case of a European call

\[ C_N(K,T) = \frac{1}{N} \sum_{i=1}^{N} C^{(i)} = e^{-rT} \left[ \frac{1}{N} \sum_{i=1}^{N} \max(S_T^{(i)} - K, 0) \right] \]
MC and QMC methods with standard and Brownian Bridge discretizations

Asian Call (32 observations)

$S=100$, $K=105$, $r=0.05$, $s=0.2$, $T=0.5$, $C=3.84$ (analytical)

Call price vrs the number of paths.

MC - slow convergences, convergence curve is highly oscillating. QMC convergence – monotonic. Convergence is much faster for Brownian bridge
Asian call. Convergence curves

**Asian Call with geometric averaging. 252 observations**

*S=100, K=105, r=0.05, s=0.2, T=1.0, C=5.56 (analytical)*

\[
\varepsilon = \left( \frac{1}{K} \sum_{k=1}^{K} (I - I_N^k)^2 \right)^{1/2}
\]

\[\varepsilon \sim N^{-\alpha}, \ 0 < \alpha < 1\]

---

**Log-log plot of the root mean square error versus the number of paths.**

Brownian bridge – much faster convergence with QMC methods: \(\sim 1/N^{0.8}\) even in high dimensions. Why?
ANOVA decomposition and Sensitivity Indices

\( x \in \Omega \rightarrow \text{Model, } f(x) \rightarrow Y \)

Consider a model

- \( x \) is a vector of input variables
- \( Y \) is the model output.

**ANOVA decomposition:**

\[
Y = f(x) = f_0 + \sum_{i=1}^{k} f_i(x_i) + \sum_{i} \sum_{j>i} f_{ij}(x_i, x_j) + \ldots + f_{1,2,\ldots,k}(x_1, x_2, \ldots, x_k),
\]

\[
\int_{0}^{1} f_{i_1\ldots i_s}(x_{i_1}, \ldots, x_{i_s}) \, dx_{i_k} = 0, \quad \forall k, \ 1 \leq k \leq s
\]

**Variance decomposition:**

\[
\sigma^2 = \sum_{i} \sigma^2_i + \sum_{i,j} \sigma^2_{ij} + \ldots \sigma^2_{1,2,\ldots,n}
\]

**Sobol’ SI:**

\[
1 = \sum_{i=1}^{k} S_i + \sum_{i<j} S_{ij} + \sum_{i<j<l} S_{ijl} + \ldots + S_{1,2,\ldots,k}
\]
ANOVA decomposition and Sensitivity Indices

\[ f(x) = f_0 + \sum_{i=1}^{k} f_i(x_i) + \sum_{i}^{j>i} f_{ij}(x_i, x_j) + \ldots + f_{1,2,\ldots,k}(x_1, x_2, \ldots, x_k), \]

\[ \int_{0}^{1} f_{i_{1}\ldots i_k}(x_{i_1}, \ldots, x_{i_k}) \, dx_{i_k} = 0, \quad \forall k, \ 1 \leq k \leq s \]

\[ \int_{0}^{1} f_{v}(x_v) f_{u}(x_u) \, dx = 0, \quad \forall v \neq u. \]

\[ f_0 = \int_{H^n} f(x) \, dx, \]

\[ f_i(x_i) = \int_{H^n} f(x) \prod_{j \neq i}^{n} dx_j - f_0, \]

\[ f_{ij}(x_i, x_j) = \int_{H^n} f(x) \prod_{k \neq i, j}^{n} dx_k - f_i(x_i) - f_j(x_j) - f_0, \ldots \]
ANOVA decomposition and Sensitivity Indices. Test case

\[ f(x_1, x_2) = f_0 + f_1(x_1) + f_2(x_2) + f_{12}(x_1, x_2), \]
\[ f(x_1, x_2) = x_1 x_2 \rightarrow f_0 = \frac{1}{4}, \]
\[ f_1(x_1) = \int_{H^n} f(x) \, dx_2 - f_0 = \frac{1}{2} x_1 - \frac{1}{4}, \]
\[ f_2(x_2) = \int_{H^n} f(x) \, dx_1 - f_0 = \frac{1}{2} x_2 - \frac{1}{4}, \]
\[ f_{12}(x_1, x_2) = x_1 x_2 - \frac{1}{2} x_1 - \frac{1}{2} x_2 + \frac{1}{4}. \]

\[ S_1 = \frac{\int_{H^n} f_1^2(x_1) \, dx_1}{\sigma^2} = \frac{3}{7}, \]
\[ S_2 = S_1 = \frac{3}{7}, \quad S_{12} = \frac{1}{7}. \]
Sobol’ Sensitivity Indices (SI)

**Definition:**

\[
S_{i_1...i_s} = \frac{\sigma_{i_1...i_s}^2}{\sigma^2}
\]

\[
\sigma_{i_1...i_s}^2 = \int_0^1 f_{i_1...i_s}^2(x_{i_1},...,x_{i_s}) \, dx_{i_1},...,x_{i_s} \quad \text{- partial variances}
\]

\[
\sigma^2 = \int_0^1 (f(x) - f_0)^2 \, dx \quad \text{- variance}
\]

**Sensitivity indices for subsets of variables:**

\[
\sigma_y^2 = \sum_{s=1}^{m} \sum_{\{i_1,\ldots,i_s\} \in K} \sigma_{i_1,...,i_s}^2
\]

The total variance:

\[
\left(\sigma_y^{\text{tot}}\right)^2 = \sigma^2 - \sigma_z^2
\]

**Corresponding global sensitivity indices:**

\[
S_y = \frac{\sigma_y^2}{\sigma^2}, \quad S_y^{\text{tot}} = \left(\sigma_y^{\text{tot}}\right)^2 / \sigma^2.
\]
How to use Sobol’ Sensitivity Indices?

\[ 0 \leq S_y \leq S_{y}^{tot} \leq 1 \]

- \( S_{y}^{tot} - S_y \) accounts for all interactions between \( y \) and \( z \), \( x=(y,z) \).

- The important indices in practice are \( S_i \) and \( S_i^{tot} \):

  \[ S_i^{tot} = 0 \rightarrow f(x) \text{ does not depend on } x_i ; \]
  \[ S_i = 1 \rightarrow f(x) \text{ only depends on } x_i ; \]
  \[ S_i = S_i^{tot} \text{ corresponds to the absence of interactions between } x_i \text{ and other variables} \]

  If \( \sum_{s=1}^{n} S_i = 1 \), then function has additive structure: \( f(x) = f_0 + \sum_{i} f_i(x_i) \)

- Fixing unessential variables:

  If \( S_{z}^{tot} \ll 1 \rightarrow f(x) \) does not depend on \( z \) so it can be fixed.

  \( f(x) \approx f(y, z_0) \rightarrow \) complexity reduction, from \( n \) to \( n-n_z \) variables.
Evaluation of Sobol’ Sensitivity Indices

Straightforward use of Anova decomposition requires

$2^n$ integral evaluations – not practical!

There are efficient formulas for evaluation of Sobol’ Sensitivity:

\[
S_y = \frac{1}{\sigma^2} \left[ \int_0^1 f(y, z)[f(y, z') - f(y', z')] \ dy dy' dz dz', \right.
\]
\[
S_y^{tot} = \frac{1}{2\sigma^2} \int_0^1 [f(y, z) - f(y', z)]^2 dy dz dz',
\]
\[
\sigma^2 = \int_0^1 f^2(y, z) \ dy dz - f_0^2
\]

Evaluation is reduced to high-dimensional integration by MC/QMC methods.

The number of function evaluations is $N(n+2)$.
Applications of Global Sensitivity Analysis

Global Sensitivity Analysis can be used to

- identify key parameters whose uncertainty most strongly affects the output;
- rank variables, fix unessential variables and reduce model complexity;
- select a model structure from a set of known competing models;
- identify functional dependencies;
- analyze efficiencies of numerical schemes.
Effective dimensions

Let $|u|$ be a cardinality of a set of variables $u$.

The effective dimension of $f(x)$ in superposition sense is the smallest integer $d_S$ such that

$$
\sum_{0<|u|<d_s} S_u \geq (1-\varepsilon), \ \varepsilon << 1
$$

It means that $f(x)$ is almost a sum of $d_S$-dimensional functions.

The effective dimension of $f(x)$ in truncation sense is the smallest integer $d_T$ such that

$$
\sum_{u \subseteq \{1,2,\ldots,d_T\}} S_u \geq (1-\varepsilon), \ \varepsilon << 1
$$

Example: $f(x) = \sum_{i=1}^{n} x_i \rightarrow d_S = 1, \ d_T = n$

$d_S$ does not depend on the order in which the input variables are sampled, $d_T$ - depends on the order $\rightarrow$ by reordering variables $d_T$ can be reduced.
Classification of functions

Type A. Variables are not equally important

\[
\frac{S^T_y}{n_y} \gg \frac{S^T_z}{n_z} \iff d_T \ll n
\]

Type B, C. Variables are equally important

\[
S_i \approx S_j \iff d_T \approx n
\]

Type B. Dominant low order indices

\[
\sum_{i=1}^{n} S_i \approx 1 \iff d_S \ll n
\]

Type C. Dominant higher order indices

\[
\sum_{i=1}^{n} S_i \ll 1 \iff d_S \approx n
\]
# Classification of functions

<table>
<thead>
<tr>
<th>Function type</th>
<th>Description</th>
<th>Relationship between $S_i$ and $S_i^T$</th>
<th>$d_T$</th>
<th>$d_S$</th>
<th>QMC is more efficient than MC</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>a few dominant variables</td>
<td>$\frac{S_y^T}{n_y} \gg \frac{S_z^T}{n_z}$</td>
<td>$&lt;&lt; n$</td>
<td>$&lt; n$</td>
<td>Yes</td>
</tr>
<tr>
<td>B</td>
<td>no unimportant subsets; important low-order interaction terms</td>
<td>$S_i \approx S_j, \forall i, j$ $S_i / S_i^T \approx 1, \forall i$</td>
<td>$\approx n$</td>
<td>$&lt;&lt; n$</td>
<td>Yes</td>
</tr>
<tr>
<td>C</td>
<td>no unimportant subsets; important high-order interaction terms</td>
<td>$S_i \approx S_j, \forall i, j$ $S_i / S_i^T &lt;&lt; 1, \forall i$</td>
<td>$\approx n$</td>
<td>$\approx n$</td>
<td>No</td>
</tr>
</tbody>
</table>

Majority of problems in finance either have low effective dimensions or effective dimensions can be reduced by using special techniques.
Effective dimensions. Test case

\[ f(x_1, x_2) = x_1 x_2, \]
\[ f_1(x_1) = \frac{1}{2} x_1 - \frac{1}{4}, \]
\[ f_2(x_2) = \frac{1}{2} x_2 - \frac{1}{4}, \]
\[ f_{12}(x_1, x_2) = x_1 x_2 - \frac{1}{2} x_1 - \frac{1}{2} x_2 + \frac{1}{4} \]

\[ S_1 = S_2 = \frac{3}{7}, \quad S_{12} = \frac{1}{7} \]
\[ S_1 + S_2 = \frac{6}{7} = 0.85 \approx 1 \rightarrow \]

effective dimensions:
\[ d_T = 2, \quad d_S = 1 \rightarrow \text{second order interactions are not important} \]
\[ f(x_1, x_2) \approx f_0 + f_1(x_1) + f_2(x_2) \]
Global Sensitivity Analysis of standard and Brownian Bridge discretizations

Apply global SA to payoff function

\[ P_A(\{Z_i\}) = \max(\overline{S}(\{Z_i\}) - K, 0), \{Z_i\}, i = 1, n \]

![Graph showing log of total sensitivity indices versus time step number.](image)

Log of total sensitivity indices versus time step number i.

Standard discretization - Si_total slowly decrease with i. Brownian bridge - Si_total of the first few variables are much larger than those of the subsequent variables.
Global Sensitivity Analysis of the standard discretization

For the standard discretization

\[ C(K,T) = e^{-rT} \int_{H^n} \max[0, \left( \prod_{i=1}^{n} S_0 \exp[(r - \frac{\sigma^2}{2})t_i + \sigma \sqrt{\frac{T}{n} \sum_{j=1}^{i} \Phi^{-1}(u_j)}] \right)^{1/n} ] - K] du_1...du_n \]

\[ Y = f(\tilde{x}), \quad \tilde{x} = (x_1, x_2, ..., x_k) - uncertain \ \text{parameters} \]

Apply global SA to a payoff function :

\[ f(\tilde{x}) = \max(0, \left[ \prod_{i=1}^{n} S_0 \exp[(r - \frac{\sigma^2}{2})t_i + \sigma \sqrt{\frac{T}{n} \sum_{j=1}^{i} \Phi^{-1}(x_j)}] \right]^{1/n} \]
Global Sensitivity Analysis of two algorithms at different n (time steps)

Table 1. Sum of the first order sensitivity indices for the standard and Brownian Bridge discretizations for various n

<table>
<thead>
<tr>
<th>n</th>
<th>Option Value</th>
<th>$\sum_i S_i$ Stand.</th>
<th>$\sum_i S_i$ BB</th>
</tr>
</thead>
<tbody>
<tr>
<td>8</td>
<td>4.13</td>
<td>0.102</td>
<td>0.41</td>
</tr>
<tr>
<td>16</td>
<td>3.94</td>
<td>0.042</td>
<td>0.38</td>
</tr>
<tr>
<td>32</td>
<td>3.84</td>
<td>0.022</td>
<td>0.37</td>
</tr>
</tbody>
</table>

The effective dimensions

Standard approximation: $d_T \approx n$
$d_S > 2$

Brownian Bridge approximation: $d_T \approx 2$
$d_S \approx 2$
Option pricing: Why Brownian Bridge is more efficient than standard discretization in the case of QMC?

- (A) The initial coordinates of LDS are much better distributed than the later high dimensional coordinates.
- (B) Low dimensional projections of low discrepancy sequences are better distributed than higher dimensional projections.

The Brownian bridge discretization

1) well distributed coordinates are used for important variables and higher not so well distributed coordinates are used for far less important variables (A)
2) the effective dimension $d_s$ is also reduced – approximation function becomes additive (B)

The standard construction does not account for the specifics of LDS distribution properties.

Application of QMC with the Brownian bridge discretization results in the $10^2$-$10^6$ time reduction of CPU time compared with MC!
Cox, Ingersoll and Ross interest rate model

\[ dy_t = (\alpha + \beta y_t) dt + \sigma \sqrt{y_t} dW_t \]

where \( \alpha > 0, \beta > 0, \sigma > 0 \)

Data: the 9-month Euribor interest rate daily time series using 250 daily observations, starting at 30 Dec 1998

The generalised method of moments (GMM) estimation is used to obtain \( \hat{\alpha}, \hat{\beta} \) and \( \hat{\sigma} \) estimates.

\[ \hat{\alpha} = 1e-005, \quad \hat{\beta} = 0.1109, \quad \hat{\sigma} = 0.1929 \]

Euler discretisation:

\[ y_{t+\Delta t} - y_t = (\alpha + \beta y_t) \Delta t + \sigma \sqrt{y_t} \sqrt{\Delta t} \varepsilon_t \]
CIR model: 50 day forecasts of the 9-month Euribor rate

MC and QMC estimators with using standard Euler and Milstein schemes

QMC produces much smoother and much faster convergence than MC
RMSE convergence for 250-days forecasting horizon

\[ \varepsilon = \left( \frac{1}{K} \sum_{k=1}^{K} (I - I_k^N) \right)^{1/2} \]

\[ \varepsilon \sim N^{-\alpha}, \quad 0 < \alpha < 1 \]

QMC estimator displays much faster convergence compared to MC regardless of the forecasting horizon (dimension)
RMSE convergence for 250-days forecasting horizon

$$\varepsilon = \left( \frac{1}{K} \sum_{k=1}^{K} (I - I_N^k)^2 \right)^{1/2}$$

$$\varepsilon \sim N^{-\alpha}, \ 0 < \alpha < 1$$

There is no notable advantage of Brownian bridge over the standard Euler scheme
Global Sensitivity Analysis of Standard and Brownian Bridge discretizations (CIR model)

Apply global SA to a function $y_T(Z_1, Z_2, ... , Z_n)$, $n$ is a number of days

Log of total sensitivity indices versus time step number $i$.

Standard discretization - $S_{i\_total}$ are constant. Brownian bridge - $S_{i\_total}$ of the first variable is much larger than those of the subsequent variables.
Global Sensitivity Analysis of Standard and Brownian Bridge discretizations (CIR model)

<table>
<thead>
<tr>
<th>$n$</th>
<th>$\sum S_i$ Standard</th>
<th>$\sum S_i$ BB</th>
</tr>
</thead>
<tbody>
<tr>
<td>32</td>
<td>0.99</td>
<td>0.99</td>
</tr>
<tr>
<td>64</td>
<td>0.99</td>
<td>0.99</td>
</tr>
<tr>
<td>128</td>
<td>0.99</td>
<td>0.99</td>
</tr>
</tbody>
</table>

Standard approximation: the effective dimension $d_T \approx n$
The effective dimension $d_S = 1$

Brownian Bridge approximation: the effective dimension $d_T \approx 2$
The effective dimension $d_S = 1$

The low effective dimension $d_S$ for both schemes = 1, hence QMC efficiency can not be further improved by changing sampling strategy
Control Variate Method

We define the control variate by

\[ X(\lambda) = X - \lambda C. \]

The control \( C \) is sampled along with \( X \) and centered at zero, i.e., \( E\{C\} = 0 \). \( \lambda \) is the control parameter so that

\[ E\{X(\lambda)\} = E\{X\}. \]

### Variance Decomposition

\[ Var(X(\lambda)) = \sigma_X^2 - 2\lambda \sigma_X \sigma_C \rho_{XC} + \lambda^2 \sigma_C^2, \]

Where \( \rho_{XC} \) is the correlation between \( X \) and \( C \).

The optimal control parameter

\[ \lambda^* = \frac{\sigma_X}{\sigma_C} \rho_{XC} = \frac{Cov(X, C)}{Var(X)} \]

is obtained by minimizing the variance \( Var(X(\lambda)) \).
Control Variate Method with Monte Carlo method

The basic Monte Carlo estimator is

\[ E \{ X \} \approx S_N := \frac{1}{N} \sum_{i=1}^{N} X^{(i)}, \]

The control variate estimator is defined by

\[ S_N^{\lambda^*} = \frac{1}{N} \sum_{i=1}^{N} (X^{(i)} - \lambda^* C^{(i)}). \]

The variance reduction ratio is

\[ \frac{Var(S_N)}{Var(S_N^{\lambda^*})} = \frac{1}{1 - \rho_{XC}^2} > 1, \text{ if } \rho_{XC} \neq 0 \]
Monte Carlo Pricing with Martingale Control Variate (MCV)

Pricing a European option by a MCV method leads to

\[
P(0, S_0, \sigma_0) \approx \frac{1}{N} \sum_{i=1}^{N} \left[ e^{-rT} H(S_T^{(i)}) - \lambda M^{(i)}(\tilde{p}) \right],
\]

where

\[
M(\tilde{p}) = \int_0^T e^{-rs} \frac{\partial \tilde{p}}{\partial x} (s, S_S, \sigma_S) \sigma_S S_S dW^*_S
\]

is a martingale with \( \tilde{P} \) being an approximation of \( P \).

\( \lambda \) : control parameter (empirically chosen as 1)

\( M(\tilde{P}) \) : martingale control.
Homogenization Method

Fouque and H. (07) * use the homogenized Black-Scholes price $P_{BS}(t, S_t; \sigma)$ to construct a martingale control $M^{(i)}(P_{BS})$.

The effective variance $\overline{\sigma^2}$ is defined as the averaging of variance function w.r.t. the invariant distribution of the volatility process.

- Martingale control variate method is very general for option pricing problems. The control is related to the accumulative value of delta-hedging portfolios.
- Martingale control variate corresponds to a smoother payoff function so that QMC methods can be effective.

European Options Pricing by Variance Reduction

blue: Basic MC samples

green: MCV samples
One-Factor Stochastic Volatility model

\[ dS_t = rS_t dt + \sigma_t S_t dW_{0t}^* \]
\[ \sigma_t = \exp \left( Y_t / 2 \right) \]
\[ dY_t = \alpha (m - Y_t) dt + \beta \left( \rho_1 dW_{0t}^* + \sqrt{1 - \rho_1^2} dW_{1t}^* \right) \]

\( Y_t \) is an Ornstein-Uhlenbeck process
\[ \alpha = 1 / \varepsilon, \quad \beta = \nu \sqrt{2 / \varepsilon} \]

so that \( Y_t \) is varying on the time scales \( \varepsilon \)
European Call Option Prices. Sobol’ LDS

\( m=-2.5, \; \nu=1, \; \rho=-0.7; \; Y_0=-2.5, \; S_0=110; \)
\( r=0.05; \; K=100, \; T=1 \)

<table>
<thead>
<tr>
<th>( \varepsilon )</th>
<th>( SE^{BMC} (mean) )</th>
<th>( SE^{MCV} (mean) )</th>
<th>( \text{Ratio}^{MCV} )</th>
</tr>
</thead>
<tbody>
<tr>
<td>1/50</td>
<td>0.352 (24.115)</td>
<td>0.039 (24.348)</td>
<td>81</td>
</tr>
<tr>
<td>1/10</td>
<td>0.303 (23.340)</td>
<td>0.038 (23.400)</td>
<td>63</td>
</tr>
<tr>
<td>10</td>
<td>0.268 (20.771)</td>
<td>0.019 (20.964)</td>
<td>193</td>
</tr>
<tr>
<td>50</td>
<td>0.276 (20.488)</td>
<td>0.017 (20.711)</td>
<td>267</td>
</tr>
<tr>
<td>100</td>
<td>0.279 (20.428)</td>
<td>0.017 (20.660)</td>
<td>277</td>
</tr>
</tbody>
</table>

*BCM – basic MC; Ratio\textsuperscript{MCV} – variance reduction ratio

Time Discretization – 128 steps; Total number of paths - 8192
Global Sensitivity Analysis of standard and Martingale Control Variate (MCV) approximations

Apply global SA to payoff function

\[ P(0, S_0, \sigma_0) \approx e^{-rT} H(S_T) - M(\tilde{P}). \]

\[ S_T = S_T(\{Z_i\}), \{Z_i\}, i = 1, n \]

Effective dimensions:
Standard: \( d_T \approx n \)
\( d_S = 2 \)

MCV: \( d_T \approx 2 \)
\( d_S = 1 \)

Log of total sensitivity indices versus time step number \( i \).

Standard discretization - \( S_{i, total} \) slowly decrease with \( i \). MCV - \( S_{i, total} \) of the first few variables are much larger than those of the subsequent variables.
Derivative based Global Sensitivity Measures and "Global Greeks"

Consider the differentiable function \( f(x_1, \ldots, x_n) \) defined in the unit hypercube.

\[
\nu_i = \int_{H^n} \left( \frac{\partial f}{\partial x_i} \right)^2 \, dx
\]

Unlike partial derivative \( \left( \frac{\partial f}{\partial x_i} \right) \) which is defined at the nominal point and by definition is local, measure \( \nu_i \) is global.

How to use in finance:

**Static hedge:** Consider plain vanilla option \( C(S, \sigma, r, t) \). Delta hedging strategy is based on a local sensitivity measure \( \Delta(S, \sigma, r, t) = \partial C / \partial S \). Parameters \( (S, \sigma, r) \) are uncertain.

Define "global delta" \( \Delta = \left( \int (\partial C / \partial S)^2 \, dS \sigma \, dr \, dt \right)^{1/2} \) and use it for a static hedge.
Similarly we can define

\[ \gamma_i = \int_{H^n} \left( \frac{\partial^2 f}{\partial x_i^2} \right)^2 dx \]

And use it to define “global gamma”

\[ \Gamma = \left( \int \left( \frac{\partial^2 C}{\partial S^2} \right)^2 dSd\sigma drdt \right)^{1/2} \]

Similarly we can define other “global greeks”

There is a link between variance based and derivative based measures

**Theorem:** \( S_i^{tot} \leq \frac{V_i}{\pi^2 \sigma^2} \)
Summary

Global Sensitivity Analysis is a general approach for uncertainty, complexity reduction and structure analysis of non-linear models. It can be widely applied in finance.

In Sobol's algorithm direction numbers is a key component to its efficiency. The Sobol Sequence generator satisfying uniformity properties A and A' has superior performance over other generators.

Quasi MC methods based on Sobol' sequences outperform MC regardless of nominal dimensionality for problems with low effective dimensions.

Great success of Quasi MC methods in finance is explained by problems structure - low effective dimensions which can be reduced by techniques such as Brownian bridge, PCA, control variate techniques
Acknowledgments

Prof. Sobol’

Dario Cziraky, Barclays Capital

Sean Han, NTHU, Taiwan