Handling Discontinuity in Pricing and Hedging: QMC with Dimension Reduction

Xiaoqun Wang

Department of Mathematical Sciences
Tsinghua University, Beijing 100084, China
Email: xwang@math.tsinghua.edu.cn

Joint work with K. S. Tan

Outline

1. Introduction

Many problem in mathematical finance can be formulated as high-dimensional integration

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

Examples:

- Pricing financial derivatives;
- Evaluating Greeks (risk measures);

The dimension can be huge (hundreds or thousands)! Source of dimensionality:

- Number of time steps in discretization;
- Number of state variables (risk factors).

Only in rare cases do explicit solutions exist (Black-Scholes formula). In most cases we have to use numerical methods (PDE, Simulation).

1. Introduction

Monte Carlo methods:

$$Q_{n,d}(f) = \frac{1}{n}\sum_{k=1}^{n} f(\mathbf{x}_k).$$

- The points $\mathbf{x}_k$ are i.i.d. samples from $U([0,1]^d)$.
- Root Mean square error:

$$\sqrt{E[I_d(f) - Q_{n,d}(f)]^2} = \frac{\sigma(f)}{\sqrt{n}}.$$

- Convergence order: $O(n^{-1/2})$
- Advantage: Independent of dim.
- Disadvantages: Very slow! Error estimate is only probabilistic.

1. Introduction

Quasi-Monte Carlo Methods:

$$Q_{n,d}(f) = \frac{1}{n}\sum_{k=1}^{n} f(\mathbf{x}_k),$$

where the points $\mathbf{x}_k$ are deterministic and cover the unit cube more uniformly — low discrepancy points. Examples:

- Digital points: Halton, Sobol', Faure, Niederreiter, ...
- Good lattice points: e.g. Korobov, Hua-Wang, Sloan, ...
- Convergence order depends on the function class, usually

  $$O\left(n^{-1}(\log n)^d\right)$$

  for functions of bounded variation, or better for smoother functions.

1. Introduction

Advantages of QMC:

- Asymptotically converges faster than MC:

$$QMC: \; O\left(n^{-1}(\log n)^d\right) \quad \text{vs.} \quad \text{MC:} \; O(n^{-1/2});$$

- Perfect one-dim projections;
- Better uniformity in initial dimensions.

Classical QMC theory does not support the practical superiority of QMC over MC in high dim (due to the factor $(\log n)^d$).

However, ...
Surprisingly, it was found in the 1990s that **QMC beat MC** for a 360-dim MBS problem (Traub & Paskov 1995, many others ...).

A Big Challenge:

- **Explain the great success of QMC** for very high dim problems!

1. Introduction

Explaining the great success of QMC:

- The theory of weighted function spaces:
  - Sloan and Wozniakowski (1998; non-constructive);
  - Hickernell and Wozniakowski (2000; non-constructive);
  - Wang (2002; Constructive, digital sequences);
  - Kuo (2003; Constructive, good lattice points);
  - Many others (Dick et al. 2006, ...)

  It is possible to obtain optimal convergence order independently of dim, if the weights decay sufficiently fast.

- Effective dimension:
  - Caflisch, Morokoff, Owen. (1997, J. Comput. Finance);
  - Wang and Fang (2003, J. Complexity);
  - Wang and Sloan (2005, SIAM J. Sci Comput.);
  - Liu and Owen (2006, JASA).

  It is well accepted that QMC works well because eff dim is small.

1. Introduction

Limitations of QMC: Two factors could significantly affect QMC:

- **Dimension:**
  - Convergence of QMC depends exponentially on dim;
  - High-dim projections exhibit patterns (lattice structure) and are no more uniform than random points for practical n.
- **Smoothness:**
  - Success of QMC in finance has been confined predominately on pricing derivative with reasonably smooth (continuous) payoff.
  - Discontinuity is common and is a major challenge in finance.

Good News: Not all kinds of discontinuities are harmful for QMC.

QMC-friendly discontinuity: Discontinuity parallels to axes is QMC-friendly, since high precision can still be expected (QMC points are found by minimizing discrepancy, which is based on rectangles with sides parallel to axes).

1. Introduction

More Challenges: **Overcoming the limitations of QMC**!

- Reducing eff dim (Path generation method, PGM)
  - Brownian bridge (BB) (Moskowitz, Caflisch 1996; Caflisch et al. 1997);
  - Principal Component Analysis (PCA) (Acworth et al. 1998);
  - Linear transformation (Imai and Tan 2006); ...

  There are many successful applications. However, BB or PCA does not offer a consistent advantage in QMC! (for digital options (Papageorgiou 2002); for weighted Asian options (Wang and Sloan 2010); ...)
  Interesting Question: Why BB or PCA may do not work well?

- Smoothing the function
  - Traditional idea: Making the function differentiable or continuous by smoothing sharp edges (conditional sampling, kernel smoothing).
  - Non-financial QMC literature: smoothing may recover the superiority of QMC (Moskowitz and Caflisch 1996, Wang 2000);
  - Financial QMC literature: Hardly any attention ...

Is it possible to reduce eff dim and to smooth function simultaneously?

1. Introduction

This paper focuses on:

- Why BB or PCA does not offer a consistent advantage in QMC?
- How to deal with discontinuity in finance?

2. Path Generation Methods (PGMs)

Derivative pricing and high-dimensional integrals

- Discounted Payoff: $g(S_{t_1},\dots,S_{t_d})$, where $S_{t_1},\cdots,S_{t_d}$ are the prices of the asset at $t_j = j\Delta t, j = 1,\dots,d, \Delta t = T/d$, T is expiration date.
- Black-Scholes model:

$$dS_t = rS_t dt + \sigma S_t dB_t.$$

- Based on risk-neutral valuation, the value of derivative at time 0 is

$$\mathbb{E}[g(S_{t_1},\dots,S_{t_d})],$$

  where $E[\cdot]$ is expectation under risk-neutral measure.
- The analytical solution to the SDE:

$$S_t = S_0 \exp\left((r-\frac{\sigma^2}{2})t + \sigma B_t\right).$$

  Simulating stock prices reduces to simulating Brownian motion (BM).

2. Path Generation Methods (PGMs)

The random vector $(B_{t_1}, \ldots, B_{t_d})^T =: \mathbf{x}$ is normally distributed with mean zero and covariance matrix $\mathbf{C}$:

$$\mathbf{C} = \Delta t \begin{pmatrix} 1 & 1 & 1 & \cdots & 1 \\ 1 & 2 & 2 & \cdots & 2 \\ 1 & 2 & 3 & \cdots & 3 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & 2 & 3 & \cdots & d \end{pmatrix}. \tag{1}$$

The discounted payoff function can be written as

$$g(S_{t_1}, \ldots, S_{t_d}) = G(\mathbf{x}), \mathbf{x} \sim N(\mathbf{0}, \mathbf{C}).$$

The derivative price can be written as a Gaussian integral:

$$V(G) = \mathbb{E}(G(\mathbf{x})) = \frac{1}{(2\pi)^{d/2}\sqrt{\det \mathbf{C}}} \int_{\mathbb{R}^d} G(\mathbf{x})\, e^{-\frac{1}{2}\mathbf{x}^T \mathbf{C}^{-1} \mathbf{x}}\, d\mathbf{x}. \tag{2}$$

Formally, we have

$$V(G) = \int_{[0,1)^d} G(A\Phi^{-1}(\mathbf{u}))d\mathbf{u} \approx \frac{1}{n} \sum_{k=1}^{n} G(A\Phi^{-1}(\mathbf{u}_k)).$$

2. Path Generation Methods (PGMs)

The Constructions of BM:

- Standard construction (STD) generates BM sequentially. Given $B_0 = 0$,

$$B_{t_j} = B_{t_{j-1}} + \sqrt{\Delta t}\ z_j, z_j \sim N(0,1), \quad j = 1, \ldots, d,$$

The generating matrix A is the *Choleskey decomposition* of $\mathbf{C}$:

$$A = A^{\mathrm{STD}} := \sqrt{\Delta t} \begin{pmatrix} 1 & 0 & \cdots & 0 \\ 1 & 1 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 1 & 1 & \cdots & 1 \end{pmatrix}.$$

2. Path Generation Methods (PGMs)

- Brownian Bridge (BB) Construction: Given a past value B_{t_i} and a future value B_{t_k}, the value B_{t_j} between them (with $t_i < t_j < t_k$) can be simulated according to BB formula.
  Assume $d = 2^\ell$ (ℓ is a non-negative integer). Given $B_0 = 0$, the BM is generated at times in order $T, T/2, T/4, 3T/4, \ldots$

$$
\begin{aligned}
B_T &= \sqrt{T}\, z_1; \\
B_{T/2} &= \frac{1}{2}(B_0 + B_T) + \sqrt{\frac{T}{4}}\, z_2; \\
B_{T/4} &= \frac{1}{2}(B_0 + B_{T/2}) + \sqrt{\frac{T}{8}}\, z_3; \\
B_{3T/4} &= \frac{1}{2}(B_{T/2} + B_T) + \sqrt{\frac{T}{8}}\, z_4; \\
&\vdots
\end{aligned}
$$

2. Path Generation Methods (PGMs)

- Principal Component Analysis (PCA) Construction:
  The corresponding generating matrix is

$$A := A^{\mathrm{PCA}} = (\sqrt{\lambda_1}\,\mathbf{v}_1, \sqrt{\lambda_2}\,\mathbf{v}_2, \ldots, \sqrt{\lambda_d}\,\mathbf{v}_d)$$

  where $\lambda_1 \geq \lambda_2 ...$, are the eigenvalues of $\mathbf{C}$ and $\mathbf{v}_1, \mathbf{v}_2, ...$ are the corresponding unit-length eigenvectors.

Optimality of PCA: PCA provides an optimal lower-dimensional approximation to the random vector $\mathbf{x}$. Also, PCA is optimal in the sense of explained variability.

2. Path Generation Methods (PGMs)

Remark:

- In MC, the choice of decomposition is unimportant: the mean square error of MC is unchanged.
- However, in QMC the PGM is crucial

  Since different PGMs may lead to

  - Different dimension structure (eff dim);
  - Different smoothness property;
  - In particular, they may lead to different discontinuity structure, upon which QMC depends crucially.
- According to the equivalent principle (Wang and Sloan 2010), there is no universally optimal PGM. A good PGM should depend on the payoff function.

2. Path Generation Methods (PGMs)

ANOVA decomposition

Any square-integrable function $f(\mathbf{x})$ can be written as

$$f(\mathbf{x}) = f_\emptyset + \sum_{j=1}^{d} f_{\{j\}}(x_j) + \sum_{1\le i<j\le d} f_{\{i,j\}}(x_i, x_j) + \cdots . \qquad (3)$$

This decomposition is ANOVA (and unique) if

$$\int_0^1 f_u(\mathbf{x}) dx_j = 0, \quad \forall j \in u.$$

Properties of ANOVA decomposition:

- **Orthogonality**: $\int_{[0,1]^d} f_u(\mathbf{x}) f_v(\mathbf{x}) d\mathbf{x} = 0, \quad u \neq v.$
- **Decomposition of the variance**:

$$\sigma^2(f) = \sum_{\emptyset \neq u \subseteq \{1,\dots,d\}} \sigma_u^2(f),$$

where $\sigma^2(f)$ and $\sigma_u^2(f)$ are the variance of f and f_u (here $\mathbf{x} \sim U[0,1]^d$).

2. Path Generation Methods (PGMs)

The effective dimensions (Caflisch et al. 1997)

Def. The effective dimension of f in truncation sense is the smallest integer d_t such that

$$\sum_{\emptyset \neq u \subseteq \{1,\ldots,d_t\}} \sigma_u^2(f) \geq p\, \sigma^2(f) \quad \text{(often} \quad p = 0.99\text{)}.$$

Def. The effective dimension of f in superposition sense is the smallest integer d_s, such that

$$\sum_{0<|u|\leq d_s} \sigma_u^2(f) \geq p\, \sigma^2(f).$$

2. Path Generation Methods (PGMs)

Def. The variance ratio of order-ℓ is defined as

$$R_\ell = \frac{1}{\sigma^2(f)} \sum_{|u|=\ell} \sigma_u^2(f), \quad \ell = 1, \ldots, d. \tag{4}$$

Def. The degree of additivity, R_1, is the sum of all first-order sensitivity indices, which measures the inherent additive structure.

Def. The mean dimension (in superposition sense) is defined as (Owen)

$$d_{ms} = \sum_{\ell=1}^{d} \ell \, R_\ell. \tag{5}$$

If $R_1 \approx 1$ or $d_{ms} \approx 1$, then the function is highly additive for which QMC is especially suitable.
These measures can be computed numerically (Wang and Fang 2003, Liu and Owen 2006).

3. No PGM Offers a Consistent Advantage

Warm-up numerical experiments:
No PGM Offers a Consistent Advantage!

Example 3.1 Digital option (Papageorgiou 2002):

$$g_1(\mathbf{S}) = \frac{1}{d} \sum_{j=1}^{d} S_j \, \mathbf{I}_{\{S_j > S_{j-1}\}}(\mathbf{S}), \qquad (6)$$

where $\mathbf{I}_{\{\cdot\}}(\mathbf{S})$ is an indicator function.

Note: All the model parameters are omitted below (they can be found in the original paper)!

3. No PGM Offers a Consistent Advantage

Table A1: The VRFs for digital options (Example 3.1)

	MC	Sobol'			Korobov		
d	STD	STD	BB	PCA	STD	BB	PCA
2	1	**2262**	50	110	**2857**	176	188
16	1	**46**	1	3	**971**	17	8
64	1	**66**	5	11	**545**	17	9
128	1	**60**	13	13	**296**	24	11

Table A2: Eff dim-related characteristics for digital options

	Degree of Additivity R_1			Mean Dimension d_{ms}		
d	STD	BB	PCA	STD	BB	PCA
2	**0.9910**	0.79	0.80	**1.0090**	1.21	1.20
16	**0.9856**	0.83	0.82	**1.0129**	1.43	2.11
64	**0.9854**	0.89	0.87	**1.0133**	1.31	2.64
128	**0.9853**	0.93	0.93	**1.0119**	1.23	2.71

3. No PGM Offers a Consistent Advantage

- The performances of STD, BB, PCA in QMC are very different.
- QMC-based STD is remarkably effective, irrespective of nominal dim and low discrepancy point sets.
- In QMC, BB and PCA behave consistently worse than STD, not only when dim is large, but also when dim is as small as 2.
- Degree of additivity and mean dim under STD are close to one, implying the resulting function is highly additive.
- Both BB and PCA lead to smaller degree of additivity and larger mean dim. Higher order ANOVA part cannot be neglected.
- The evidences from VRFs, R_1 and d_{ms} consistently support the superiority of STD, relative to BB and PCA in QMC.

In QMC, the digital option is “STD-friendly” while unfriendly for BB and PCA.

3. No PGM Offers a Consistent Advantage

Example 3.2 Asian options with knock-out feature at the maturity (Glasserman et al. 1999).

$$g_2(\mathbf{S}) = \max(S_A - K, 0)\, \mathbf{I}_{\{S_d \le H\}}(\mathbf{S}), \tag{7}$$

where

$$S_A = \sum_{j=1}^{d} w_j\, S_j \quad \text{with} \quad \sum_{j=1}^{d} w_j = 1, \tag{8}$$

(with equal weights $w_j = 1/d$) and K is strike price and H is barrier.

If at expiration the price of the underlying asset is below the barrier H, then the option pays $\max(S_A - K, 0)$ as with an ordinary Asian call option, but if the final price is above the barrier H, the option pays nothing.

3. No PGM Offers a Consistent Advantage

Table A3: The VRFs for Example 3.2

		MC	Sobol'			Korobov		
d	H	STD	STD	BB	PCA	STD	BB	PCA
128	120	1	1	**105**	5	1	**44**	4
	140	1	2	**213**	7	2	**115**	7
	160	1	2	**257**	16	3	**89**	15

Table A4: Eff dim-related characteristics for Example 3.2

		Degree of Additivity R_1			Mean Dimension d_{ms}		
d	H	STD	BB	PCA	STD	BB	PCA
128	120	0.40	**0.71**	0.49	16.61	**1.32**	3.77
	140	0.30	**0.86**	0.64	10.58	**1.15**	2.69
	160	0.71	**0.91**	0.84	5.69	**1.10**	1.79

3. No PGM Offers a Consistent Advantage

- The VRFs again reveal the significant difference in performance among STD, BB and PCA in QMC.
- BB is the most effective and outperforms remarkably STD and PCA. STD or PCA in QMC behaves no better or only slightly better than MC.
- BB leads to the largest degree of additivity and the smallest mean dim (while STD is the worst). This provides additional support for the superiority of BB.
- The option in this example is “BB-friendly”, but is unfriendly for STD and PCA.

3. No PGM Offers a Consistent Advantage

Example 3.3 Asian options with a modified knock-out feature).
The payoff function of the option is defined as

$$g_3(\mathbf{S}) = \max(S_A - K, 0)\, \mathbf{I}_{\{\overline{S} \leq H\}}(\mathbf{S}) \quad \text{with} \quad \overline{S} = \prod_{j=1}^{d} S_j^{v_j}, \qquad (9)$$

where $\overline{S}$ is a specific weighted geometric average of the asset prices with respect to the weights v_j's, defined as follows :
(... see the original paper)

3. No PGM Offers a Consistent Advantage

Table A5: The VRFs for Example 3.3

		MC	Sobol'			Korobov		
d	H	STD	STD	BB	PCA	STD	BB	PCA
128	120	1	1	6	**3101**	1	4	**815**
	140	1	2	17	**2525**	3	16	**1727**
	160	1	4	52	**1474**	8	38	**1249**

Table A6: The eff dim-related characteristics for Example 3.3

		Degree of Additivity R_1			Mean Dimension d_{ms}		
d	H	STD	BB	PCA	STD	BB	PCA
128	120	0.17	0.36	**0.983**	16.92	2.86	**1.02**
	140	0.48	0.74	**0.994**	5.74	1.59	**1.01**
	160	0.79	0.90	**0.996**	2.04	1.17	**1.01**

3. No PGM Offers a Consistent Advantage

- PCA in QMC has the most remarkable success. STD or BB in QMC behaves no better or only slightly better than MC.
- PCA leads to the largest degree of additivity and the smallest mean dim. STD is the worst.
- The option in this example is “PCA-friendly”, but is unfriendly for STD and BB.

In summary,

- We provide “STD-”, “BB-” and “PCA-friendly” financial options, respectively (while unfriendly for others).
- No PGM offers a consistent advantage. A PGM that works very well in one example may perform badly in another.
- WHY?

4. Impact of PGMs on the Nature of Discontinuity

It is instructive to first analyze the discontinuity of the indicator function

$$h_1(\mathbf{S}) = \mathbf{I}_{\{S_j > S_{j-1}\}}(\mathbf{S}), \text{for some } j = 1, \ldots, d. \tag{10}$$

Its jumps occur when

$$S_{j-1} = S_j.$$

This condition is equivalent to (under any PGM)

$$B_{j-1} - B_j = \frac{\Delta t}{\sigma} \left(r - \sigma^2/2 \right). \tag{11}$$

When the paths are generated by STD, we have

$$B_j = B_{j-1} + \sqrt{\Delta t}\, z_j, \quad z_j \sim N(0,1).$$

Thus under STD the condition (11) is equivalent to

$$z_j = -\frac{\sqrt{\Delta t}}{\sigma} \left(r - \sigma^2/2 \right) =: C_1, \text{ or } u_j = \Phi(C_1). \tag{12}$$

An important observation is that when the paths are generated by STD, the discontinuity occurs only on axis-parallel hyperplane, which is QMC-friendly.

4. Impact of PGMs on the Nature of Discontinuity

On the other hand, when the paths are generated by BB or PCA, the discontinuity occurs on a manifold with a nonlinear equation ...
(see the original paper)

The discontinuity is in general no longer QMC-friendly and the function has infinite variation.

The payoff function of a digital option in Example 3.1 is essentially the sum of the indicator functions of the form (10).

4. Impact of PGMs on the Nature of Discontinuity

Theorem

Under Black-Scholes model, consider a PGM with a generating matrix A. Suppose the payoff of a financial derivative $G(\mathbf{x})$ with $\mathbf{x} \sim N(\mathbf{0}, \mathbf{C})$ is transformed to $G(A\Phi^{-1}(\mathbf{u}))$ by $\mathbf{x} = A\mathbf{z}$ with $AA^T = \mathbf{C}$, followed by $\mathbf{z} = \Phi^{-1}(\mathbf{u})$.

For digital option (Example 3.1):

- *Under STD, the jumps of the function $G(A\Phi^{-1}(\mathbf{u}))$ occur on the axis-parallel hyperplanes*

$$u_j = \Phi(C_1), \quad j = 1, \ldots, d.$$

- *Under BB or PCA, its jumps occur on the manifold*

$$\beta_{j,1}\Phi^{-1}(u_1) + \cdots + \beta_{j,d}\Phi^{-1}(u_d) = C_2,$$

where $\beta_{j,k} = a_{j-1,k} - a_{j,k}$ for $k = 1, \ldots, d$.

4. Impact of PGMs on the Nature of Discontinuity

Theorem

- *For Asian option with a knock-out feature at maturity (Example 3.2):*
  - *Under BB, the jumps of the final function* $G(A\Phi^{-1}(\mathbf{u}))$ *occur on the axis-parallel hyperplane*

$$u_1 = \Phi(C_3/\sqrt{T}).$$

  - *Under STD or PCA, its jumps occur on manifold*

$$a_{d,1}\Phi^{-1}(u_1) + \cdots + a_{a,d}\Phi^{-1}(u_d) = C_3.$$

- *For the payoff in Example 3.3:*
  - *Under PCA, the jumps of the final integrand* $G(A\Phi^{-1}(\mathbf{u}))$ *occur on the axis-parallel hyperplane*

$$u_1 = \Phi(C_5).$$

  - *Under STD or BB, its jumps occur on manifold* $\mathbf{V}^T A\Phi^{-1}(\mathbf{u}) = C_4.$

4. Impact of PGMs on the Nature of Discontinuity

Remark:

- For the same problem different PGMs may lead to functions with different structure of discontinuity (QMC-friendly or not).
- QMC-friendly discontinuity corresponds to much higher QMC accuracy than for unfriendly discontinuity (as observed in Sec 3).
- This explains why different PGMs can have different performance for the same problem in QMC. In particular, this explains why BB or PCA does not offer a consistent advantage.
- Discontinuity property is a dominate factor that determines QMC accuracy. This calls for a significant concern.

5. Handling Discontinuity in Pricing and Hedging

Many functions in pricing and hedging may be written as

$$g(\mathbf{S}) = f(\mathbf{S})\,\mathbf{I}_{\{h(\mathbf{S})>0\}}(\mathbf{S}). \qquad (13)$$

- A binary (or digital) Asian option

$$g(\mathbf{S}) = \mathbf{I}_{\{S_A>K\}}(\mathbf{S}).$$

- The pathwise estimate for the delta of arithmetic Asian option

$$g(\mathbf{S}) = e^{-rT}\frac{S_A}{S_0}\,\mathbf{I}_{\{S_A>K\}}(\mathbf{S}).$$

Below we consider an indicator function of the form

$$\Lambda(\mathbf{x}) = \mathbf{I}_{\{h(\mathbf{q}^T\mathbf{x})<H\}}(\mathbf{x}), \quad \mathbf{x} = (x_1,\ldots,x_d)^T \sim N(\mathbf{0},\mathbf{C}), \qquad (14)$$

for some vector $\mathbf{q} = (q_1,\ldots,q_d)^T$ and function $h(\cdot)$ on $\mathbb{R}$.

5. Handling Discontinuity in Pricing and Hedging

Theorem

Let $\mathbf{C}$ be a $d \times d$ positive definite matrix and let A_0 be a fixed decomposition matrix such that $A_0 A_0^T = \mathbf{C}$. Suppose that the indicator function $\Lambda(\mathbf{x})$ has the form:

$$\Lambda(\mathbf{x}) = \mathbf{I}_{\{h(\mathbf{q}^T\mathbf{x}) < H\}}(\mathbf{x}), \quad \mathbf{x} = (x_1, \ldots, x_d)^T \sim N(\mathbf{0}, \mathbf{C}). \tag{15}$$

If U is a $d \times d$ orthogonal matrix, whose first column $\mathbf{U}_1$ is given by

$$\mathbf{U}_1 = \frac{1}{D} A_0^T \mathbf{q}, \; D := \sqrt{\mathbf{q}^T \mathbf{C}\, \mathbf{q}}, \tag{16}$$

the remaining columns are arbitrary as long as they satisfy the orthogonality conditions, then by the transformation $\mathbf{x} = A_0 U\mathbf{z}$, the function $h(\mathbf{q}^T\mathbf{x})$ involved in $\Lambda(\mathbf{x})$ is transformed to a function depending only on the first component of $\mathbf{z}$:

$$h(\mathbf{q}^T\mathbf{x}) = h(Dz_1).$$

5. Handling Discontinuity in Pricing and Hedging

Theorem

Consequently, the indicator function $\Lambda(\mathbf{x})$ is transformed to

$$\Lambda(\mathbf{x}) = \mathbf{I}_{\{h(Dz_1) < H\}}(\mathbf{z}), \quad \mathbf{z} = (z_1, \ldots, z_d)^T \sim N(\mathbf{0}, I_d).$$

If $h(\cdot)$ is strictly increasing on $\mathbb{R}$ and if $\mathbf{I}_{\{h(Dz_1) < H\}}(\mathbf{z})$ is further transformed by $\mathbf{z} = \Phi^{-1}(\mathbf{u})$, then $\Lambda(\mathbf{x})$ is transformed to one-dim function:

$$\Lambda(\mathbf{x}) = \mathbf{I}_{\{u_1 < c\}}(\mathbf{u}), \quad \mathbf{u} = (u_1, \ldots, u_d)^T \sim U(0, 1)^d,$$

where $c = \Phi(D^{-1} h^{-1}(H))$ is a constant. Its discontinuities occur only on the axis-parallel hyperplane $u_1 = c$, which are QMC-friendly.

5. Handling Discontinuity in Pricing and Hedging

Remark:

- Theorem offers a new PGM (denoted by OT, Orthogonal Transformation) for simulating BM with generating matrix A_0U. The resulting function has discontinuities only on an axis-parallel hyperplane, which are QMC-friendly.
- Traditionally, PGMs are designed primarily to reduce eff dim. Here we present the idea to use PGM to "smooth" the function, opening a new way to improve QMC.
- Once the difficulty of discontinuity has been overcome (at least partially), good performance of QMC can be expected.

5. Handling Discontinuity in Pricing and Hedging

Example 5.1. We consider an option with a payoff

$$g(\mathbf{S}) = \max(S_A - K, 0)\, \mathbf{I}_{\{S_{j_0} > S_{i_0}\}}(\mathbf{S})$$

for some $0 \leq i_0 < j_0 \leq d$ (for computation $i_0 = d/4 - 1, j_0 = 3d/4 - 1$).

The indicator function has the form (15) studied in Theorem 4, corresponding to $\mathbf{q}$ with $q_{i_0} = 1$, $q_{j_0} = -1$ and $q_j = 0$ for $j \neq i_0, j_0$.

Use our new method, the resulting function has discontinuities only on an axis-parallel hyperplane, which are QMC-friendly.

On the contrary, STD, BB and PCA do not necessarily.

5. Handling Discontinuity in Pricing and Hedging

Table B1: The VRFs for Example 5.1

d	K	MC STD	Sobol' STD	Sobol' BB	Sobol' PCA	Sobol' OT	Korobov STD	Korobov BB	Korobov PCA	Korobov OT
128	90	1	4	25	11	**937**	5	5	4	**92**
	100	1	3	41	18	**694**	7	9	7	**64**
	110	1	2	59	39	**456**	5	10	23	**40**

Table B2: The eff dim-related characteristics for Example 5.1

d	K	Degree of Additivity R_1 STD	BB	PCA	OT	Mean Dimension d_{ms} STD	BB	PCA	OT
128	90	0.80 (0.06)	0.85 (0.76)	**0.88** (0.86)	0.87 (**0.9980**)	3.18	1.40	1.77	**1.14**
	100	0.73 (0.02)	0.86 (0.84)	**0.93** (0.93)	0.79 (**0.9968**)	2.44	1.29	1.43	**1.22**
	110	0.54 (0.02)	0.79 (0.88)	**0.97** (0.97)	0.67 (**0.9955**)	2.20	1.29	**1.20**	1.32

5. Handling Discontinuity in Pricing and Hedging

- OT leads to the significantly highest VRFs in all cases. This confirms the power of aligning the discontinuity with the axes.
- The other PGMs in QMC, behave only slightly better than MC.
- OT does not necessarily lead to the highest degree of additivity nor the lowest mean dim (PCA mostly has the greatest degree of additivity while OT mostly has the lowest mean dim). Under OT, the first two variables capture the most cumulative variance.
- OT is the most accurate method based on VRF. This suggests that among the two key factors which determine the performance of a PGM in QMC, the nature of discontinuity, as opposed to the eff dim, seems more important for this example.

5. Handling Discontinuity in Pricing and Hedging

Example 5.3. We consider a modification of the option (9)

$$g(\mathbf{S}) = \max(S_A - K, 0) \cdot \mathbf{I}_{\{S_G < H\}}(\mathbf{S}),$$

where

$$S_G = \prod_{j=1}^{d} S_j^{v_j} \text{ with } \sum_{j=1}^{d} v_j = 1,$$

is the weighted geometric average of the underlying asset prices with respect to the weights $v_1, \ldots, v_d$.
The indicator function $\mathbf{I}_{\{S_G < H\}}(\mathbf{S})$ has the form in Theorem 4, corresponding to the vector $\mathbf{q} = \mathbf{v}$. Consequently, Theorem 4 is applicable, leading to QMC-friendly jumps.

5. Handling Discontinuity in Pricing and Hedging

Table B5: The VRFs for Example 5.3

		MC	Sobol'				Korobov			
d	H	STD	STD	BB	PCA	OT	STD	BB	PCA	OT
128	120	1	1	7	29	**4207**	1	5	31	**3543**
	140	1	4	31	92	**2109**	5	22	115	**1927**
	160	1	5	146	718	**1101**	14	79	644	**899**

Table B6: The eff dim-related characteristics Example 5.3

		Degree of Additivity R_1				Mean Dimension d_{ms}			
d	H	STD	BB	PCA	OT	STD	BB	PCA	OT
128	120	0.23	0.38	0.85	**0.9994**	12.90	2.67	1.29	**1.0005**
	140	0.75	0.84	0.97	**0.9996**	3.10	1.32	1.04	**1.0004**
	160	0.81	0.92	0.99	**0.9999**	1.33	1.09	1.01	**1.0010**

5. Handling Discontinuity in Pricing and Hedging

- The superiority of OT in QMC is obvious with VRFs in thousands.
- Knock-out level has large impact on the performance of PGM. For OT, larger VRF is achieved for smaller knock-out level (implying that ...), while for other PGMs the inverse is true.
- Both the degree of additivity and the mean dim under OT are almost one, regardless of knock-out level or nominal dim. This implies that under OT the function is essentially one-dim.
- In comparison, the resulting function using STD is far from being close to additive nor close to having eff dim one for $H = 120$.

5. Handling Discontinuity in Pricing and Hedging

Example 5.4. We consider a binary (or digital) Asian option:

$$g(\mathbf{S}) = \mathbf{I}_{\{S_A < K\}}(\mathbf{S}).$$

Theorem 4 cannot be applied directly.

We seek a good PGM for an "auxiliary" indicator function whereby OT can be implemented easily. The PGM, which is QMC-friendly to the auxiliary problem, is applied to the original problem.

A natural auxiliary indicator function is $\mathbf{I}_{\{S_G < K\}}(\mathbf{S})$, for which Theorem 4 can be applied successfully.

(The following results are for weights (B), see the original paper)

5. Handling Discontinuity in Pricing and Hedging

Table B7: The VRFs for Example 5.4

d	weights	K	MC STD	Sobol' STD	Sobol' BB	Sobol' PCA	Sobol' OT	Korobov STD	Korobov BB	Korobov PCA	Korobov OT
128	(B)	90	1	2	58	8	**1179**	3	39	6	**2525**
		100	1	3	50	12	**2332**	3	36	8	**2750**
		110	1	4	40	11	**2407**	2	47	7	**2818**

Table B8: The eff dim-related characteristics for Example 5.4

d	weights	K	Degree of Additivity R_1 STD	BB	PCA	OT	Mean Dimension d_{ms} STD	BB	PCA	OT
128	(B)	90	0.41	0.95	0.66	**0.9997**	11.28	1.13	2.42	**1.0046**
		100	0.67	0.95	0.70	**0.9985**	10.32	1.12	2.28	**1.0038**
		110	0.68	0.96	0.73	**0.9994**	10.21	1.12	2.24	**1.0036**

5. Handling Discontinuity in Pricing and Hedging

- OT has the best performance, leading to large VRF. The other PGMs in QMC behave only slightly better than MC.
- OT leads to large degree of additivity and small mean dim, and strike price has very small impact. On the contrary, STD leads to small degree of additivity and large mean dim, and strike price has significant impact. BB and PCA are between them.

5. Handling Discontinuity in Pricing and Hedging

Example 5.5 (Greeks by pathwise method) Let us estimate the delta of an arithmetic Asian option: $\max(S_A - K, 0)$. The pathwise estimate for the delta is

$$g(\mathbf{S}) = e^{-rT} \mathbf{I}_{\{S_A > K\}}(\mathbf{S}) \, \frac{S_A}{S_0}. \tag{17}$$

Similar indicator functions appear in pathwise estimates of other Greeks, such as *vega, rho, theta*.

Note that the payoff of the arithmetic Asian option is continuous, but using pathwise approach to estimate Greeks induces discontinuity.

The discontinuity structure of (17) cannot be expressed in a form (15). We use the same strategy as Example 5.4 and the sam weights (B), see the original paper.

5. Handling Discontinuity in Pricing and Hedging

Table B9: The VRFs for Example 5.5

d	weights	K	MC	Sobol'				Korobov			
			STD	STD	BB	PCA	OT	STD	BB	PCA	OT
128	(B)	90	1	5	102	17	**927**	5	74	12	**1746**
		100	1	5	75	19	**1806**	5	56	13	**1907**
		110	1	5	55	15	**1924**	3	66	10	**2362**

Table B10: The eff dim-related characteristics for Example 5.5

d	weights	K	Degree of Additivity R_1				Mean Dimension d_{ms}			
			STD	BB	PCA	OT	STD	BB	PCA	OT
128	(B)	90	0.76	0.98	0.84	**0.9996**	6.21	1.07	1.71	**1.0025**
		100	0.80	0.97	0.82	**0.9999**	6.84	1.08	1.80	**1.0024**
		110	0.73	0.97	0.80	**0.9994**	7.50	1.08	1.88	**1.0025**

5. Handling Discontinuity in Pricing and Hedging

- Again, OT outperforms other PGMs consistently in QMC, while other PGMs in QMC behave only slightly better than MC.
- OT leads to large degree of additivity and small mean dim, implying the obtained functions are highly additive.
- In some cases we cannot make the discontinuity only occurring exactly on axis-parallel hyperplanes, however, by using Theorem 4 we can meticulously “localize” the discontinuity.
- This presents a general principle of handling the discontinuity ...

6. Conclusions

- Discontinuity plays a crucial role in QMC.
- No PGM offer a consistent advantage.
- PGMs have significant impact on discontinuity structure. This explains why BB or PCA does not offer a consistent advantage.
- For discontinuous functions, using traditional PGMs in QMC should be very careful, since ...
- By carefully designing a PGM, we may ensure the discontinuity to be QMC-friendly, and generally have small eff dim, both are crucial for the success of QMC.
- The new method outperforms other PGMs remarkably in QMC.
- The new method and idea can be generalized easily ...

Thank you!