Smoothing the Payoff for Efficient Computation of Option Pricing in Time-Stepping Setting

1 Continuous time, theoretical motivation

The purpose of this project is to approximate $\mathrm{E}\left[g(\mathbf{X}_T)\right]$, where g is a certain payoff function and $\mathbf{X} = (X_1, \ldots, X_d)$ is described by the following SDE

$$dX_i = a_i(X)dt + \sum_{j=1}^{d} b_{ij}(X)dW_j. \tag{1.1}$$

Without loss of generality, we assume that the $\{W_j\}_{j=1}^d$ are uncorrelated (the correlation terms can be included the diffusion terms b_j).

First, we start by representing hierarchically $\mathbf{W}$. In fact, we can write

$$\begin{aligned} W_j(t) &= \frac{t}{T}W_j(T) + B_j(t) \\ &= \frac{t}{\sqrt{T}}Y_j + B_j(t), \end{aligned} \tag{1.2}$$

with $Y_j \sim \mathcal{N}(0,1)$ iid and $\{B_j\}_{j=1}^d$ are the Brownian bridges.

Now, we aim to represent $\mathbf{Y} = (Y_1, \ldots, Y_d)$ hierarchically by using a discrete Brownian bridge (Bb) namely

$$\mathbf{Y} = \underbrace{\mathbb{P}_0\mathbf{Y}}_{\text{One dimensional projection}} + \underbrace{\mathbb{P}_\perp\mathbf{Y}}_{\text{Projection on the complementary}}, \tag{1.3}$$

where we write $\mathbb{P}_0\mathbf{Y} = (\mathbf{Y}, \mathbf{v})\mathbf{v}$, with $(.,.)$ denotes the scalar product and $||\ \mathbf{v}\ || = 1$. We can easily show that $Z_v := (Y, v)$ is normal with $\mathrm{E}\left[Z_v\right] = 0$ and $\mathrm{Var}(Z_v) = 1$. Furthermore, we can write

$$\begin{aligned} Y_j &= Z_v v_j + (\mathbb{P}_\perp\mathbf{Y})_j \\ &= Z_v v_j + (Z_v^\perp)_j. \end{aligned} \tag{1.4}$$

The first aim of the project is to determine the optimal direction $\mathbf{v}$. By optimal direction, we mean the direction that maximizes the smoothing effect, that is the one given by

$$\sup_{\substack{\mathbf{v}\in\mathbb{R}^{d\times 1}\\ ||\mathbf{v}||=1}} \left(\max_{1\le i\le d} \left| \frac{\partial^2 g}{\partial^2 v_i} \right| \right). \tag{1.5}$$

We need to check if it is better to consider the second derivative as a metric for optimality or not. Also, one may check if formulation (1.5) is correct. For instance, one may consider also the following metric

$$\sup_{\substack{\mathbf{v}\in\mathbb{R}^{d\times 1}\\ ||\mathbf{v}||=1}} \left(\mathrm{Var}\left[g(\widehat{X}_T) \right] \right), \tag{1.6}$$

where $\widehat{X}$ is defined in (1.9).

Going back to the SDE (1.1), we have

$$dX_i = a_i(X)dt + \sum_{j=1}^{d} b_{ij}(X)Y_j \frac{dt}{\sqrt{T}} + \sum_{j=1}^{d} b_{ij}(X)Y_j \frac{dt}{\sqrt{T}} + \sum_{j=1}^{d} b_{ij}(X)dB_j. \tag{1.7}$$

Using (1.4), then we have

$$dX_i = \left(a_i(X) + \sum_{j=1}^{d} b_{ij}(X)\frac{Z_v v_j}{\sqrt{T}} \right) dt + \left(\sum_{j=1}^{d} b_{ij}(X)\frac{(Z_v^{\perp})_j}{\sqrt{T}} \right) dt + \sum_{j=1}^{d} b_{ij}(X)dB_j. \tag{1.8}$$

1.1 First approach (Brutal)

We assume that the first term in the right-hand side of (1.8) is the dominant term compared to the remaining terms (We need to formulate clearly what we mean by this statement. Also I see some issues with this assumption, since it implicitly implies from (1.4) that $Y \approx Z_v v$, which is already a strong assumption about the solution v.), and denote the approximate process $\widehat{X}$, whose dynamics are given by

$$\frac{d\widehat{X}_i}{dt} = a_i(\widehat{X}) + \sum_{j=1}^{d} b_{ij}(\widehat{X})\frac{Z_v v_j}{\sqrt{T}}, \quad 1 \le i \le d, \tag{1.9}$$

with mean

$$\frac{d\mathrm{E}\left[\widehat{X}_i\right]}{dt} = \mathrm{E}\left[a_i(\widehat{X})\right] + \sum_{j=1}^{d} \mathrm{E}\left[b_{ij}(\widehat{X})\frac{Z_v v_j}{\sqrt{T}} \right], \quad 1 \le i \le d,$$

and second moment

$$\begin{aligned} \frac{d\mathrm{E}\left[\widehat{X}_i^2\right]}{dt} &= 2\mathrm{E}\left[\widehat{X}_i \frac{d\widehat{X}_i}{dt} \right] \\ &= 2\mathrm{E}\left[\widehat{X}_i a_i(\widehat{X})\right] + \sum_{j=1}^{d} \mathrm{E}\left[\widehat{X}_i b_{ij}(\widehat{X})\frac{Z_v v_j}{\sqrt{T}} \right], \quad 1 \le i \le d. \end{aligned}$$

Now, let us expand $\widehat{X}$ around Z_v, that is

$$\widehat{X}_{(Z_v)} = \widehat{X}_{(0)} + \widehat{X}'_{(0)} Z_v + \dots \tag{1.10}$$

Then, we have

$$\begin{aligned}
\frac{d\mathrm{E}\left[\widehat{X}_i\right]}{dt} &\approx \mathrm{E}\left[a_i(\widehat{X}_{(0)})\right] + \sum_{j=1}^{d} \mathrm{E}\left[b_{ij}(\widehat{X}_{(0)}) \frac{Z_v v_j}{\sqrt{T}}\right], \quad 1 \le i \le d, \\
&= a_i(\widehat{X}_{(0)}) + \sum_{j=1}^{d} b_{ij}(\widehat{X}_{(0)}) \frac{\mathrm{E}\left[Z_v\right] v_j}{\sqrt{T}}, \quad 1 \le i \le d, \\
&= a_i(\widehat{X}_{(0)}), \quad 1 \le i \le d. \qquad (1.11)
\end{aligned}$$

Similarly, if we approximate the second moment, we get

$$\begin{aligned}
\frac{d\mathrm{E}\left[\widehat{X}_i^2\right]}{dt} &\approx 2\mathrm{E}\left[\widehat{X}_{i,(0)} a_i(\widehat{X}_{(0)})\right] + \sum_{j=1}^{d} \mathrm{E}\left[\widehat{X}_{i,(0)} b_{ij}(\widehat{X}_{(0)}) \frac{Z_v v_j}{\sqrt{T}}\right], \quad 1 \le i \le d, \\
&\approx 2\widehat{X}_{i,(0)} a_i(\widehat{X}_{(0)}) + \sum_{j=1}^{d} \widehat{X}_{i,(0)} b_{ij}(\widehat{X}_{(0)}) \frac{\mathrm{E}\left[Z_v\right] v_j}{\sqrt{T}}, \quad 1 \le i \le d, \\
&\approx 2\widehat{X}_{i,(0)} a_i(\widehat{X}_{(0)}), \quad 1 \le i \le d. \qquad (1.12)
\end{aligned}$$

The idea then is maximize, at the final time, the smoothing effect, represented by $\mathrm{Var}\left[g(\widehat{X}_T)\right]$, where $\widehat{X}_T \sim \mathcal{N}(\mu_T, \Sigma_T)$, such that μ_T, Σ_T are computed using (1.11) and (1.12).

There is an issue that I see here! In fact, if we just use the zero order expansion of $\widehat{X}$ to get (1.11) and (1.12), then we end up with ODEs that does not depend on v. On the other hand, to get dependence with respect to v we need to push the expansion of $\widehat{X}$ at least up to the first order which will involve $\widehat{X}'$, and which can make the approach complicated for instance if one consider Heston model for example.

2 Discrete time, practical motivation

To motivate our purposes, we consider the basket option under multi-dimensional GBM model where the process $\mathbf{X}$ is the discretized d-dimensional Black-Scholes model and the payoff function g is given by

$$g(\mathbf{X}(T)) = \max\left(\sum_{j=1}^{d} c_j X^{(j)}(T) - K, 0\right). \tag{2.1}$$

Precisely, we are interested in the d-dimensional lognormal example where the dynamics of the stock are given by

$$dX_t^{(j)} = \sigma^{(j)} X_t^{(j)} dW_t^{(j)}, \tag{2.2}$$

where $\{W^{(1)},\dots,W^{(d)}\}$ are correlated Brownian motions with correlations ρ_{ij}.

We denote by $(Z_1^{(j)},\dots,Z_N^{(j)})$ the N Gaussian independent rdvs that will be used to construct the path of the j-th asset $\overline{X}^{(j)}$, where $1\le j\le d$ (d denotes the number of underlyings considered in the basket). We denote $\psi^{(j)}:(Z_1^{(j)},\dots,Z_N^{(j)})\to(B_1^{(j)},\dots,B_N^{(j)})$ the mapping of Brownian bridge construction and by $\Phi:(\Delta t,\widetilde{B}_1^{(1)},\dots,\widetilde{B}_N^{(1)},\dots,\widetilde{B}_1^{(d)},\dots,\widetilde{B}_N^{(d)})\to\left(\overline{X}_T^{(1)},\dots,\overline{X}_T^{(d)}\right)$, the mapping consisting of the time-stepping scheme, where $\widetilde{\mathbf{B}}$ is the correlated Brownian bridge that can be obtained from the non correlated Brownian bridge $\mathbf{B}$ through multiplication by the correlation matrix, we denote this transformation by $T:\left(B_1^{(1)},\dots,B_N^{(1)},\dots,B_1^{(d)},\dots,B_N^{(d)}\right)\to\left(\widetilde{B}_1^{(1)},\dots,\widetilde{B}_N^{(1)},\dots,\widetilde{B}_1^{(d)},\dots,\widetilde{B}_N^{(d)}\right)$. Then, we can express the option price as

$$\begin{aligned}
\mathrm{E}\left[g(\mathbf{X}(T))\right] &\approx \mathrm{E}\left[g\left(\overline{X}_T^{(1)},\dots,\overline{X}_T^{(d)}\right)\right]\\
&= \mathrm{E}\left[g\left(\Phi\circ T\right)\left(B_1^{(1)},\dots,B_N^{(1)},\dots,B_1^{(d)},\dots,B_N^{(d)}\right)\right]\\
&= \mathrm{E}\left[g\left(\Phi\circ T\right)\left(\psi^{(1)}(Z_1^{(1)},\dots,Z_N^{(1)}),\dots,\psi^{(d)}(Z_1^{(d)},\dots,Z_N^{(d)})\right)\right]\\
&= \int_{\mathbb{R}^{d\times N}} G(z_1^{(1)},\dots,z_N^{(1)},\dots,z_1^{(d)},\dots,z_N^{(d)}))\rho_{d\times N}(\mathbf{z})dz_1^{(1)}\dots dz_N^{(1)}\dots z_1^{(d)}\dots dz_N^{(d)}, \qquad (2.3)
\end{aligned}$$

where

$$\rho_{d\times N}(\mathbf{z})=\frac{1}{(2\pi)^{d\times N/2}}e^{-\frac{1}{2}\mathbf{z}^T\mathbf{z}}.$$

In the discrete case, we can show that the numerical approximation of $X^{(j)}(T)$ satisfies

$$\begin{aligned}
\overline{X}^{(j)}(T) &= X_0^{(j)}\prod_{i=0}^{N-1}\left[1+\frac{\sigma^{(j)}}{\sqrt{T}}Z_1^{(j)}\Delta t+\sigma^{(j)}\Delta\widetilde{B}_i^{(j)}\right],\quad 1\le j\le d\\
&= \prod_{i=0}^{N-1} f_i^{(j)}(Z_1^{(j)}),\quad 1\le j\le d. \qquad (2.4)
\end{aligned}$$

2.1 Step 1: Numerical smoothing

The first step of our idea is to smoothen the problem by solving the root finding problem in one dimension after using a sub-optimal linear mapping for the coarsest factors of the Brownian increments $\mathbf{Z}_1=(Z_1^{(1)},\dots,Z_1^{(d)})$. In fact, let us define for a certain $d\times d$ matrix $\mathcal{A}$, the linear mapping

$$\mathbf{Y}=\mathcal{A}\mathbf{Z}_1. \qquad (2.5)$$

Then from (2.4), we have

$$\begin{aligned}
\overline{X}^{(j)}(T) &= \prod_{i=0}^{N-1} f_i^{(j)}(\mathcal{A}^{-1}\mathbf{Y})_j,\quad 1\le j\le d,\\
&= \prod_{i=0}^{N-1} g_i^{(j)}(Y_1,\mathbf{Y}_{-1})\quad 1\le j\le d \qquad (2.6)
\end{aligned}$$

where, with defining $\mathcal{A}^{\text{inv}} = \mathcal{A}^{-1}$, we have

$$
\begin{aligned}
g_i^{(j)}(Y_1, \mathbf{Y}_{-1}) &= X_0^{(j)} \left[1 + \frac{\sigma^{(j)}}{\sqrt{T}} \left(\sum_{i=1}^{d} A_{ji}^{\text{inv}} Y_i \right) \Delta t + \sigma^{(j)} \Delta \widetilde{B}_i^{(j)} \right] \\
&= X_0^{(j)} \left[1 + \frac{\sigma^{(j)} \Delta t}{\sqrt{T}} A_{j1}^{\text{inv}} Y_1 - \frac{\sigma^{(j)}}{\sqrt{T}} \left(\sum_{i=2}^{d} A_{ji}^{\text{inv}} Y_i \right) \Delta t + \sigma^{(j)} \Delta \widetilde{B}_i^{(j)} \right] \qquad (2.7)
\end{aligned}
$$

Therefore, in order to determine Y_1^*, we need to solve

$$
x = \sum_{j=1}^{d} c_j \prod_{i=0}^{N-1} g_i^{(j)}(Y_1^*(x), \mathbf{Y}_{-1}), \qquad (2.8)
$$

which implies that the location of the kink point for the approximate problem is equivalent to finding the roots of the polynomial $P(Y_1^*(K))$, given by

$$
P(Y_1^*(K)) = \left(\sum_{j=1}^{d} c_j \prod_{i=0}^{N-1} g_i^{(j)}(Y_1^*) \right) - K. \qquad (2.9)
$$

Using **Newton iteration method**, we use the expression $P' = \frac{dP}{dY_1^*}$, and we can easily show that

$$
P'(W_1) = \sum_{j=1}^{d} c_j \frac{\sigma^{(j)} \Delta t A_{j1}^{\text{inv}}}{\sqrt{T}} \left(\prod_{i=0}^{N-1} g_i^{(j)}(Y_1) \right) \left[\sum_{i=0}^{N-1} \frac{1}{g_i^{(j)}(Y_1)} \right]. \qquad (2.10)
$$

Remark 2.1. For our purposes, we suggest already that the coarsest factors of the Brownian increments are the most important ones, compared to the remaining factors. However, one may expect that in case we want to optimize over the choice of the linear mapping $\mathcal{A}$, and which direction is the most important for the kink location, one needs then to solve

$$
\sup_{\substack{\mathcal{A} \in \mathbb{R}^{d \times d} \\ \mathcal{A} \text{ is a rotation}}} \left(\max_{1 \le i \le d} \left| \frac{\partial g}{\partial Y_i} \right| \right),
$$

which becomes hard to solve when d increases.

Remark 2.2. A general choice for $\mathcal{A}$ should in the family of rotations. However, we think that a sufficiently good matrix $\mathcal{A}$ would be the one leading to $Y_1 = \sum_{i=1}^{d} Z_1^{(i)}$ up to re-scaling.

2.2 Step 2: Integration

At this stage, we want to perform the pre-integrating step with respect to W_1^*. In fact, we have from (5.5)

$$\begin{aligned}
\mathrm{E}\left[g(\mathbf{X}(T))\right] &= \int_{\mathbb{R}^{d\times N}} G(z_1^{(1)},\dots,z_N^{(1)},\dots,z_1^{(d)},\dots,z_N^{(d)}))\rho_{d\times N}(\mathbf{z})dz_1^{(1)}\dots dz_N^{(1)}\dots z_1^{(d)}\dots dz_N^{(d)} \\
&= \int_{\mathbb{R}^{dN-1}} \left(\int_{\mathbb{R}} G(y_1,\mathbf{y}_{-1},\mathbf{z}_{-1}^{(1)},\dots,\mathbf{z}_{-1}^{(d)})\rho_{y_1}(y_1)dy_1\right)\rho_{d-1}(\mathbf{y}_{-1})d\mathbf{y}_{-1}\rho_{d\times(N-1)}(\mathbf{z}_{-1}^{(1)},\dots,\mathbf{z}_{-1}^{(d)})d\mathbf{z}_{-1}^{(1)}\dots d\mathbf{z}_{-1}^{(d)} \\
&= \int_{\mathbb{R}^{dN-1}} h(\mathbf{y}_{-1},\mathbf{z}_{-1}^{(1)},\dots,\mathbf{z}_{-1}^{(d)})\rho_{d-1}(\mathbf{y}_{-1})d\mathbf{y}_{-1}\rho_{d\times(N-1)}(\mathbf{z}_{-1}^{(1)},\dots,\mathbf{z}_{-1}^{(d)})d\mathbf{z}_{-1}^{(1)}\dots d\mathbf{z}_{-1}^{(d)}, \\
&= \mathrm{E}\left[h(\mathbf{y}_{-1},\mathbf{z}_{-1}^{(1)},\dots,\mathbf{z}_{-1}^{(d)})\right],
\end{aligned} \tag{2.11}$$

where

$$\begin{aligned}
h(\mathbf{y}_{-1},\mathbf{z}_{-1}^{(1)},\dots,\mathbf{z}_{-1}^{(d)}) &= \int_{\mathbb{R}} G(y_1,\mathbf{y}_{-1},\mathbf{z}_{-1}^{(1)},\dots,\mathbf{z}_{-1}^{(d)})\rho_{y_1}(y_1)dy_1 \\
&= \int_{-\infty}^{y_1^*} G(y_1,\mathbf{y}_{-1},\mathbf{z}_{-1}^{(1)},\dots,\mathbf{z}_{-1}^{(d)})\rho_{y_1}(y_1)dy_1 \\
&+ \int_{y_1^*}^{+\infty} G(y_1,\mathbf{y}_{-1},\mathbf{z}_{-1}^{(1)},\dots,\mathbf{z}_{-1}^{(d)})\rho_{y_1}(y_1)dW_1
\end{aligned} \tag{2.12}$$

2.3 The best call option case under GBM and Heston model

The second example that we consider for multi-dimension is the best call option under GBM and Heston model. I am in process of formulating this but somehow we agreed that the first potential directions of smoothing for the Heston model will be the first factors related to the asset prices.

3 Introduction

3.1 The goal and outline of the project

The first goal of the project is to approximate $\mathrm{E}\left[f(X(t))\right]$, using multi-index stochastic collocation(MISC) method, proposed in [11], where

- The payoff $f:\mathbb{R}^d\to\mathbb{R}$ has either jumps or kinks. Possible choices of f that we wanted to test are:
  - hockey-stick function, i.e., put or call payoff functions;
  - indicator functions (both relevant in finance (binary option,...) and in other applications of estimation of probabilities of certain events);
  - delta-functions for density estimation (and derivatives thereof for estimation of derivatives of the density).

  More specifically, f should be the composition of one of the above with a smooth function. (For instance, the basket option payoff as a function of the log-prices of the underlying.)

- The process X is simulated via a time-stepping scheme. Possible choices that we wanted to test are
  - The one/multi dimensional discretized Black-Scholes(BS) process where we compare different ways to identify the location of the kink, such as:
    * Exact location of the continuous problem
    * Exact location of the discrete problem by root finding of a polynomial in y.
    * Newton iteration.
  - A relative simple interest rate model or stochastic volatility model, for instance CIR or Heston models: In fact, the impact of the Brownian bridge will disappear in the limit, which may make the effect of the smoothing, but also of the errors in the kink location difficult to identify. For this reason, we suggest to study a more complicated 1-dimensional problem next. We suggest to use a CIR process. To avoid complications at the boundary, we suggest "nice" parameter choices, such that the discretized process is very unlikely to hit the boundary (Feller condition).
  - The multi dimensional discretized Black-Scholes(BS) process: Here, we suggest to return to the Black-Scholes model, but in multi-dimensional case. In this case, linearizing the exponential, suggest that a good variable to use for smoothing might be the sum of the final values of the Brownian motion. In general, though, one should probably eventually identify the optimal direction(s) for smoothing via the duals algorithmic differentiation.

The desired outcome is a paper including

- Theoretical results including: i) an analiticity proof for the integrand in the time stepping setting, ii) a numerical analysis of the schemes involved, such as Newton iteration, etc.
- Applications that tests the examples above.

What has beed achieved so far:

1. Numerical outputs:
   - **Example 1**: Tests for the basket option with the smoothing trick as in [2]: in that example we checked the performance of MISC without time stepping scheme and also compare the results with reference [2] (See Section 7.7). (Done).
   - **Example 2**: The one dimensional binary option under discretized BS model (see Section 7.4). The results are promising (See Section 7.4) (Done).
   - **Example 3**: The one dimensional call option under discretized BS model (see Section 7.5). The results are promising (See Section 7.5) (Done).
   - **Example 4**: The multi dimensional basket call option under discretized BS model (see Section 7.2). (Under process).
2. Theoretical outputs:
   - Heuristic proof of analiticity.

3.2 Literature review

Many option pricing problems require the computation of multivariate integrals. The dimension of these integrals is determined by the number of independent stochastic factors (e.g. the number of time steps in the time discretization or the number of assets under consideration). The high dimension of these integrals can be treated with dimension-adaptive quadrature methods to have the desired convergence behavior.

Unfortunately, in many cases, the integrand contains either kinks and jumps. In fact, an option is normally considered worthless if the value falls below a predetermined strike price. A kink (discontinuity in the gradients) is present when the payoff function is continuous, while a jump (discontinuity in the function) exists when the payoff coreesponds to a binary or other digital options. The existence of kinks or jumps in the integrand heavily degrades the performance of quadrature formulas. In this work, we are interested in solving this problem by using adaptive sparse grids (SG) methods coupled with suitable transformations. The main idea is to find lines or areas of discontinuity and to employ suitable transformations of the integration domain. Then by a pre-integration (smoothing) step with respect to the dimension containing the kink/jump, we end up with integrating only over the smooth parts of the integrand and the fast convergence of the sparse grid method can be regained.

One can ignore the kinks and jumps, and apply directly a method for integration over $\mathbb{R}^d$. Despite the significant progress in SG methods [3] for high dimensional integration of smooth integrands, few works have been done to deal with cases involving integrands with kinks or jumps due to the decreasing performance of SG methods in the presence of kinks and jumps.

Some works [8, 2, 9, 10, 16] adressed similar kind of problems, characterized by the presence of kinks and jumps, but with much more emphasis on Quasi Monte Carlo (QMC). In [8, 9, 10], an analysis of the performance of Quasi Monte Carlo (QMC) and SG methods has been conducted, in the presence of kinks and jumps. In [8, 9], the authors studied the terms of the ANOVA decomposition of functions with kinks defined on d-dimensional Euclidean space $\mathbb{R}^d$, and showed that under some assumptions all but the the highest order ANOVA term of the 2^d ANOVA terms can be smooth for the case of an arithmetic Asian option with the Brownian bridge construction. Furthermore, [10] extended the work in [8, 9] from kinks to jumps for the case of an arithmetic average digital Asian option with the principal component analysis (PCA). The main findings in [8, 9] was obtained for an integrand of the form $f(\mathbf{x}) = \max(\phi(\mathbf{x}), 0)$ with ϕ being smooth. In fact, by assuming i) the d-dimensional function ϕ has a positive partial derivative with respect to x_j for some $j \in \{1, \ldots, d\}$, ii) certain growth conditions at infinity are satisfied, the authors showed that the ANOVA terms of f that do not depend on the variable x_j are smooth. We note that [8, 9, 10] focus more on theoretical aspects of applying QMC in such a setting. On the other hand, we focus more on specific practical problems, where we add the adaptivity paradigm to the picture.

A recent work [16] adresses similar kind of problems using QMC. Being very much related to [2], the authors i) assume that the conditional expectation can be computed explicitly, by imposing very strong assumptions. ii) Secondly, they use PCA on the gradients to reduce the effective dimension. In our work, we do not make such strong assumptions, which is why we need numerical methods, more precisely root finding and the quadrature in the first direction.

4 Problem formulation and Setting

In the context of option pricing, we aim at approximating the option price, $E[g(\mathbf{X}(t))]$, where $g : \mathbb{R}^d \to \mathbb{R}$ is the payoff function and where the process of the asset prices $\mathbf{X} \in \mathbb{R}^d$ solves

$$\mathbf{X}(t) = \mathbf{X}(0) + \int_0^t a(s, \mathbf{X}(s))ds + \sum_{\ell=1}^{\ell_0} \int_0^t b^\ell(s, \mathbf{X}(s))dW^\ell(s)$$

Let us denote by $\Phi : (\mathbf{z}_1, \ldots, \mathbf{z}_N) \to \overline{\mathbf{X}}_T$, the mapping consisting of the time-stepping scheme, where $\{\mathbf{z}_i\}_{i=1}^N$ are independent d-dimensional Gaussian random vectors, and N is the number of time steps. Without loss of Generality, we assume that Φ may include pre-processing transformations to reduce the effective dimension. We also assume that $d = 1$ and the extension to higher dimension is trivial.

In this setting, we are interetsed in the basic problem of approximating

$$\begin{aligned} E[g(\mathbf{X}(T))] &\approx E[g(\overline{\mathbf{X}}_T)] \\ &= \int_{\mathbb{R}^N} g \circ \Phi(\mathbf{z})d\mathbf{z} = \int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} g \circ \Phi(z_1, \ldots, z_N)\rho_d(\mathbf{z})dz_1, \ldots, dz_N \\ &= I_N(g \circ \Phi), \end{aligned} \tag{4.1}$$

with

$$\rho_N(\mathbf{z}) = \frac{1}{(2\pi)^{N/2}} e^{-\frac{1}{2}\mathbf{z}^T\mathbf{z}}.$$

where ρ is a continuous and strictly positive probability density function on $\mathbb{R}$ and g is a real-valued function integrable with respect to ρ_N.

In this context, we work mainly with two possible structures of payoff function g. In fact, for the cases of call/put options, the payoff g has a kink and will be of the form

$$g(\mathbf{x}) = \max(\phi(\mathbf{x}), 0).$$

One can also encounter jumps in the payoff when working with binary digital options. In this case, g is given by

$$g(\mathbf{x}) = \mathbf{1}_{(\phi(\mathbf{x}) \geq 0)}.$$

We introduce the notation $\mathbf{x} = (x_j, \mathbf{x}_{-j})$, where $\mathbf{x}_{-j}$ denotes the vector of length $d-1$ denoting all the variables other than x_j. Then, if we assume for some $j \in \{1, \ldots, d\}$

$$\frac{\partial \phi}{\partial x_j}(\mathbf{x}) > 0, \ \forall \mathbf{x} \in \mathbb{R}^d \ \textbf{(Monotonicity condition)} \tag{4.2}$$

$$\lim_{x \to +\infty} \phi(\mathbf{x}) = \lim_{x \to +\infty} \phi(x_j, \mathbf{x}_{-j}) = +\infty, \text{ or } \frac{\partial^2 \phi}{\partial x_j^2}(\mathbf{x}) \ \textbf{(Growth condition)}, \tag{4.3}$$

then, using Fubini's theorem, we can rewrite (4.1) as

$$\begin{aligned} I_N(g \circ \Phi) &= \int_{\mathbb{R}^{d-1}} \left(\int_{-\infty}^{\infty} g \circ \Phi(z_j, \mathbf{z}_{-j})\rho(z_j)dz_j \right) \rho_{z-1}(\mathbf{z}_{-j})d\mathbf{z}_{-j}, \\ &= \mathrm{E}\left[E\left[g \circ \Phi(z_j, \mathbf{z}_{-j}) \mid z_j\right]\right] \end{aligned} \tag{4.4}$$

where we evaluate the inner integral for each $\mathbf{z}_{-j}$ and which results in a smooth integrand for the outer $(N-1)$-dimensional integral.

We note that conditions ((4.2) and (4.3)) imply that for each $\mathbf{z}_{-j}$, the function $\phi \circ \Phi(z_j, \mathbf{z}_{-j})$ either has a simple root z_j or is positive for all $z_j \in \mathbb{R}$.

We generally do not have a closed form for the inside integral in (4.4). Therefore, the pre-integration (conditional sampling) step should be performed numerically.

5 Details of our approach

In the following, we describe our approach which basically can be seen as a two stage method. In the first step, we use root finding procedure to get the inner integral in (4.4), then in a second stage we employ adaptive quadraure, multi-index stochastic collocation (MISC), to compute the obtained smooth integrand.

5.1 The one dimensional case

For illustration purposes, let us focus on the one dimensional case, where under the risk-neutral measure, the undelying asset follows the geometric Brownian motion (GBM)

$$dX_t = rX_t dt + \sigma X_t dB_t, \tag{5.1}$$

where r is the risk-free rate, σ is the volatility and B_t is the standard Brownian motion. The analytical solution to (5.1) is

$$X_t = X_0 \exp((r - \sigma^2)t + \sigma B_t).$$

The Brownian motion B_t can be constructed either sequentially using a standard random walk construction or hierarchically using Brownian bridge (Bb) construction. To make an effective use of MISC, which profits from anisotropy, we use the Bb construction since it produces dimensions with different importance for MISC (creates anisotropy), contrary to random walk procedure for which all the dimension of the stochastic space have equal importance (isotropic). We explain the Bb construction in Section 5.4.1.

Let us denote by $\psi : (z_1, \ldots, z_N) \to (B_1, \ldots, B_N)$ the mapping of Bb construction and by $\Phi : (B_1, \ldots, B_N) \to \overline{X}_T$, the mapping consisting of the time-stepping scheme. Then, we can express the option price as

$$\begin{aligned}
\mathrm{E}\left[g(X(T))\right] &\approx E[g(\overline{\mathbf{X}}_T)] \\
&= \mathrm{E}\left[g\left(\Phi \circ \psi\right)(z_1, \ldots, z_N)\right] \\
&= \int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} G(z_1, \ldots, z_N) \rho_N(\mathbf{z}) dz_1, \ldots, dz_N \\
&= I_N(G), \qquad (5.2)
\end{aligned}$$

where $G = g \circ \Phi \circ \psi$ and

$$\rho_N(\mathbf{z}) = \frac{1}{(2\pi)^{N/2}} e^{-\frac{1}{2}\mathbf{z}^T \mathbf{z}}.$$

Now, we can easily apply the procedure of pre-integration of section 4, where we can assume that the payoff function g can be either the maximum or indicator function and $\phi = \Phi \circ \psi$. The remaining ingredient is to determine with respect to which variable z_j we will integrate.

Claiming that pre-integrating with respect to z_1 is the optimal option then from (5.2), we have

$$\begin{aligned} E[g(\overline{\mathbf{X}}_T)] &= \int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} G(z_1, \ldots, z_N) \rho_N(\mathbf{z}) dz_1, \ldots, dz_N \\ &= \int_{\mathbb{R}^{N-1}} \left(\int_{-\infty}^{\infty} G(z_1, \mathbf{z}_{-1}) \rho(z_1) dz_1 \right) \rho_{N-1}(\mathbf{z}_{-1}) d\mathbf{z}_{-1} \\ &= \int_{\mathbb{R}^{N-1}} h(\mathbf{z}_{-1}) \rho_{N-1}(\mathbf{z}_{-1}) d\mathbf{z}_{-1}, \\ &= \mathrm{E}\left[h(\mathbf{z}_{-1})\right] \end{aligned} \tag{5.3}$$

where $h(\mathbf{z}_{-1}) = \int_{-\infty}^{\infty} G(z_1, \mathbf{z}_{-1}) \rho(z_1) dz_1 = E\left[G\left(z_1, \ldots, z_N\right) \mid z_1\right]$.

Since g can have a kink or jump. Computing $h(\mathbf{z}_{-1})$ in the pre-integration step should be carried carefully to not deteriorate the smoothness of h. This can be done by applying a root finding procedure and then computing the uni-variate integral by summing the terms coming from integrating in each region where g is smooth. In Sections (5.6,5.7), we explain those points.

Once we perform stage 1 procedure, we use multi-index stochastic collocation (MISC) procedure, suggested in [11], to compute the expectation $\mathrm{E}\left[h(\mathbf{z}_{-1})\right]$. We describe the general strategy for the multi-index construction in Section 5.3.

If we denote by $\mathcal{E}_{\text{tot}}$ the total error of approximating the expectation in (5.2) using the MISC estimator, Q_N, then we have a natural error decomposition

$$\begin{aligned} \mathcal{E}_{\text{tot}} &\leq |C_{\text{RB}} - I_N(G)| + |I_N(G) - Q_N| \\ &\leq \mathcal{E}_B(N) + \mathcal{E}_Q(\text{TOL}_{\text{MISC}}, N), \end{aligned} \tag{5.4}$$

where $\mathcal{E}_Q$ is the quadrature error, $\mathcal{E}_B$ is the bias, and $I_N(G)$ is the biased price computed with N time steps as given by (5.2).

5.2 Extension to the high dimensional case

In the high dimensional case, we denote by $(z_1^{(i)}, \ldots, z_N^{(i)})$ the N Gaussian independent rdvs that will be used to construct the path of the i-th asset $\overline{X}^{(i)}$, where $1 \leq i \leq d$ (d denotes the number of underlyings considered in the basket). We keep the same notations by denoting $\psi : (z_1^{(i)}, \ldots, z_N^{(i)}) \to (B_1, \ldots, B_N)$ the mapping of Bb construction and by $\Phi : (B_1^{(i)}, \ldots, B_N^{(i)}) \to \overline{X}_T^{(i)}$, the mapping consisting of the time-stepping scheme. Then, we can express the option price as

$$\begin{aligned} \mathrm{E}\left[g(\mathbf{X}(T))\right] &\approx \mathrm{E}\left[g\left(\Phi \circ \psi\right)\left(z_1^{(1)}, \ldots, z_N^{(1)}, \ldots, z_1^{(d)}, \ldots, z_N^{(d)}\right)\right] \\ &= \int_{\mathbb{R}^{d \times N}} G(z_1^{(1)}, \ldots, z_N^{(1)}, \ldots, z_1^{(d)}, \ldots, z_N^{(d)})) \rho_{d \times N}(\mathbf{z}) dz_1^{(1)} \ldots dz_N^{(1)} \ldots z_1^{(d)} \ldots dz_N^{(d)}, \end{aligned} \tag{5.5}$$

where $G = g \circ \Phi \circ \psi$ and

$$\rho_{d \times N}(\mathbf{z}) = \frac{1}{(2\pi)^{d \times N/2}} e^{-\frac{1}{2}\mathbf{z}^T \mathbf{z}}.$$

Now, performing the pre-integrating step with respect to the coarse rdvs $(z_1^{(1)}, \dots, z_1^{(d)})$, results in

$$\begin{aligned}
\mathrm{E}\left[g(\mathbf{X}(T))\right] &= \int_{\mathbb{R}^{d\times N}} G(z_1^{(1)}, \dots, z_N^{(1)}, \dots, z_1^{(d)}, \dots, z_N^{(d)}))\rho_{d\times N}(\mathbf{z}) dz_1^{(1)} \dots dz_N^{(1)} \dots z_1^{(d)} \dots dz_N^{(d)} \\
&= \int_{\mathbb{R}^{d\times(N-1)}} \left(\int_{\mathbb{R}^d} G(z_1^{(1)}, \mathbf{z}_{-1}^{(1)}, \dots, z_1^{(d)}, \mathbf{z}_{-1}^{(d)}) \rho_d(z_1^{(1)}, \dots, z_1^{(d)}) dz_1^{(1)} \dots dz_1^{(d)} \right) \rho_{d\times(N-1)}(\mathbf{z}_{-1}^{(1)}, \dots, \mathbf{z}_{-1}^{(d)}) d\mathbf{z}_{-1}^{(1)} \dots d\mathbf{z}_{-1}^{(d)} \\
&= \int_{\mathbb{R}^{d\times(N-1)}} h(\mathbf{z}_{-1}^{(1)}, \dots, \mathbf{z}_{-1}^{(d)}) \rho_{d\times(N-1)}(\mathbf{z}_{-1}^{(1)}, \dots, \mathbf{z}_{-1}^{(d)}) d\mathbf{z}_{-1}^{(1)} \dots d\mathbf{z}_{-1}^{(d)}, \qquad (5.6) \\
&= \mathrm{E}\left[h(\mathbf{z}_{-1}^{(1)}, \dots, \mathbf{z}_{-1}^{(d)}) \right]
\end{aligned}$$

where $h(\mathbf{z}_{-1}^{(1)}, \dots, \mathbf{z}_{-1}^{(d)}) = \int_{\mathbb{R}^d} G(z_1^{(1)}, \mathbf{z}_{-1}^{(1)}, \dots, z_1^{(d)}, \mathbf{z}_{-1}^{(d)}) \rho_d(z_1^{(1)}, \dots, z_1^{(d)}) dz_1^{(1)} \dots dz_1^{(d)}$.

5.3 The MISC method

We assume that we want to approximate the expected value $\mathrm{E}[f(Y)]$ of an analytic function $f\colon \Gamma \to \mathbb{R}$ using a tensorization of quadrature formulas over Γ.

To introduce simplified notations, we start with the one-dimensional case. Let us denote by β a non negative integer, referred to as a "stochastic discretization level", and by $m : \mathbb{N} \to \mathbb{N}$ a strictly increasing function with $m(0) = 0$ and $m(1) = 1$, that we call "level-to-nodes function". At level β, we consider a set of $m(\beta)$ distinct quadrature points in $\mathbb{R}$, $\mathcal{H}^{m(\beta)} = \{y_\beta^1, y_\beta^2, \dots, y_\beta^{m(\beta)}\} \subset \mathbb{R}$, and a set of quadrature weights, $\boldsymbol{\omega}^{m(\beta)} = \{\omega_\beta^1, \omega_\beta^2, \dots, \omega_\beta^{m(\beta)}\}$. We also let $C^0(\mathbb{R})$ be the set of real-valued continuous functions over $\mathbb{R}$. We then define the quadrature operator as

$$Q^{m(\beta)} : C^0(\mathbb{R}) \to \mathbb{R}, \quad Q^{m(\beta)}[f] = \sum_{j=1}^{m(\beta)} f(y_\beta^j)\omega_\beta^j.$$

In our case, we have in (5.6) a multi-variate integration problem with, $f := h$, $\mathbf{Y} = (\mathbf{z}_{-1}^{(1)}, \dots, \mathbf{z}_{-1}^{(d)})$, and $\Gamma = \mathbb{R}^{d(N-1)}$, in the previous notations. Therefore, we define for any multi-index $\boldsymbol{\beta} \in \mathbb{N}^{d(N-1)}$

$$Q^{m(\boldsymbol{\beta})} : C^0(\mathbb{R}^{d(N-1)}) \to \mathbb{R}, \quad Q^{m(\boldsymbol{\beta})} = \bigotimes_{n=1}^{d(N-1)} Q^{m(\beta_n)},$$

where the n-th quadrature operator is understood to act only on the n-th variable of f. Practically, we obtain the value of $Q^{m(\boldsymbol{\beta})}[f]$ by considering the tensor grid $\mathcal{T}^{m(\boldsymbol{\beta})} = \times_{n=1}^{d(N-1)} \mathcal{H}^{m(\beta_n)}$ with cardinality $\#\mathcal{T}^{m(\boldsymbol{\beta})} = \prod_{n=1}^{d(N-1)} m(\beta_n)$ and computing

$$Q^{m(\boldsymbol{\beta})}[f] = \sum_{j=1}^{\#\mathcal{T}^{m(\boldsymbol{\beta})}} f(\widehat{y}_j)\overline{\omega}_j,$$

where $\widehat{y}_j \in \mathcal{T}^{m(\boldsymbol{\beta})}$ and $\overline{\omega}_j$ are products of weights of the univariate quadrature rules.

Remark 5.1. We note that the quadrature points are chosen to optimize the convergence properties of the quadrature error. For instance, in our context, since we are dealing with Gaussian densities, using Gauss-Hermite quadrature points is the appropriate choice.

A direct approximation $\mathrm{E}\left[f[\mathbf{Y}]\right] \approx Q^{\boldsymbol{\beta}}[f]$ is not an appropriate option due to the well-known "curse of dimensionality". We use MISC, which is a hierarchical adaptive sparse grids quadrature strategy that uses stochastic discretizations and classic sparsification approach to obtain an effective approximation scheme for $\mathrm{E}\left[f\right]$.

For the sake of concreteness, in our setting, we are left with a $d(N-1)$-dimensional Gaussian random input, which is chosen independently, resulting in $d(N-1)$ numerical parameters for MISC, which we use as the basis of the multi-index construction. For a multi-index $\boldsymbol{\beta} = (\beta_n)_{n=1}^{d(N-1)} \in \mathbb{N}^{d(N-1)}$, we denote by $Q_N^{\boldsymbol{\beta}}$, the result of approximating (5.6) with a number of quadrature points in the i-th dimension equal to $m(\beta_i)$. We further define the set of differences $\Delta Q_N^{\boldsymbol{\beta}}$ as follows: for a single index $1 \leq i \leq d(N-1)$, let

$$\Delta_i Q_N^{\boldsymbol{\beta}} = \begin{cases} Q_N^{\boldsymbol{\beta}} - Q_N^{\boldsymbol{\beta}'}, \text{ with } \boldsymbol{\beta}' = \boldsymbol{\beta} - e_i, \text{ if } \boldsymbol{\beta}_i > 0 \\ Q_N^{\boldsymbol{\beta}}, \quad \text{otherwise} \end{cases}$$

where e_i denotes the ith $d(N-1)$-dimensional unit vector. Then, $\Delta Q_N^{\boldsymbol{\beta}}$ is defined as

$$\Delta Q_N^{\boldsymbol{\beta}} = \left(\prod_{i=1}^{d(N-1)} \Delta_i \right) Q_N^{\boldsymbol{\beta}}.$$

Given the definition of $I^N(G)$ by (5.6), we have the telescoping property

$$I^N(G) = Q_N^{\infty} = \sum_{\beta_1=0}^{\infty} \cdots \sum_{\beta_{d(N-1)}=0}^{\infty} \Delta Q_N^{(\beta_1,\ldots,\beta_{d(N-1)})} = \sum_{\boldsymbol{\beta} \in \mathbb{N}^{d(N-1)}} \Delta Q_N^{\boldsymbol{\beta}}.$$

The MISC estimator used for approximating (5.6), and using a set of multi-indices $\mathcal{I} \subset \mathbb{N}^{d(N-1)}$ is given by

$$Q_N^{\mathcal{I}} = \sum_{\boldsymbol{\beta} \in \mathcal{I}} \Delta Q_N^{\boldsymbol{\beta}}. \tag{5.7}$$

The quadrature error in this case is given by

$$\mathcal{E}_Q(\mathrm{TOL}_{\mathrm{MISC}}, N) = \left| Q_N^{\infty} - Q_N^{\mathcal{I}} \right| \leq \sum_{\ell \in \mathbb{N}^{d(N-1)} \setminus \mathcal{I}} \left| \Delta Q_N^{\ell} \right|. \tag{5.8}$$

We define the work contribution, $\Delta \mathcal{W}_{\boldsymbol{\beta}}$, to be the computational cost required to add $\Delta Q_N^{\boldsymbol{\beta}}$ to $Q_N^{\mathcal{I}}$, and the error contribution, $\Delta E_{\boldsymbol{\beta}}$, to be a measure of how much the quadrature error, defined in (5.8), would decrease once $\Delta Q_N^{\boldsymbol{\beta}}$ has been added to $Q_N^{\mathcal{I}}$, that is

$$\begin{aligned} \Delta \mathcal{W}_{\boldsymbol{\beta}} &= \mathrm{Work}[Q_N^{\mathcal{I} \cup \{\boldsymbol{\beta}\}}] - \mathrm{Work}[Q_N^{\mathcal{I}}] \\ \Delta E_{\boldsymbol{\beta}} &= \left| Q_N^{\mathcal{I} \cup \{\boldsymbol{\beta}\}} - Q_N^{\mathcal{I}} \right|. \end{aligned} \tag{5.9}$$

The construction of the optimal $\mathcal{I}$ will be done by profit thresholding, that is, for a certain threshold value $\overline{T}$, and a profit of a hierarchical surplus defined by

$$P_{\boldsymbol{\beta}} = \frac{|\Delta E_{\boldsymbol{\beta}}|}{\Delta \mathcal{W}_{\boldsymbol{\beta}}},$$

the optimal index set $\mathcal{I}$ for MISC is given by $\mathcal{I} = \{\boldsymbol{\beta} : P_{\boldsymbol{\beta}} \geq \overline{T}\}$.

5.4 Path generation methods (PGM)

In the literature of adaptive sparse grids and QMC, several hierarchical path generation methods (PGMs) or transformation methods have been proposed to reduce the effective dimension. Among these transformations, we cite the Brownian bridge (Bb) construction [14, 4, 13, 12], the principal component analysis (PCA) [1] and the linear transformation (LT) [**?**], etc ...

Assume that one wants to compute $\mathrm{E}\left[g(B)\right]$, where B is a Brownian motion with index set $[0,T]$. In most applications this is can be reasonably approximated by $\mathrm{E}\left[\tilde{g}\left(B_{\frac{T}{N}},\ldots,B_{\frac{TN}{N}}\right)\right]$, where $\tilde{g}$ is a function of the set of discrete Brownian paths.

There are three classical methods for sampling from $\left(B_{\frac{T}{N}},\ldots,B_{\frac{TN}{N}}\right)$ given a standard normal vector Z, namely the forward method, the Brownian bridge (Bb) construction and the principal component analysis (PCA) construction. All of these constructions may be written in the form $\left(B_{\frac{T}{N}},\ldots,B_{\frac{TN}{N}}\right)=AZ$, where A is an $N\times N$ real matrix with

$$AA^T=\Sigma:=\left(\frac{T}{N}\min(j,k)\right)_{j,k=1}^{N}=\frac{T}{N}\begin{bmatrix}1&1&1&\ldots&1\\1&2&2&\ldots&2\\1&2&3&\ldots&3\\\vdots&\vdots&\vdots&\ddots&\vdots\\1&2&3&\ldots&N\end{bmatrix}. \tag{5.10}$$

For instance, the matrix A corresponding to the forward method is given by

$$A^F=\sqrt{\frac{T}{N}}\begin{bmatrix}1&0&\ldots&0\\1&1&\ldots&0\\\vdots&\vdots&\ddots&\vdots\\1&1&\ldots&1\end{bmatrix}.$$

In the case of Bb construction, details about the construction are given in Section 5.4.1, and the corresponding matrix A, For $N=8$ is given by

$$A^{\mathrm{Bb}}=\sqrt{T}\begin{bmatrix}\frac{1}{8}&\frac{1}{8}&\frac{\sqrt{2}}{8}&0&\frac{2}{8}&0&0&0\\\frac{2}{8}&\frac{2}{8}&\sqrt{2}\frac{2}{8}&0&0&0&0&0\\\frac{3}{8}&\frac{3}{8}&\frac{\sqrt{2}}{8}&0&0&\frac{2}{8}&0&0\\\frac{4}{8}&\frac{4}{8}&0&0&0&0&0&0\\\frac{5}{8}&\frac{3}{8}&0&\frac{\sqrt{2}}{8}&0&0&\frac{2}{8}&0\\\frac{6}{8}&\frac{2}{8}&0&\sqrt{2}\frac{2}{8}&0&0&0&0\\\frac{7}{8}&\frac{1}{8}&0&\frac{\sqrt{2}}{8}&0&0&0&\frac{2}{8}\\1&0&0&0&0&0&0&0\end{bmatrix}.$$

When doing PCA construction [1], we have $A^{\mathrm{PCA}}=VD$, where $\Sigma=VD^2V^T$ is the singular value decomposition of Σ.

5.4.1 Brownian bridge (Bb) construction

In our context, sampling the Brownian motion can be constructed either sequentially using a standard random walk construction or hierarchically using other hierarchical PGM as listed above.

For our purposes, to make an effective use of MISC, which profits from anisotropy, we use the Bb construction since it produces dimensions with different importance for MISC (creates anisotropy), contrary to random walk procedure for which all the dimension of the stochastic space have equal importance (isotropic). This pre-transformation reduces the effective dimension dimension of the problem and as a consquence accelerates the MISC procedure by reducing the computational cost.

Let us denote $\{t_i\}_{i=0}^N$ the grid of time steps, then the Bb construction [7] consists of the following: given a past value B_{t_i} and a future value B_{t_k}, the value B_{t_j} (with $t_i < t_j < t_k$) can be generated according to the formula:

$$B_{t_j} = (1-\rho)B_{t_i} + \rho B_{t_k} + \sqrt{\rho(1-\rho)(k-i)\Delta t}z,\ z \sim \mathcal{N}(0,1), \tag{5.11}$$

where $\rho = \frac{j-i}{k-i}$. In particular, if N is a power of 2, then given $B_0 = 0$, Bb generates the Brownian motion at times $T, T/2, T/4, 3T/4, \ldots$ according

$$\begin{aligned} B_T &= \sqrt{T}z_1 \\ B_{T/2} &= \frac{1}{2}(B_0 + B_T) + \sqrt{T/4}z_2 = \frac{\sqrt{T}}{2}z_1 + \frac{\sqrt{T}}{2}z_2 \\ B_{T/4} &= \frac{1}{2}(B_0 + B_{T/2}) + \sqrt{T/8}z_3 = \frac{\sqrt{T}}{4}z_1 + \frac{\sqrt{T}}{4}z_2 + \sqrt{T/8}z_3 \\ &\vdots \end{aligned} \tag{5.12}$$

where $\{z_j\}_{j=1}^N$ are independent standard normal variables.

Describing construction 5.12: we first generate the final value B_T, then sample $B_{T/2}$ conditional on the values of B_T and B_0, and proceed by progressively filling in intermediate values. Bb uses the first several coordinates of the low-discrepancy points to determine the general shape of the Brownian path, and the last few coordinates influence only the fine detail of the path. Therefore, the most important values that determine the large scale structure of Brownian motion are the first components of $\mathbf{z} = (z_1, \ldots, z_N)$.

5.5 Richardson extrapolation

Another transformation that we coupled with MISC is Richardson extrapolation [15]. In fact, applying level K_{R} (level of extrapolation) of Richardson extrapolation reduces dramatically the bias and as a consequence reduces the number of time steps N needed in the coarsest level to achieve a certain error tolerance. This means basically that Richardson extrapolation directly reduces the total dimension of the integration problem for achieving some error tolerance.

We recall that the Euler scheme has weak order one so that

$$\left|\mathrm{E}\left[f(\widehat{X}_T^h)\right] - \mathrm{E}\left[f(X_T)\right]\right| \leq Ch \tag{5.13}$$

for some constant C, all sufficiently small h and suitably smooth f. It can be easily shown that (5.13) can be improved to

$$\mathrm{E}\left[f(\widehat{X}_T^h)\right] = \mathrm{E}\left[f(X_T)\right] + ch + \mathcal{O}\left(h^2\right), \tag{5.14}$$

where c depends on f.

Applying (5.14) with discretization step $2h$, we obtain

$$\mathrm{E}\left[f(\widehat{X}_T^{2h})\right] = \mathrm{E}\left[f(X_T)\right] + 2ch + \mathcal{O}\left(h^2\right),$$

implying

$$2\mathrm{E}\left[f(\widehat{X}_T^{2h})\right] - \mathrm{E}\left[f(\widehat{X}_T^{h})\right] = \mathrm{E}\left[f(X_T)\right] + \mathcal{O}\left(h^2\right),$$

For higher levels of extrapolations, we use the following: Let us denote by $h_J = h_0 2^{-J}$ the grid sizes (where h_0 is the coarsest grid size), by K_{R} the level of the Richardson extrapolation, and by $I(J, K_{\mathrm{R}})$ the approximation of $\mathrm{E}\left[f((X_T)\right]$ by terms up to level K_{R} (leading to a weak error of order K_{R}), then we have the following recursion

$$I(J, K_{\mathrm{R}}) = \frac{2^{K_{\mathrm{R}}}\left[I(J, K_{\mathrm{R}} - 1) - I(J-1, K_{\mathrm{R}} - 1)\right]}{2^{K_{\mathrm{R}}} - 1}, \quad J = 1, 2, \ldots, K_{\mathrm{R}} = 1, 2, \ldots$$

Remark 5.2. We emphasize that through our work, we are interested in the pre-asymptotic regime (small number of time steps), and the use of Richardson extrapolation is justified by our observed experimental results in that regime (see Section **??**), which in particular, show an order one of convergence for the weak error. Although, we do not claim that the observed rates will scale well in the asymptotic regime, we observed that the pre-asymptotic regime is enough to get sufficiently accurate estimates for the option prices.

5.6 Root Finding

Without loss of generality, we can assume that the integration domain can be divided into two parts Ω_1, and Ω_2 such that the integrand f is smooth and positive in Ω_1 whereas $f(\mathbf{x}) = 0$ in Ω_2. Therefore,

$$If := \int_{\Omega_1} f(\mathbf{x}) d\mathbf{x} \tag{5.15}$$

This situation may arise when the integrand is non-differentiable or noncontinuous along the boundary between Ω_1 and Ω_2. For these problems, kinks and jumps can efficiently be identified by a one-dimensional root finding. Then, the kinks and jumps can be transformed to the boundary of integration domain such that they no longer deteriorate the performance of the numerical methods. In fact, we compute the zeros of the integrand with respect to the last dimension. In this dimension, then, e.g., Newton's method or bisection can be used to identify the point which separates Ω_1 and Ω_2. In our project, we use Newton 's iteration solver.

Let us call y the mapping such that: $y : \mathbf{z}_1 \to z^{\mathrm{kink}}$, where z^{kink} is the "location of irregularity", *i.e.*, g is not smooth at the point $\phi \circ \Phi \circ \Psi(z^{\mathrm{kink}}, \mathbf{z}_{-1})$. Generally, there might be (for given $\mathbf{z}_{-1}$

- no solution, i.e., the integrand in the definition of $h(\mathbf{z}_{-1})$ above is smooth (*best case*);
- a unique solution;
- multiple solutions.

Generally, we need to assume that we are in the first or second case. Specifically, we need that

$$\mathbf{z}_{-1} \mapsto h(\mathbf{z}_{-1}) \text{ and } \mathbf{z}_{-1} \mapsto \widehat{h}(\mathbf{z}_{-1})$$

are smooth, where $\widehat{h}$ denotes the numerical approximation of h based on a grid containing $y(\mathbf{z}_{-1})$. In particular, y itself should be smooth in $\mathbf{z}_{-1}$. This would already be challenging in practice in the third case. Moreover, in the general situation we expect the number of solutions y to increase when the discretization of the SDE gets finer.

In many situations, case 2 (which is thought to include case 1) can be guaranteed by monotonicity (I think we need to add also the growth condition. For instance, in the case of one-dimensional SDEs with z_1 representing the terminal value of the underlying Brownian motion (and $\mathbf{z}_{-1}$ representing the Brownian bridge), this can often be seen from the SDE itself. Specifically, if each increment "dX" is increasing in z_1, no matter the value of X, then the solution X_T must be increasing in z_1. This is easily seen to be true in examples such as the Black-Scholes model and the CIR process. (Strictly speaking, we have to distinguish between the continuous and discrete time solutions. In these examples, it does not matter.) On the other hand, it is also quite simple to construct counter examples, where monotonicity fails, for instance SDEs for which the "volatility" changes sign, such as a trigonometric function.[1]

Even in multi-dimensional settings, such monotonicity conditions can hold in specific situations. For instance, in case of a basket option in a multivariate Black Scholes framework, we can choose a linear combination z_1 of the terminal values of the driving Bm, such that the basket is a monotone function of z_1. (The coefficients of the linear combination will depend on the correlations and the weights of the basket.) However, in that case this may actually not correspond to the optimal "rotation" in terms of optimizing the smoothing effect.

5.7 Description of the Domain Decomposition and Suitable Transformation

The payoff function is not smooth due to the nature of the option. In fact, the holder would not exercise the option if a purchase or sale of the underlying asset would lead to a loss. As a result, the discontinuity of the payoff function carries over to the integrand. In this case, The integrand shows a kink or even a jump with respect to a manifold. Since some (mixed) derivatives are not bounded at these manifolds, the smoothness requirements for the sparse grid method are clearly not fulfilled any more.

The first step consists of identifying the areas of discontinuity or non-differentiability. Then, we decompose the total integration domain Ω into sub-domains Ω_i, $i = 1, \ldots, n$ such that the integrand is smooth in the interior of Ω_i and such that all kinks and jumps are located along the boundary of these areas. This procedure results in integrating several smooth functions, instead of one discontinuous function. The total integral is then given as the sum of the separate integrals, *i.e.*

$$If := \int_{\Omega} f(\mathbf{x}) d\mathbf{x} = \sum_{i=1}^{n} \int_{\Omega_i} f(\mathbf{x}) d\mathbf{x} \tag{5.16}$$

[1] Actually, in every such case the simple remedy is to replace the volatility by its absolute value, which does not change the law of the solution. Hence, there does not seem to be a one-dimensional counter-example.

In this way, the fast convergence of SG can be regained whereas the costs only increase by a constant (the number of terms in the sum), provided the cost required for the decomposition is sufficiently small such that it can be neglected.

In general, such a decomposition is even more expensive than to integrate the function. Nevertheless, for some problem classes, the areas of discontinuity have a particular simple form, which allows to decompose the integration domain with costs that are much smaller than the benefit which results from the decomposition. In this work, we consider those cases.

In the literature, there two classes that have been tackled. In the first one, we have the information that the kinks are part of the integration domain where the integrand is zero and can thus be identified by root finding as proposed in [5].

In the second class, we have the information that the discontinuities are located on hyperplanes, which allows a decomposition first into polyhedrons and then into orthants as discussed in [6]. In this work, we start by the first class of problems.

6 Error discussion

6.1 Errors in smoothing

For the analysis it is useful to assume that $\widehat{h}$ is a smooth function of $\mathbf{z}_{-1}$, but in reality this is not going to be true. Specifically, if the true location y of the non-smoothness in the system were available, we could actually guarantee $\widehat{h}$ to be smooth, for instance by choosing

$$\widehat{h}(\mathbf{z}_{-1}) = \sum_{k=-K}^{K} \eta_k g\left(\phi \circ \Phi \circ \Psi\left(\zeta_k(y(\mathbf{z}_{-1})), \mathbf{z}_{-1}\right)\right),$$

for points $\zeta_k \in \mathbb{R}$ with $\zeta_0 = y$ and corresponding weights η_k.[2] However, in reality we have to numerical approximate y by $\overline{y}$ with error $|y - \overline{y}| \leq \delta$. Now, the actual integrand in $\mathbf{z}_{-1}$ becomes

$$\overline{h}(\mathbf{z}_{-1}) \coloneqq \sum_{k=-K}^{K} \eta_k g\left(\phi \circ \Phi \circ \Psi\left(\zeta_k(\overline{y}(\mathbf{z}_{-1})), \mathbf{z}_{-1}\right)\right),$$

which we cannot assume to be smooth anymore. On the other hand, if $\zeta_k(y)$ is a continuous function of y and y and $\overline{y}$ are continuous in $\mathbf{z}_{-1}$, then *eventually* we will have

$$\left\|\widehat{h} - \overline{h}\right\|_\infty \leq \mathrm{TOL}, \quad \left\|h - \overline{h}\right\|_\infty \leq \mathrm{TOL},$$

i.e., the smooth functions h and $\widehat{h}$ are close to the integrand $\overline{h}$. (Of course, this may depend on us choosing a good enough quadrature ζ!)

Remark 6.1. If the adaptive collocation used for computing the integral of $\overline{h}$ depends on derivatives (or difference quotients) of its integrand $\overline{h}$, then we may also need to make sure that derivatives of $\overline{h}$ are close enough to derivatives of $\widehat{h}$ or h. This may require higher order solution methods for determining y.

[2] Of course, the points ζ_k have to be chosen in a systematic manner depending on y.

Remark 6.2. In some important cases, f may be trivial (e.g., $\equiv 0$). In these cases, we may be able to make sure that $\overline{y}$ never crosses the "location of non-smoothness". Then even $\overline{h}$ is smooth.

Remark 6.3. We expect that the global error of our procedure will be bounded by the weak error which is in our case of order $O(\Delta t)$. In this case, the overall complexity of our procedure will be of order $O(TOL^{-1})$. We note that this rate can be improved up to $O(TOL^{-\frac{1}{2}})$ if we use **Richardson extrapolation**. Another way that can improve the complexity could be based on **Cubature on Wiener Space** (This is left for a future work). The aimed complexity rate illustrates the contribution of our procedure which outperforms Monte Carlo forward Euler (MC-FE) and multi-level MC-FE, having complexity rates of order $O(TOL^{-3})$ and $O(TOL^{-2}log(TOL)^2)$ respectively.

Remark 6.4.

We need to check the impact of the error caused by the Newton iteration on the integration error. In the worst case, we expect that if the error in the Newton iteration is of order $O(\epsilon)$ than the integration error will be of order $\log(\epsilon)$. But we need to check that too.

7 Numerical experiments

7.1 The discretized 1D Black-Scholes

The first two examples that we will test later are the single binary and call options under BS model where the process X is the discretized one dimensional Black-Scholes model and the payoff function g is the indicator or maximum function, and which has a kink. Precisely, we are interested in the 1-d lognormal example where the dynamics of the stock are given by

$$dX_t = \sigma X_t dB_t, \tag{7.1}$$

where $\{B_t, 0 \leq t \leq T\}$ is a standard one-dimensional Brownian motion. In the discrete case, the numerical approximation of $X(T)$, using N time steps ($\Delta t = \frac{T}{N}$), satisfies

$$\begin{aligned} \overline{X}_T &= \Phi(\Delta t, z_1, \Delta B_0, \ldots, \Delta B_{N-1}), \\ &= \Phi(\Delta t, \Psi(z_1, \ldots, z_N)), \end{aligned} \tag{7.2}$$

wehre $(z_1, \ldots, z_N)$ are standard Gaussian random variables, for some path function Φ and Brownian bridge map Ψ as described in Section 5.4.1. As explained in Section 5, the first step of our approach is determining the location of irregularity (kink). In the following, we want to compare different ways for identifying the location of the kink for this model.

7.1.1 Determining the kink location

Exact location of the kink for the continuous problem

Let us denote y_* an invertible function that satisfies

$$X(T; y_*(x), B) = x. \tag{7.3}$$

We can easily prove that the expression of y_* for model given by (7.1) is given by

$$y_*(x) = \left(\log(x/x_0) + T\sigma^2/2\right) \frac{1}{\sqrt{T}\sigma}, \tag{7.4}$$

and since the kink for Black-Scholes model occurs at $x = K$, where K is the strike price then the exact location of the continuous problem is given by

$$y_*(K) = \left(\log(K/x_0) + T\sigma^2/2\right) \frac{1}{\sqrt{T}\sigma}. \tag{7.5}$$

Exact location of the kink for the discrete problem

The discrete problem of model (7.1) is solved by simulating

$$\begin{aligned} X_{t_1} &= X_{t_0}\left[1 + \frac{\sigma}{\sqrt{T}} z_1 \Delta t + \sigma \Delta B_0\right] \\ X_{t_2} &= X_{t_1}\left[1 + \frac{\sigma}{\sqrt{T}} z_1 \Delta t + \sigma \Delta B_1\right] \\ \vdots &= \vdots \\ X_{t_N} &= X_{t_{N-1}}\left[1 + \frac{\sigma}{\sqrt{T}} z_1 \Delta t + \sigma \Delta B_{N-1}\right] \end{aligned} \tag{7.6}$$

implying that

$$\overline{X}(T) = X_0 \prod_{i=0}^{N-1}\left[1 + \frac{\sigma}{\sqrt{T}} z_1 \Delta t + \sigma \Delta B_i\right]. \tag{7.7}$$

Therefore, in order to determine y_*, we need to solve

$$x = \overline{X}(T; y_*, B) = X_0 \prod_{i=0}^{N-1}\left[1 + \frac{\sigma}{\sqrt{T}} y_*(x) \Delta t + \sigma \Delta B_i\right], \tag{7.8}$$

which implies that the location of the kink point for the approximate problem is equivalent to finding the roots of the polynomial $P(y_*(K))$, given by

$$P(y_*(K)) = \prod_{i=0}^{N-1}\left[1 + \frac{\sigma}{\sqrt{T}} y_*(K) \Delta t + \sigma \Delta B_i\right] - \frac{K}{X_0}. \tag{7.9}$$

The exact location of the kink can be obtained exactly by solving exactly $P(y_*(K)) = 0$.

Approximate location of the discrete problem

Here, we try to find the roots of polynomial $P(y_*(K))$, given by (7.9), by using **Newton iteration method**. In this case, we need the expression $P' = \frac{dP}{dy_*}$. If we denote $f_i(y) = 1 + \frac{\sigma}{\sqrt{T}} y \Delta t + \sigma \Delta B_i$, then we can easily show that

$$P'(y) = \frac{\sigma \Delta t}{\sqrt{T}} \left(\prod_{i=0}^{N-1} f_i(y)\right)\left[\sum_{i=0}^{N-1} \frac{1}{f_i(y)}\right] \tag{7.10}$$

Therefore, in this case, the integrand $h(\mathbf{z}_{-1})$ (as expressed in (5.3)) is given by

$$h(\mathbf{z}_{-1}) = \int_{\Omega} \max\left[(\Phi \circ \Psi(T; z_1, \mathbf{z}_{-1})) - K, 0\right] \rho_1(z_1) dz_1. \tag{7.11}$$

We get the kink point by running Newton iteration for root solving of the polynomial P as expressed in (7.9) with a precision of 10^{-10}. We decompose the total integration domain Ω into sub-domains Ω_i, $i = 1, 2$ such that the integrand is smooth in the interior of Ω_i and such that the kink is located along the boundary of these areas. The total integral is then given as the sum of the separate integrals, *i.e.*

$$\begin{aligned} h(\mathbf{z}_{-1}) &:= \int_{\Omega} \max\left[(\Phi \circ \Psi(T; z_1, \mathbf{z}_{-1})) - K, 0\right] \rho_1(z_1) dz_1 \\ &= \sum_{i=1}^{2} \int_{\Omega_i} \max\left[(\Phi \circ \Psi(T; z_1, \mathbf{z}_{-1})) - K, 0\right] \rho_d(z_1) dz_1, \end{aligned} \tag{7.12}$$

where we use Gauss-laguerre quadrature with β points to get each part.

7.2 The basket call option under the discretized multi-dimensional Black-Scholes

In the following, I present two ways of solving the root finding problem in multi-dimension. The first way, presented in Section 7.2.1, is an extension of Section 7.1, and the numerical results in Section 7.6 are based on applying this way of numerical smoothing. In the second way, presented in Section 7.2.2, we tried a different approach that is inspired by the work of [2]. However, the way I did it, it seems to me that we do not need numerical smoothing and the smoothing can be done analytically.

7.2.1 First way

In this suggested way, I try to extend the way we solved the one dimensional problem, as in Section 7.1, to higher dimensions.

We consider the basket option under multi-dimensional BS model where the process $\mathbf{X}$ is the discretized d-dimensional Black-Scholes model and the payoff function g is given by

$$g(\mathbf{X}(T)) = \max\left(\sum_{i=1}^{d} \omega_i X^{(i)}(T) - K, 0\right) \tag{7.13}$$

Precisely, we are interested in the d-dimensional lognormal example where the dynamics of the stock are given by

$$dX_t^{(i)} = \sigma^{(i)} X_t^{(i)} dB_t^{(i)}, \tag{7.14}$$

where $\{B^{(1)}, \ldots, B^{(d)}\}$ are correlated Brownian motions with correlations ρ_{ij}.

In the discrete case, the numerical approximation of $X^{(j)}(T)$ satisfies

$$\begin{aligned}\overline{X}_T^{(j)} &= \Phi(\Delta t, z_1^{(j)}, \Delta B_0^{(j)}, \ldots, \Delta B_{N-1}^{(j)}), \quad 1 \le j \le d, \\ &= \Phi(\Delta t, \Psi(z_1^{(j)}, \ldots, z_N^{(j)})), \quad 1 \le j \le d,\end{aligned} \tag{7.15}$$

for some path function Φ and Brownian bridge map Ψ as described in Section 5.4.1.

Using results from section 7.1, we have

$$\overline{X}^{(j)}(T) = X_0^{(j)} \prod_{i=0}^{N-1} \left[1 + \frac{\sigma^{(j)}}{\sqrt{T}} z_1^{(j)} \Delta t + \sigma^{(j)} \Delta B_i^{(j)} \right], \quad 1 \le j \le d. \tag{7.16}$$

Therefore, in order to determine $\mathbf{y}_* = (y_*^{(1)}, \ldots, y_*^{(d)})$, we need to solve

$$\mathbf{x} = \sum_{j=1}^{d} \omega_j X_0^{(j)} \prod_{i=0}^{N-1} \left[1 + \frac{\sigma^{(j)}}{\sqrt{T}} y_*^{(j)}(\mathbf{x}) \Delta t + \sigma^{(j)} \Delta B_i^{(j)} \right], \tag{7.17}$$

which implies that the location of the kink point for the approximate problem is equivalent to finding the roots of the polynomial $P(\mathbf{y}_*(K))$, given by

$$\begin{aligned} P(\mathbf{y}_*(K)) &= \sum_{j=1}^{d} \omega_j X_0^{(j)} \prod_{i=0}^{N-1} \left[1 + \frac{\sigma^{(j)}}{\sqrt{T}} y_*^{(j)} \Delta t + \sigma^{(j)} \Delta B_i^{(j)} \right] - K, \\ &= \sum_{j=1}^{d} \omega_j X_0^{(j)} \left(\prod_{i=0}^{N-1} \left[1 + \frac{\sigma^{(j)}}{\sqrt{T}} y_*^{(j)} \Delta t + \sigma^{(j)} \Delta B_i^{(j)} \right] - \frac{K}{\omega_j X_0^{(j)} d} \right) \\ &= \sum_{j=1}^{d} \omega_j X_0^{(j)} P^{(j)} \left(y_*^{(j)} \left(\frac{K}{\omega_j X_0^{(j)} d} \right) \right) \end{aligned} \tag{7.18}$$

Therefore, the problem of finding the kink in the d-dimensional problem is brought to finding the location of the kink for each dimension by solving the root of the polynomial $P^{(j)}\left(y_*^{(j)} \left(\frac{K}{\omega_j X_0^{(j)}} \right) \right)$.

Using **Newton iteration method**, we use the expression $P' = \frac{dP^{(j)}}{dy_*^{(j)}}$. If we denote $f_i^{(j)}(y) = 1 + \frac{\sigma^{(j)}}{\sqrt{T}} y \Delta t + \sigma^{(j)} \Delta B_i^{(j)}$, then we can easily show that

$$P'^{(j)}(y) = \frac{\sigma^{(j)} \Delta t}{\sqrt{T}} \left(\prod_{i=0}^{N-1} f_i^{(j)}(y) \right) \left[\sum_{i=0}^{N-1} \frac{1}{f_i^{(j)}(y)} \right]. \tag{7.19}$$

Therefore, in this case, the integrand $h(\mathbf{z}_{-1}^{(1)}, \ldots, \mathbf{z}_{-1}^{(d)})$ (as expressed in (5.6)) is given by

$$h(\mathbf{z}_{-1}^{(1)}, \ldots, \mathbf{z}_{-1}^{(d)}) = \int_\Omega \max \left[\left(\sum_{j=1}^{d} \Phi \circ \Psi(T; z_1^{(j)}, \mathbf{z}_{-1}^{(j)}) \right) - K, 0 \right] \rho_d(z_1^{(1)}, \ldots, z_1^{(d)}) dz_1^{(1)} \ldots dz_1^{(d)}. \tag{7.20}$$

We get the kink point by running Newton iteration in each dimension seperately for root solving of each polynomial $P^{(j)}$ as expressed in (7.18) with a precision of 10^{-10}. We decompose the total integration domain Ω into sub-domains Ω_i, $i = 1, 2, \ldots, 2^d$ such that the integrand is smooth in the interior of Ω_i and such that the kink is located along the boundary of these areas. The total integral is then given as the sum of the separate integrals, *i.e.*

$$h(\mathbf{z}_{-1}^{(1)}, \ldots, \mathbf{z}_{-1}^{(d)}) := \int_{\Omega} \max\left[\left(\sum_{j=1}^{d} \Phi \circ \Psi(T; z_1^{(j)}, \mathbf{z}_{-1}^{(j)})\right) - K, 0\right] \rho_d(z_1^{(1)}, \ldots, z_1^{(d)}) dz_1^{(1)} \ldots dz_1^{(d)}$$

$$= \sum_{i=1}^{2^d} \int_{\Omega_i} \max\left[\left(\sum_{j=1}^{d} \Phi \circ \Psi(T; z_1^{(j)}, \mathbf{z}_{-1}^{(j)})\right) - K, 0\right] \rho_d(z_1^{(1)}, \ldots, z_1^{(d)}) dz_1^{(1)} \ldots dz_1^{(d)}, \tag{7.21}$$

where we use Gauss-laguerre quadrature with β points to get each part.

7.2.2 Second way

The i-th asset $X^{(i)}$ of the basket ($i = 1, \ldots, d$) is given by

$$X_{k\Delta t}^{(i)} = X_0^{(i)} \exp\left(-\frac{\sigma_i^2}{2} k\Delta t + \sigma_i B_{k\Delta t}^{(i)}\right), \quad 1 \le k \le N, \tag{7.22}$$

where $X_0^{(i)}$ is the current price of the i-th asset, σ_i is the volatility of the i-th asset and $B = (B^{(1)}, \ldots, B^{(d)})$ is a d-dimensional Brownian motion. The correlation between $B^{(i)}$ and $B^{(j)}$ is denoted by ρ_{ij}. The payoff function of the European basket option is given by

$$f(Z) = \max\left(\sum_{i=1}^{d} c_i X_T^{(i)}(Z) - K, 0\right), \tag{7.23}$$

where c_i is the corresponding weight of the i-th asset, $Z \in \mathbb{R}^{N \times d}$ is standard Gaussian vector, and

$$X_{k\Delta t}^{(i)}(Z) = X_0^{(i)} e^{-\frac{\sigma_i^2}{2} k\Delta t} \exp\left(\sum_{j=1}^{dN} C_{(k-1)d+i,j} Z_j\right). \tag{7.24}$$

where C is a $dN \times dN$-matrix with $CC^T = \overline{\Sigma} := R \otimes \Sigma$ (Σ is $N \times N$ matrix given by (5.10)), and R is an $d \times d$-matrix with $R_{ij} = \sqrt{T/N}\rho_{ij}\sigma_i\sigma_j$.

Therefore, we have

$$\begin{aligned} X_T^{(i)}(Z) = X_{N\Delta t}^{(i)}(Z) &= X_0^{(i)} e^{-\frac{\sigma_i^2}{2} T} \exp\left(\sum_{j=1}^{dN} C_{(N-1)d+i,j} Z_j\right) \\ &= \tilde{w}^{(i)} \exp\left(A_{i*} Z\right) \\ &= \tilde{w}^{(i)} \exp\left(\widetilde{Z}^{(i)}\right), \end{aligned} \tag{7.25}$$

where $\tilde{w}^{(i)} = X_0^{(i)} e^{-\frac{\sigma_i^2}{2}T}$, A is a $d \times dN$ matrix such that $A := \left(C_{(N-1)d+i,j}\right)_{1\le i\le d,\, 1\le j\le dN}$, A_{i*} is the i-th row vector of A, and $\widetilde{Z} = AZ$.

Let us denote by $\widetilde{\Sigma}$ the covariance matrix of $\widetilde{Z}$, then $\widetilde{\Sigma} = AA^T$. If we use the same approach in [2], Specifically Lemma 3.1, then we can write

$$\widetilde{\Sigma} = \widetilde{V}\widetilde{D}\widetilde{V}^T \tag{7.26}$$

such that $\tilde{D} = \mathrm{diag}(\lambda_1^2, \ldots, \lambda_d^2)$ is $d \times d$ diagonal matrix and $\widetilde{V} \in \mathbb{R}^{d\times d}$ is an invertibe matrix, with the property that $\widetilde{V}_{i,1} \equiv 1$, $i = 1, \ldots, d$.

Going back to (7.23), and replacing $\widetilde{Z}$ by $\widetilde{V}Y$, such that $Y := \widetilde{V}^{-1}\widetilde{Z} \sim \mathcal{N}(0, \widetilde{D})$, we obtain

$$\begin{aligned}
f(Z) &= \max\left(\sum_{i=1}^{d} c_i X_T^{(i)}(Z) - K, 0\right) \\
&= \max\left(\sum_{i=1}^{d} w^{(i)} \exp\left(\widetilde{Z}^{(i)}\right) - K, 0\right) \\
&= \max\left(\sum_{i=1}^{d} w^{(i)} \exp\left((\widetilde{V}Y)^{(i)}\right) - K, 0\right) \\
&= \max\left(\sum_{i=1}^{d} w^{(i)} \exp\left(Y_1 + \sum_{j=2}^{d} \widetilde{V}_{ij} Y_j\right) - K, 0\right) \\
&= \max\left(h(Y_2, \ldots, Y_d) e^{Y_1} - K, 0\right), \qquad (7.27)
\end{aligned}$$

where $w^{(i)} = c_i \tilde{w}^{(i)}$ and $h(y_2, \ldots, y_d) := \sum_{i=1}^{d} w^{(i)} \exp\left(\sum_{i=1}^{d} w^{(i)} \sum_{j=2}^{d} \widetilde{V}_{ij} Y_j\right)$.

Giving (7.27), and Lemma 3.3 in [2], it seems to me that we do not need numerical smoothing and the smoothing can be done analytically using conditional expectation formula.

7.3 Summary of numerical results

We conduct our experiments for 3 different examples under discretized BS model: i) single binary, ii) single call, and basket call options (We just includeded results for 2-dimensional basket following the way suugested in Section 7.2.1, I still have issue for higher dimensions that we need to discuss, and this is maybe related to the way I am solving the problem in high dimension).

In Sections 7.4.1 ,7.5.1 and 7.6.1, we estimate the weak error (Bias) of MC combined with root finding, for the different examples, for 2 scenarios involving with/without Richardson extrapolation. The conclusions of this section are:

- Without Richardson extrapolation: For all cases, we get a weak error of order Δt, with different constants (See tables (7.2, 7.12,7.22) for the corresponding bias values as well the statistical errors).
- With Richardson extrapolation: For the case of binary and 2-dimensional basket call options, we get a weak error of order almost Δt^2. For the case of single call, we get a weak error of order higher than Δt^2 (See tables (7.7,7.17) for the corresponding bias values as well the statistical errors).

Remark 7.1. We emphasize that the reported weak rates correspond to the pre-asymptotic regime that we are interested in. We are not interested on estimating the rates specifically but rather a sufficient precise estimate of the weak error (Bias), $\mathcal{E}_B(N)$, for different time steps N, in order to get the biased MC solution for a given discretization, that we denoted $Q^N(\infty)$ in Section 5.3. For a fixed discretization, the corresponding biased solution will be set as a reference solution to the MISC method in order to estimate the quadrature error $\mathcal{E}_Q(TOL_{\text{MISC}}, N)$.

In Sections 7.4.2, 7.5.2 and 7.6.2, we show tables and plots reporting the different errors involved in MC method (bias and statistical error), and in MISC (Quadrature error). The quadrature error (see (5.8)) is computed by subtracting the MISC solution from the biased solution, computed with sufficiently large number of samples (to kill the statistical error). Given that both methods, MC and MISC, have the same bias, the computational time of MC and MISC is compared such that the statistical error is almost equal to the stable quadrature error produced by MISC.

The conclusions of those sections are:

- For the case of single binary option (see Section 7.4), MISC coupled with Richardson extrapolation is 3.5 times faster than MC coupled with Richardson extrapolation, to achieve a total relative error around 0.3% (See tables (7.9,7.10)). Applying Richardson extrapolation brought a significant improvement for MISC (compare tables (7.4,7.5) (no Richardson), tables (7.9,7.10) (Richardson (level 1)) and corresponding complexity plots).
- For the case of single call option (see Section 7.5), MISC is 241 times faster than MC, to achieve a total relative error below 0.5% (See tables (7.14,7.15)). Applying Richardson extrapolation brought a significant improvement for MISC (compare tables (7.14,7.15) (no Richardson), tables (7.19,7.20) (Richardson (level 1)) and corresponding complexity plots).
- For the case of 2-dimensional basket call option (see Section 7.6), MISC is 80 times faster than MC, to achieve a total relative error around 0.8% (See tables (7.24,7.25)).

7.4 Results for the single binary option example

In this case, the integrand $h(\mathbf{z}_{-1})$ is given by

$$\begin{aligned} h(\mathbf{z}_{-1}) &= \int \mathbf{1}_{\Phi\circ\Psi(T;z_1,\mathbf{z}_{-1})>K} \frac{1}{\sqrt{2\pi}} \exp(-z_1^2/2) dy \\ &= P(Y > y_*(K)), \end{aligned} \tag{7.28}$$

where $y_*(x)$, is an invertible function that satisfies

$$\Phi \circ \Psi(T; y_*(x), \mathbf{z}_{-1}) = x \tag{7.29}$$

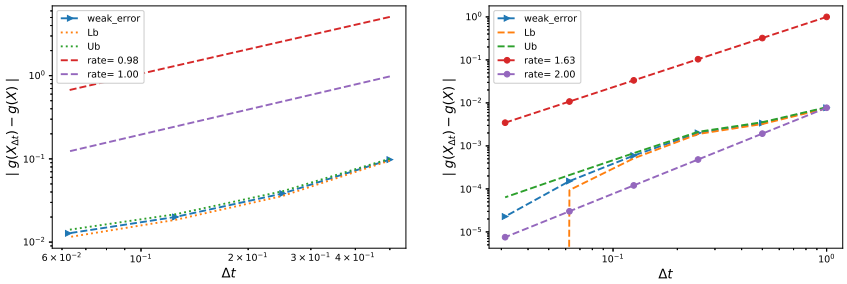
We get the kink point by running Newton iteration with a precision of 10^{-10}.

The paramters that we used in our numerical experiments are: $T = 1$, $\sigma = 0.4$ and $S_0 = K = 100$. The exact value of this case is 0.42074029.

7.4.1 Weak error plots

In this section, we include the results of weak error rates, for the binary option example, for 2 scenarios, without/with Richardson extrapolation (level 1). We note that the weak errors plotted here correspond to relative errors. The upper and lower bounds are 95% confidence interval.

From figure 7.1, we see that we get a weak error of order Δt for the case without Richardson and we observe an improvement in the rate and the constant when using level 1 of Richardson extrapolation. The corresponding values of the Bias are reported in tables (7.2,7.7).



(a) (b)

Figure 7.1: The rate of convergence of the weak error for the binary option. a) $|\mathrm{E}\left[g(X_{\Delta t})\right] - g(X)|$ b)$\left|\mathrm{E}\left[2g(X_{\Delta t/2}) - g(X_{\Delta t})\right] - g(X)\right|$.

7.4.2 Comparing relative errors

Without Richardson extrapolation

Method \Steps	2	4	8	16
MISC ($TOL_{\text{MISC}} = 5.10^{-1}$)	0.4620	0.4404	0.4299	0.4250
MISC ($TOL_{\text{MISC}} = 10^{-1}$)	0.4620	0.4404	0.4301	0.4250
MISC ($TOL_{\text{MISC}} = 5.10^{-2}$)	0.4620	0.4403	0.4300	0.4250
MISC ($TOL_{\text{MISC}} = 10^{-2}$)	0.4620	0.4406	0.4300	0.4250
MISC ($TOL_{\text{MISC}} = 10^{-3}$)	0.4620	0.4406	0.4301	0.4254
MC method ($M = 10^4$)	0.4620	0.4369	0.4292	0.4261

Table 7.1: Binary option price of the different methods for different number of time steps, without Richardson extrapolation.

Method \Steps	2	4	8	16
MC Bias ($M = 10^4$)	**0.0980** (0.0412)	**0.0384** (0.0162)	**0.0201** (0.0085)	**0.0127** (0.0053)
MC Statistical error ($M = 10^4$)	**1.2e − 03** (5.0e−04)	**1.2e − 03** (5.0e−04)	**8.0e − 04** (3.4e−04)	**6.5e − 04** (2.7e−04)

Table 7.2: Bias and statistical errors of MC for computing binary option price for different number of time steps, without Richardson extrapolation. The numbers between parentheses are the corresponding absolute errors.

Method \Steps	2	4	8	16
MISC ($TOL_{\text{MISC}} = 5.10^{-1}$)	**1.0e − 05** (2.4e−05)	**0.0083** (0.0035)	**0.0017** (0.0007)	**0.0026** (0.0011)
MISC ($TOL_{\text{MISC}} = 10^{-1}$)	**1.0e − 05** (2.4e−05)	**0.0083** (0.0035)	**0.0021** (0.0009)	**0.0026** (0.0011)
MISC ($TOL_{\text{MISC}} = 5.10^{-2}$)	**1.0e − 05** (2.4e−05)	**0.0081** (0.0034)	**0.0019** (0.0008)	**0.0026** (0.0011)
MISC ($TOL_{\text{MISC}} = 10^{-2}$)	**1.0e − 05** (2.4e−05)	**0.0088** (0.0037)	**0.0019** (0.0008)	**0.0026** (0.0011)
MISC ($TOL_{\text{MISC}} = 10^{-3}$)	**1.0e − 05** (2.4e−05)	**0.0088** (0.0037)	**0.0021** (0.0009)	**0.0017** (0.0007)

Table 7.3: Quadrature error of MISC, with different tolerances, to compute binary option price for different number of time steps, without Richardson extrapolation. The numbers between parentheses are the corresponding absolute errors. The values marked in red correspond to stable quadrature errors for MISC, and will be used for complexity comparison against MC.

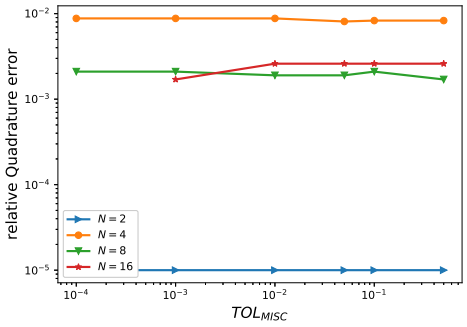


Figure 7.2: Relative quadrature error of MISC, with different tolerances, to compute binary option price for different number of time steps, without Richardson extrapolation.

Method \Steps	4	8	16
MISC ($TOL_{\text{MISC}} = 5.10^{-1}$)	**0.0467**	**0.0218**	**0.0153**
MISC ($TOL_{\text{MISC}} = 10^{-1}$)	**0.0467**	**0.0222**	**0.0153**
MISC ($TOL_{\text{MISC}} = 5.10^{-2}$)	**0.0465**	**0.0220**	**0.0153**
MISC ($TOL_{\text{MISC}} = 10^{-2}$)	**0.0472**	**0.0220**	**0.0153**
MISC ($TOL_{\text{MISC}} = 10^{-3}$)	**0.0472**	**0.0222**	**0.0144**
MC+root finding	**0.0471**	**0.0221**	**0.0141**
MC	**0.0467**	**0.0222**	**0.0146**

Table 7.4: Total relative error of MISC, with different tolerances, and MC to compute binary option price for different number of time steps, without Richardson extrapolation. The values marked in red, for MISC method, correspond to the total relative errors associated with stable quadrature errors for MISC, and will be used for complexity comparison against MC.

Method \Steps	4	8	16
MISC ($TOL_{\text{MISC}} = 5.10^{-1}$)	0.8	2	9
MISC ($TOL_{\text{MISC}} = 10^{-1}$)	0.8	9	49
MISC ($TOL_{\text{MISC}} = 5.10^{-2}$)	1.3	12	59
MISC ($TOL_{\text{MISC}} = 10^{-2}$)	2	14	63
MISC ($TOL_{\text{MISC}} = 10^{-3}$)	2	34	1090
MC+root finding method	9	96	141
MC method	7	95	152
Ratio of (MC+root finding)/(MISC)	4.5	11	0.13
Ratio of (MC)/(MISC)	3.5	11	0.14

Table 7.5: Comparison of the computational time of MC and MISC, used to compute binary option price for different number of time steps, without Richardson extrapolation. The average computational time of MC is computed over 10 runs.

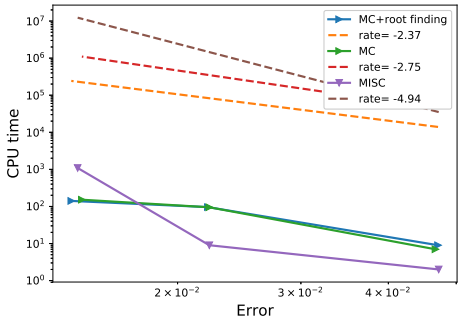


Figure 7.3: Complexity plot for MC and MISC for the case without Richardson extrapolation.

With Richardson extrapolation (level 1)

Method \Steps	$1-2$	$2-4$	$4-8$	$8-16$
MISC ($TOL_{\text{MISC}} = 5.10^{-1}$)	0.4239	0.4188	0.4191	0.4200
MISC ($TOL_{\text{MISC}} = 10^{-1}$)	0.4239	0.4188	0.4191	0.4199
MISC ($TOL_{\text{MISC}} = 5.10^{-2}$)	0.4239	0.4188	0.4190	0.4199
MISC ($TOL_{\text{MISC}} = 10^{-2}$)	0.4239	0.4192	0.4194	0.4199
MC method ($M = 5.10^6$)	0.4240	0.4224	0.4216	0.4210

Table 7.6: Binary option price of the different methods for different number of time steps, with Richardson extrapolation (level 1).

Method \Steps	$1-2$	$2-4$	$4-8$	$8-16$
MC Bias ($M = 5.10^6$)	**0.0077** (0.0032)	**0.0039** (0.0016)	**0.0020** (0.0008)	**0.0006** (0.0003)
MC Statistical error ($M = 5.10^6$)	**1.1e − 04** (4.6e−05)	**8.4e − 05** (3.5e−05)	**6.0e − 05** (2.5e−05)	**4.2e − 05** (1.8e−05)

Table 7.7: Bias and statistical errors of MC for computing binary option price for different number of time steps, with Richardson extrapolation (level 1). The numbers between parentheses are the corresponding absolute errors.

Method \Steps	$1-2$	$2-4$	$4-8$	$8-16$
MISC ($TOL_{\text{MISC}} = 5.10^{-1}$)	**2.4e − 04** (1.0e−04)	**8.6e − 03** (3.6e−03)	**5.9e − 03** (2.5e−03)	**2.4e − 03** (1.0e−03)
MISC ($TOL_{\text{MISC}} = 10^{-1}$)	**2.4e − 04** (1.0e−04)	**8.6e − 03** (3.6e−03)	**5.9e − 03** (2.5e−03)	**2.6e − 03** (1.1e−04)
MISC ($TOL_{\text{MISC}} = 5.10^{-2}$)	**2.4e − 04** (1.0e−04)	**8.6e − 03** (3.6e−03)	**6.2e − 03** (2.6e−03)	**2.6e − 03** (1.1e−04)
MISC ($TOL_{\text{MISC}} = 10^{-2}$)	**2.4e − 04** (1.0e−04)	**7.6e − 03** (3.2e−03)	**5.2e − 03** (2.2e−03)	**2.6e − 03** (1.1e−04)

Table 7.8: Quadrature error of MISC, with different tolerances, to compute binary option price for different number of time steps, with Richardson extrapolation (level 1). The numbers between parentheses are the corresponding absolute errors. The values marked in red correspond to stable quadrature errors for MISC, and will be used for complexity comparison against MC.

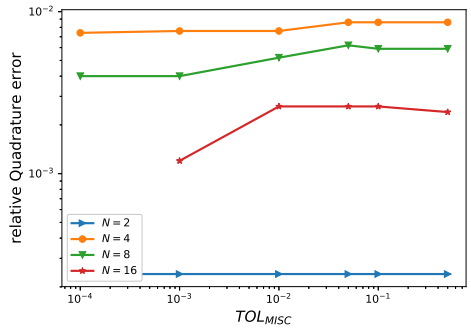


Figure 7.4: Relative quadrature error of MISC, with different tolerances, to compute binary option price of the different tolerances for different number of time steps, with Richardson extrapolation.

Method \Steps	$2-4$	$4-8$	$8-16$
MISC ($TOL_{\text{MISC}} = 5.10^{-1}$)	**0.0125**	**0.0079**	**0.0030**
MISC ($TOL_{\text{MISC}} = 10^{-1}$)	**0.0125**	**0.0079**	**0.0032**
MISC ($TOL_{\text{MISC}} = 5.10^{-2}$)	**0.0125**	**0.0082**	**0.0032**
MISC ($TOL_{\text{MISC}} = 10^{-2}$)	**0.0115**	**0.0072**	**0.0032**
MC+root finding	**0.0111**	**0.0070**	**0.0035**
MC	**0.0116**	**0.0073**	**0.0029**

Table 7.9: Total relative error of MISC, with different tolerances, and MC to compute binary option price for different number of time steps, with Richardson extrapolation (level 1). The values marked in red, for MISC method, correspond to the total relative errors associated with stable quadrature errors for MISC, and will be used for complexity comparison against MC.

Method \Steps	$2-4$	$4-8$	$8-16$
MISC ($TOL_{\text{MISC}} = 5.10^{-1}$)	1	4	9
MISC ($TOL_{\text{MISC}} = 10^{-1}$)	1	8	42
MISC ($TOL_{\text{MISC}} = 5.10^{-2}$)	1	10	72
MISC ($TOL_{\text{MISC}} = 10^{-2}$)	4	15	78
MC+root finding method	48	62	122
MC	12	23	147
Ratio of (MC+root finding)/(MISC)	12	4.1	2.9
Ratio of (MC)/(MISC)	3	1.5	3.5

Table 7.10: Comparison of the computational time of MC and MISC, used to compute binary option price for different number of time steps, with Richardson extrapolation (level 1). The average computational time of MC is computed over 10 runs.

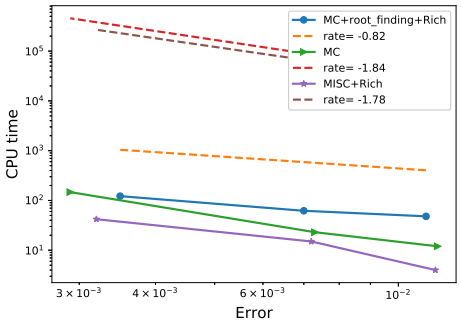


Figure 7.5: Complexity plot for MC and MISC for the case with Richardson extrapolation.

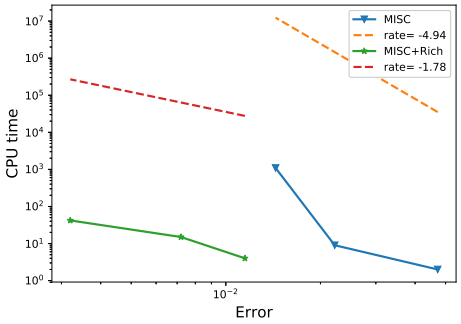


Figure 7.6: Complexity plot for MISC without and with Richardson extrapolation, for the binary option.

7.5 Results for the single call option example

In this case, the integrand $h(\mathbf{z}_{-1})$ is given by

$$h(\mathbf{z}_{-1}) = \int_{\Omega} \max\left(\Phi \circ \Psi(T; z_1, \mathbf{z}_{-1}) - K, 0\right) \frac{1}{\sqrt{2\pi}} \exp(-z_1^2/2) dz_1 \tag{7.30}$$

We get the kink point by running Newton iteration with a precision of 10^{-10}. We decompose the total integration domain Ω into sub-domains Ω_i, $i = 1, 2$ such that the integrand is smooth in the interior of Ω_i and such that the kink is located along the boundary of these areas. The total integral is then given as the sum of the separate integrals, *i.e.*

$$\begin{aligned} h(\mathbf{z}_{-1}) &:= \int_{\Omega} \max\left(\Phi \circ \Psi(T; z_1, \mathbf{z}_{-1}) - K, 0\right) \frac{1}{\sqrt{2\pi}} \exp(-z_1^2/2) dy \\ &= \sum_{i=1}^{2} \int_{\Omega_i} \max\left(\Phi \circ \Psi(T; z_1, \mathbf{z}_{-1}) - K, 0\right) \frac{1}{\sqrt{2\pi}} \exp(-z_1^2/2) dz_1, \end{aligned} \tag{7.31}$$

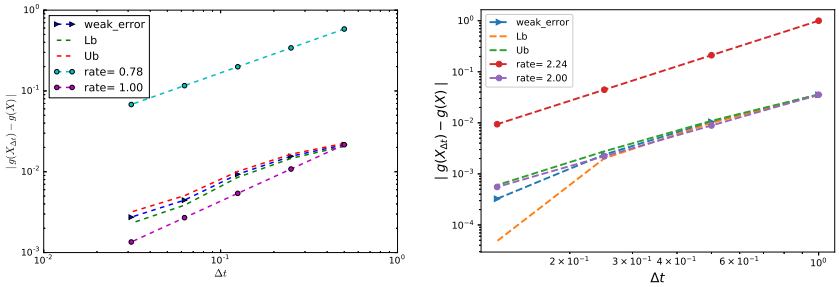
where we use Gauss-laguerre quadrature with β points to get each part.

The paramters that we used in our numerical experiments are: $T = 1$, $\sigma = 0.4$ and $S_0 = K = 100$. The exact value of this case is 15.85193755.

7.5.1 Weak error plots

In this section, we include the results of weak error rates for the call option for 2 scenarios, without/with Richardson extrapolation (level 1). We note that the weak errors plotted here correspond to relative errors. The upper and lower bounds are 95% confidence interval.

We can see from figure 7.7 that we get a weak error of order Δt for the case without Richardson extrapolation and an improvement in the rate and the constant when using level 1 of Richardson extrapolation, approximately of order Δt^2.



(a) (b)

Figure 7.7: The rate of convergence of the weak error for the call option using MC. a) $|\mathrm{E}\left[g(X_{\Delta t})\right] - g(X)|$ b)$\left|\mathrm{E}\left[2g(X_{\Delta t/2}) - g(X_{\Delta t})\right] - g(X)\right|$

7.5.2 Comparing relative errors

Without Richardson extrapolation

Method \Steps	2	4	8	16
MISC ($TOL_{\text{MISC}} = 5.10^{-1}, \beta = 32$)	16.184	16.070	15.998	15.930
MISC ($TOL_{\text{MISC}} = 10^{-1}, \beta = 32$)	16.184	16.070	15.996	15.928
MISC ($TOL_{\text{MISC}} = 5.10^{-2}, \beta = 32$)	16.184	16.070	15.996	15.928
MISC ($TOL_{\text{MISC}} = 10^{-2}, \beta = 32$)	16.184	16.103	15.996	15.928
MC method ($M = 10^5$)	16.194	16.099	15.999	15.923

Table 7.11: Call option price of the different methods for different number of time steps, without Richardson extrapolation.

Method \Steps	2	4	8	16
MC Bias ($M = 10^5$)	**0.0216** (0.3424)	**0.0156** (0.2473)	**0.0093** (0.1474)	**0.0045** (0.0713)
MC Statistical error ($M = 10^5$)	**4.4e − 04** (7.5e−03)	**4.7e − 04** (7.5e−03)	**4.0e − 04** (6.3e−03)	**3.1e − 04** (4.9e−03)

Table 7.12: Bias and statistical errors of MC for computing call option price for different number of time steps, without Richardson extrapolation. The numbers between parentheses are the corresponding absolute errors.

Method \Steps	2	4	8	16
MISC ($TOL_{\text{MISC}} = 5.10^{-1}$)	**0.0007** (0.0103)	**0.0018** (0.0290)	**6.3e − 05** (0.0010)	**0.0004** (0.0070)
MISC ($TOL_{\text{MISC}} = 10^{-1}$)	**0.0007** (0.0103)	**0.0018** (0.0290)	**0.0002** (0.0030)	**0.0001** (0.0020)
MISC ($TOL_{\text{MISC}} = 5.10^{-2}$)	**0.0007** (0.0103)	**0.0018** (0.0290)	**0.0002** (0.0030)	**0.0001** (0.0020)
MISC ($TOL_{\text{MISC}} = 10^{-2}$)	**0.0007** (0.0103)	**0.0003** (0.0040)	**0.0002** (0.0030)	**0.0001** (0.0020)

Table 7.13: Quadrature error of MISC, with different tolerances, to compute call option price for different number of time steps, without Richardson extrapolation. The numbers between parentheses are the corresponding absolute errors. The values marked in red correspond to stable quadrature errors for MISC, and will be used for complexity comparison against MC.

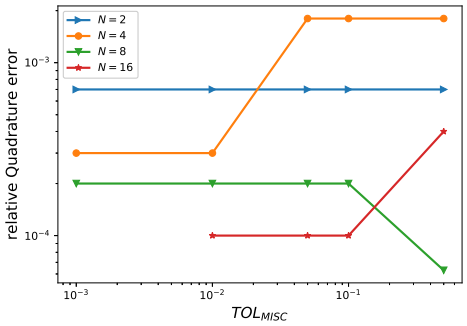


Figure 7.8: Relative quadrature error of MISC, with different tolerances, to compute call option price for different number of time steps, without Richardson extrapolation.

Method \Steps	2	4	8	16
MISC ($TOL_{\text{MISC}} = 5.10^{-1}$)	**0.0223**	**0.0174**	**0.0094**	**0.0049**
MISC ($TOL_{\text{MISC}} = 10^{-1}$)	**0.0223**	**0.0174**	**0.0095**	**0.0046**
MISC ($TOL_{\text{MISC}} = 5.10^{-2}$)	**0.0223**	**0.0174**	**0.0095**	**0.0046**
MISC ($TOL_{\text{MISC}} = 10^{-2}$)	**0.0223**	**0.0159**	**0.0095**	**0.0046**
MC +root finding	**0.0223**	**0.0159**	**0.0095**	**0.0046**
MC	**0.0223**	**0.0159**	**0.0095**	**0.0046**

Table 7.14: Total relative error of MISC, with different tolerances, and MC to compute call option price for different number of time steps, without Richardson extrapolation. The values marked in red, for MISC method, correspond to the total relative errors associated with stable quadrature errors for MISC, and will be used for complexity comparison against MC.

Method \Steps	2	4	8	16
MISC ($TOL_{\text{MISC}} = 5.10^{-1}$)	0.3	3	17	473
MISC ($TOL_{\text{MISC}} = 10^{-1}$)	0.3	3	58	656
MISC ($TOL_{\text{MISC}} = 5.10^{-2}$)	0.3	3	73	731
MISC ($TOL_{\text{MISC}} = 10^{-2}$)	0.3	6	108	1972
MC method +root finding	1328	8140	21400	70200
MC method	1450	9990	32790	158108
Ratio of (MC+root finding)/(MISC)	4400	1400	369	107
Ratio of (MC)/(MISC)	4800	1665	565	241

Table 7.15: Comparison of the computational time of MC and MISC, used to compute call option price for different number of time steps, without Richardson extrapolation. The average computational time of MC is computed over 10 runs.

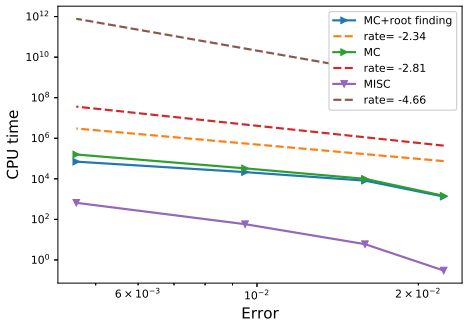


Figure 7.9: Complexity plot for MC and MISC for the case without Richardson extrapolation.

With Richardson extrapolation (level 1)

Method \Steps	$1-2$	$2-4$	$4-8$	$8-16$
MISC ($TOL_{\text{MISC}} = 5.10^{-1}$)	16.4108	16.0254	15.8912	15.8621
MISC ($TOL_{\text{MISC}} = 10^{-1}$)	16.4108	16.0254	15.8883	15.8603
MISC ($TOL_{\text{MISC}} = 5.10^{-2}$)	16.4108	16.0218	15.8885	15.8600
MC method ($M = 5.10^6$)	16.4147	16.0184	15.8900	15.8567

Table 7.16: Call option price of the different methods for different number of time steps, with Richardson extrapolation (level 1).

Method \Steps	$1-2$	$2-4$	$4-8$	$8-16$
MC Bias ($M = 5.10^6$)	**0.0355** (0.5627)	**0.0105** (0.1664)	**0.0024** (0.0380)	**0.0003** (0.0048)
MC Statistical error ($M = 5.10^6$)	**2.8e − 04** (4.4e−03)	**2.4e − 04** (3.8e−03)	**1.9e − 04** (3.0e−03)	**1.4e − 04** (2.2e−03)

Table 7.17: Bias and statistical errors of MC for computing Call option price for different number of time steps, with Richardson extrapolation (level 1). The numbers between parentheses are the corresponding absolute errors.

Method \Steps	$1-2$	$2-4$	$4-8$	$8-16$
MISC ($TOL_{\text{MISC}} = 5.10^{-1}$)	**2.5e − 04** (0.0039)	**4.4e − 04** (0.0070)	**7.6e − 05** (0.0012)	**3.4e − 04** (0.0054)
MISC ($TOL_{\text{MISC}} = 10^{-1}$)	**2.5e − 04** (0.0039)	**4.4e − 04** (0.0070)	**1.1e − 04** (0.0017)	**2.3e − 04** (0.0036)
MISC ($TOL_{\text{MISC}} = 5.10^{-2}$)	**2.5e − 04** (0.0039)	**2.1e − 04** (0.0034)	**9.5e − 05** (0.0015)	**2.1e − 04** (0.0033)

Table 7.18: Quadrature error of MISC, with different tolerances, to compute call option price for different number of time steps, with Richardson extrapolation (level 1). The numbers between parentheses are the corresponding absolute errors. The values marked in red correspond to stable quadrature errors for MISC, and will be used for complexity comparison against MC.

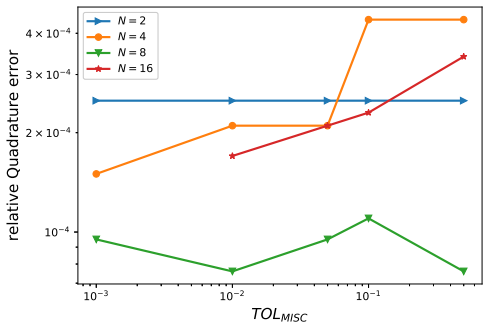


Figure 7.10: Relative quadrature error of MISC, with different tolerances, to compute call option price for different number of time steps, with Richardson extrapolation.

Method \Steps	$1-2$	$2-4$	$4-8$	$8-16$
MISC ($TOL_{\text{MISC}} = 5.10^{-1}$)	**0.0358**	**0.0109**	**0.0025**	**0.0006**
MISC ($TOL_{\text{MISC}} = 10^{-1}$)	**0.0358**	**0.0109**	**0.0025**	**0.0005**
MISC ($TOL_{\text{MISC}} = 5.10^{-2}$)	**0.0358**	**0.0107**	**0.0025**	**0.0005**

Table 7.19: Total relative error of MISC, with different tolerances, to compute call option price for different number of time steps, with Richardson extrapolation (level 1). The values marked in red, for MISC method, correspond to the total relative errors associated with stable quadrature errors for MISC, and will be used for complexity comparison against MC.

Method \Steps	$1-2$	$2-4$	$4-8$	$8-16$
MISC ($TOL_{\text{MISC}} = 5.10^{-1}$)	0.3	4	56	713
MISC ($TOL_{\text{MISC}} = 10^{-1}$)	0.3	4	107	1126
MISC ($TOL_{\text{MISC}} = 5.10^{-2}$)	0.3	9	135	1253

Table 7.20: Comparison of the computational time of MISC, used to compute call option price for different number of time steps, with Richardson extrapolation (level 1). The average computational time of MC is computed over 10 runs.

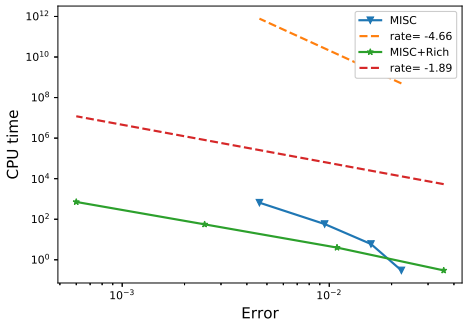


Figure 7.11: Complexity plot for MISC without and with Richardson extrapolation, for the call option.

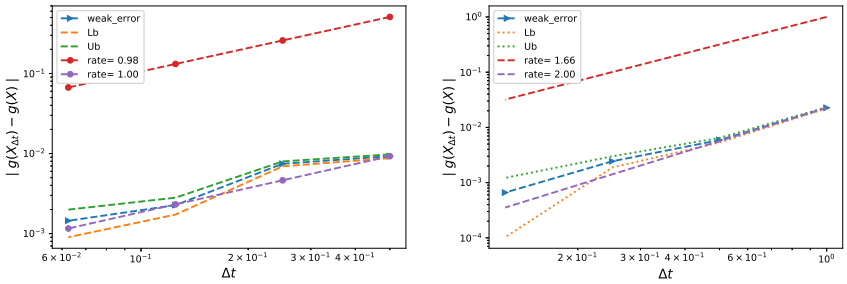
7.6 Result for the 2-dimensional Basket call option

7.6.1 Weak error plots

We consider the case of 2-dimensional Basket call option, with parameters: $S^{(1,2)} = 100$, $K = 100$, $\sigma^{(1,2)} = 0.4$, $\rho = 0.3$, $r = 0$, $T = 1$. The MC value of this case (with large number of samples) is 12.900784 with a statistical error equal to 0.001033 (Provided by Premia).

In this section, we include the results of weak error rates for 2 scenarios, without/with Richardson extrapolation (level 1). We note that the weak errors plotted here correspond to relative errors.

We can see from figure 7.12 that we get a weak error of order Δt for the case without Richardson extrapolation and almost Δt^2 for the case with Richardson extrapolation.



(a) (b)

Figure 7.12: The rate of convergence of the weak error for the Basket (2-dimensional) call option using MC. a) $|\mathrm{E}\left[g(X_{\Delta t})\right] - g(X)|$ b) $\left|\mathrm{E}\left[2g(X_{\Delta t/2}) - g(X_{\Delta t})\right] - g(X)\right|$

7.6.2 Comparing relative errors

Without Richardson extrapolation

Method \Steps	2	4	8	16
MISC ($TOL_{\text{MISC}} = 5.10^{-1}, \beta = 16$)	13.0009	12.8577	12.8854	12.9000
MISC ($TOL_{\text{MISC}} = 10^{-1}, \beta = 16$)	13.0009	12.9912	12.9556	12.9231
MISC ($TOL_{\text{MISC}} = 5.10^{-2}, \beta = 16$)	13.0009	13.0049	12.9550	–
MISC ($TOL_{\text{MISC}} = 10^{-2}, \beta = 16$)	13.0548	13.0046	12.9550	–
MC method ($M = 4.10^7$)	13.0203	12.9969	12.9301	12.9194

Table 7.21: Basket (2-dimensional) call option price of the different methods for different number of time steps, without Richardson extrapolation.

Method \Steps	2	4	8	16
MC Bias ($M = 4.10^7$)	**0.0093** (0.1195)	**0.0075** (0.0968)	**0.0023** (0.0297)	**0.0014** (0.0181)
MC Statistical error ($M = 4.10^7$)	**2.5e − 04** (3.2e−03)	**2.7e − 04** (3.5e−03)	**2.8e − 04** (3.6e−03)	**2.8e − 04** (3.6e−03)

Table 7.22: Bias and statistical errors of MC for computing Basket (2-dimensional) call option price for different number of time steps, without Richardson extrapolation. The numbers between parentheses are the corresponding absolute errors.

Method \Steps	2	4	8	16
MISC ($TOL_{\text{MISC}} = 5.10^{-1}$)	**0.0015** (0.0194)	**0.0108** (0.1392)	**0.0035** (0.0447−)	**0.0015** (0.0194)
MISC ($TOL_{\text{MISC}} = 10^{-1}$)	**0.0015** (0.0194)	**4.4e − 04** (0.0057)	**0.0020** (0.0255)	**2.9e − 04** (0.0037)
MISC ($TOL_{\text{MISC}} = 5.10^{-2}$)	**0.0015** (0.0194)	**6.2e − 04** (0.0080)	**0.0019** (0.0249)	– (–)
MISC ($TOL_{\text{MISC}} = 10^{-2}$)	**0.0027** (0.0345)	**6.0e − 04** (0.0077)	**0.0019** (0.0249)	– (–)

Table 7.23: Quadrature error of MISC, with different tolerances, to compute Basket (2-dimensional) call option price for different number of time steps, without Richardson extrapolation. The numbers between parentheses are the corresponding absolute errors. The values marked in red correspond to stable quadrature errors for MISC, and will be used for complexity comparison against MC.

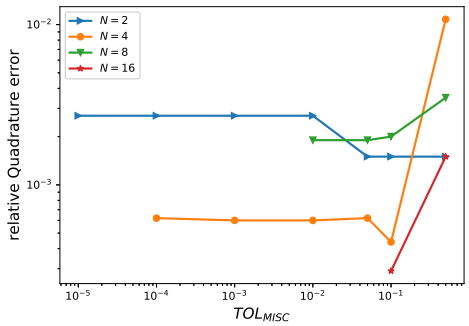


Figure 7.13: Relative quadrature error of MISC,with different tolerances, to compute Basket (2-dimensional) call option price for different number of time steps, without Richardson extrapolation.

Method \Steps	2	4	8	16
MISC ($TOL_{\text{MISC}} = 5.10^{-1}$)	**0.0108**	**0.0183**	**0.0058**	**0.0029**
MISC ($TOL_{\text{MISC}} = 10^{-1}$)	**0.0108**	**0.0079**	**0.0043**	**0.0017**
MISC ($TOL_{\text{MISC}} = 5.10^{-2}$)	**0.0108**	**0.0081**	**0.0042**	–
MISC ($TOL_{\text{MISC}} = 10^{-2}$)	**0.0120**	**0.0081**	**0.0042**	–
MC +root finding	**0.0120**	**0.0081**	**0.0042**	**0.0029**
MC	**0.0119**	**0.0081**	**0.0042**	**0.0029**

Table 7.24: Total relative error of MISC, with different tolerances, and MC to compute Basket (2-dimensional) call option price for different number of time steps, without Richardson extrapolation. The values marked in red, for MISC method, correspond to the total relative errors associated with stable quadrature errors for MISC, and will be used for complexity comparison against MC.

Method \Steps	2	4	8	16
MISC ($TOL_{\text{MISC}} = 5.10^{-1}$)	3	15	197	1048
MISC ($TOL_{\text{MISC}} = 10^{-1}$)	3	34	217	7324
MISC ($TOL_{\text{MISC}} = 5.10^{-2}$)	3	44	917	–
MISC ($TOL_{\text{MISC}} = 10^{-2}$)	4	98	1825	–
MC method +root finding	557	13564	1218	1749
MC method	164	3518	451	1112
Ratio of (MC+root finding)/(MISC)	139	308	5.6	1.7
Ratio of (MC)/(MISC)	41	80	2.1	1.1

Table 7.25: Comparison of the computational time of MC and MISC, used to compute Basket (2-dimensional) call option price for different number of time steps, without Richardson extrapolation. The average computational time of MC is computed over 10 runs.

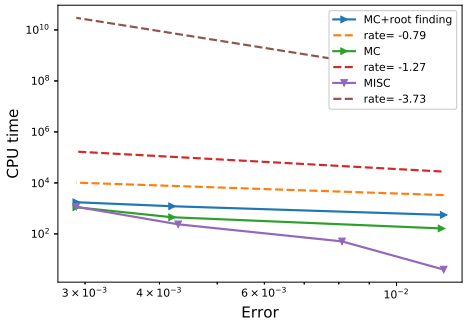


Figure 7.14: Complexity plot for MC and MISC for the case without Richardson extrapolation.

7.7 The basket option with the smoothing trick as in [2]

The third experiment that we consider is the pricing of a European basket call option in a Black-Scholes model. The basket is composed of d assets ($d = 3, 8, 25$) and we use the same trick of smoothing the integrand that was proposed in [2]. In this case, the dimension of the parameter space $N = d-1$. The interpolation over the parameter space is based on the tensorized Lagrangian interpolation technique with Gaussian points. The aim of this section is to check the performance of MISC without time stepping. The main goal is to extend this case to the time stepping framework in the next section 7.2

7.7.1 Results using MISC

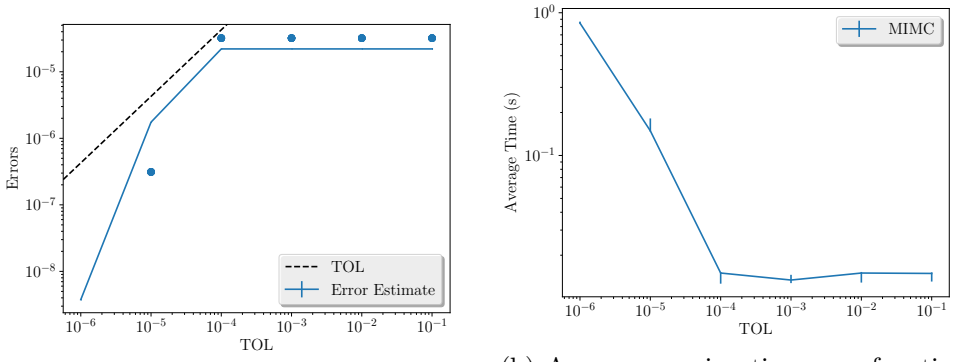
In table 7.26, we summarize the observed complexity rates for different tested settings for the basket example. From this table, we can check that even with the 25 dimensional case, the complexity rate in terms of the elapsed time is at least order 1, which is better than MC, which is 2. Detailed plots for each case are given by figures (7.15, 7.16) for $d = 3$, figures (7.17, 7.18) for $d = 8$ and figures (7.19, 7.20) for $d = 25$. Mainly, from the plots, we checked that we achive the prescribed

tolerance using MISC, the convergence rates of mixed differences which is a basic assumption for using MISC (we observe exponential decay of error rates wrt to the number of quadrature points) and finally the complexity rates. In the next Section, we try to extend these results to the time stepping framework.

# assets \	3	8	25
rate	$-1/3$	$-9/20$	$-16/25$

Table 7.26: Complexity rates of the different experiments for the basket option using BS model

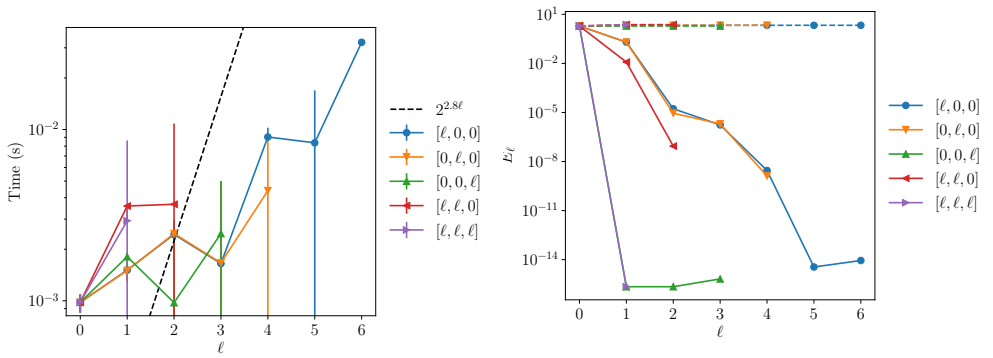
Case of 3-dimensional Basket



(a) Error estimate

(b) Average running time as a function of TOL

Figure 7.15: Convergence and complexity results for the 3-dimensional basket option using BS model.

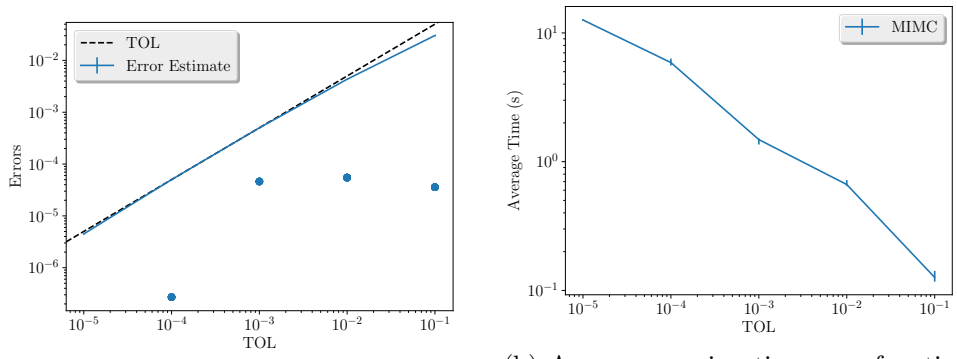


(a) Average Computational time per level.

(b) The convergence rate of mixed differences per level.

Figure 7.16: Convergence and work rates for discretization levels for the 3-dimensional basket option using BS model.

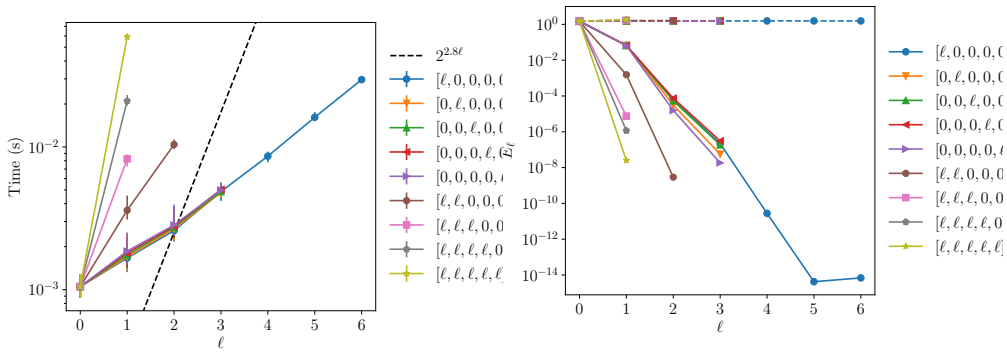
Case of 8-dimensional Basket



(a) Error estimate

(b) Average running time as a function of TOL

Figure 7.17: Convergence and complexity results for the 8-dimensional basket option using BS model.

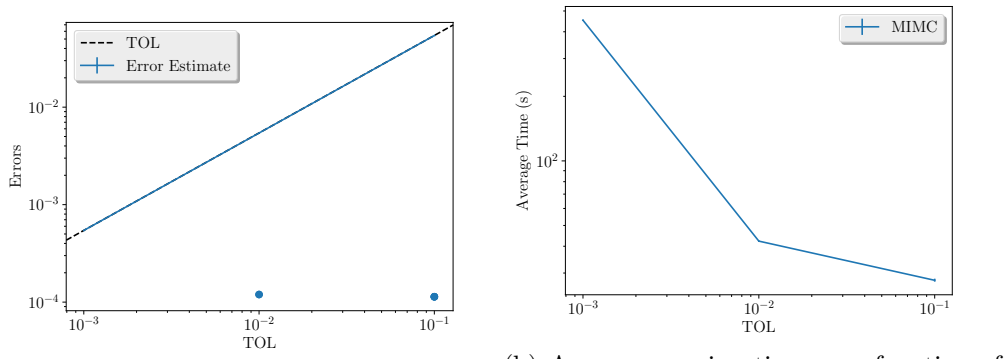


(a) Average Computational time per level.

(b) The convergence rate of mixed differences per level.

Figure 7.18: Convergence and work rates for discretization levels for the 8-dimensional basket option using BS model.

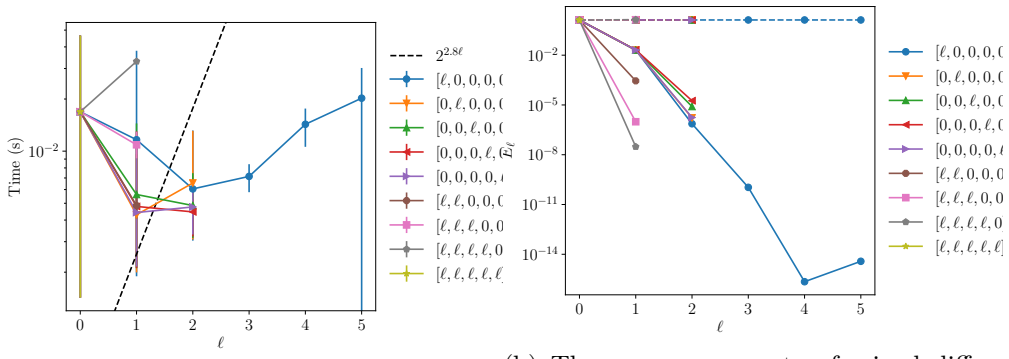
Case of 25-dimensional Basket



(a) Error estimate

(b) Average running time as a function of TOL

Figure 7.19: Convergence and complexity results for the 25-dimensional basket option using BS model.



(a) Average Computational time per level.

(b) The convergence rate of mixed differences per level.

Figure 7.20: Convergence and work rates for discretization levels for the 25-dimensional basket option using BS model.

References Cited

[1] Peter A Acworth, Mark Broadie, and Paul Glasserman. A comparison of some monte carlo and quasi monte carlo techniques for option pricing. In *Monte Carlo and Quasi-Monte Carlo Methods 1996*, pages 1–18. Springer, 1998.

[2] CHRISTIAN BAYER, MARKUS SIEBENMORGEN, and RAUL TEMPONE. Smoothing the payoff for efficient computation of basket option pricing.

[3] Hans-Joachim Bungartz and Michael Griebel. Sparse grids. *Acta numerica*, 13:147–269, 2004.

[4] Russel E Caflisch, William J Morokoff, and Art B Owen. *Valuation of mortgage backed securities using Brownian bridges to reduce effective dimension*. 1997.

[5] Thomas Gerstner. Sparse grid quadrature methods for computational finance.

[6] Thomas Gerstner and Markus Holtz. Valuation of performance-dependent options. *Applied Mathematical Finance*, 15(1):1–20, 2008.

[7] Paul Glasserman. *Monte Carlo methods in financial engineering*. Springer, New York, 2004.

[8] Michael Griebel, Frances Kuo, and Ian Sloan. The smoothing effect of integration in ˆ{} and the anova decomposition. *Mathematics of Computation*, 82(281):383–400, 2013.

[9] Michael Griebel, Frances Kuo, and Ian Sloan. Note on the smoothing effect of integration in ˆ{} and the anova decomposition. *Mathematics of Computation*, 86(306):1847–1854, 2017.

[10] Andreas Griewank, Frances Y Kuo, Hernan Leövey, and Ian H Sloan. High dimensional integration of kinks and jumps–smoothing by preintegration. *arXiv preprint arXiv:1712.00920*, 2017.

[11] Abdul-Lateef Haji-Ali, Fabio Nobile, Lorenzo Tamellini, and Raul Tempone. Multi-index stochastic collocation for random pdes. *Computer Methods in Applied Mechanics and Engineering*, 306:95–122, 2016.

[12] Gerhard Larcher, Gunther Leobacher, and Klaus Scheicher. On the tractability of the brownian bridge algorithm. *Journal of Complexity*, 19(4):511–528, 2003.

[13] William J Morokoff. Generating quasi-random paths for stochastic processes. *Siam Review*, 40(4):765–788, 1998.

[14] Bradley Moskowitz and Russel E Caflisch. Smoothness and dimension reduction in quasi-monte carlo methods. *Mathematical and Computer Modelling*, 23(8):37–54, 1996.

[15] Denis Talay and Luciano Tubaro. Expansion of the global error for numerical schemes solving stochastic differential equations. *Stochastic analysis and applications*, 8(4):483–509, 1990.

[16] Ye Xiao and Xiaoqun Wang. Conditional quasi-monte carlo methods and dimension reduction for option pricing and hedging with discontinuous functions. *Journal of Computational and Applied Mathematics*, 343:289–308, 2018.