Variance Reduction for Monte Carlo Methods

DUONG Rayhaan

August 2025

Contents

1 Introduction

1.1 Abstract

In this paper, we aim to price an arithmetic basket option using Monte Carlo methods, since no closed-form formula is available for this product. The efficiency of a Monte Carlo estimator is directly related to the variance of the underlying random variables. We therefore investigate several variance reduction techniques and analyze their sensitivity to the initial parameters of the problem. We begin with the naive Monte Carlo estimator (MC), then introduce the control variate approach based on a geometric basket option (VC). Afterwards, we present a new control variate method proposed by Curran [1](NVC), as well as a hybrid approach that combines Curran's method with the use of conditional expectations [2] (NVC C). Eventually, we will experiment with real market data.

1.2 Objectives

We aim to analyze how the effectiveness of different methods evolves as the problem parameters change. Effectiveness is assessed using two indicators: the first is a variance ratio, and the second—more interestingly—is the effort ratio. After discussing the theoretical framework, we will implement the approach on real market data.

1.3 Key results

Courte synthèse qui introduit la suite du document.

2 Theory

2.1 Naive Monte Carlo (MV)

The standard Monte Carlo method is often used in numerical applications when it is necessary to compute an expectation for which no easily computable closed-form solution exists. We use the following Monte Carlo (MC) estimator to calculate $I = \mathbb{E}[f(X)]$ where $f(X)$ is in L^1 (L^2 if we want to use the Central Limit Theorem (CLT)).

$$\frac{1}{N}\sum_{i=1}^{N} f(X_i) \xrightarrow{\text{a.s.}} I$$

Note that the convergence is guaranteed by the Strong Law of Large Numbers (SLLN). What matters a lot is to have a small asymptotic variance, because by relying on the CLT, we know that, for N large and a given $\alpha \in]0,1[$:

$$\mathbb{P}(\overline{X}_N - q_{1-\alpha/2}\frac{\sigma_N}{\sqrt{N}} < I < \overline{X}_N + q_{1-\alpha/2}\frac{\sigma_N}{\sqrt{N}}) \approx 1-\alpha$$

and we have, inside the probability, the following inequality :

$$\mid X_N - I \mid \leq q_{1-\alpha/2}\frac{\sigma_N}{\sqrt{N}}$$

We use this standard method as a baseline to assess other variance reduction techniques. In other words, all indicator ratios are computed with respect to this estimator.

2.2 Variate Control using a geometric basket option (VC)

Rather than detailing the theory rigorously, we will focus on explaining the core ideas.

A few words on the control variate method: it works based on the following principle :

$$\mathbb{E}[f(X)] = \mathbb{E}[f(X) - c(Y - \mathbb{E}[Y])]$$

We choose Y such that :

- Var(f(X) - cY) < Var(f(X))
- It exists a closed formula for $\mathbb{E}[Y]$
- Sampling from f(X) is equivalent to sampling from f(X) - cY

The reason we choose the geometric basket option as a control variate is that the arithmetic basket option (A) is highly correlated with the geometric basket option (G) in most, but not all, cases. We want correlated variables. The worst-case scenario is when the variables are independent, as their variances simply add up.

2.3 New Variate Control (Curran, 1994) (NVC)

A closed-form formula exists for the geometric basket option

$$G = e^{\sum_{i=1}^{n} w_i \ln(S_i^T)},$$

whose payoff is $(G - K)^+$. Using it as a control variate reduces the variance when estimating the arithmetic basket option $(A - K)^+$. The intuition behind this choice relies on the fact that when $G > K$, we also have $A > K$ (due to convexity). Therefore, whenever $G > K$, the random variable and the control variate take exactly the same value, implying a high covariance that effectively reduces the variance; furthermore, we can compute a closed formula of its expectation. Empirically, the optimal coefficient c is very close to 1 in this case, so we simply take $c = 1$.

The resulting control variate estimator is

$$(A-K)^+ - (A-K)^+\mathbf{1}_{\{G>K\}} + \mathbb{E}\big[(A-K)^+\mathbf{1}_{\{G>K\}}\big] = \mathbb{E}\big[(A-K)\mathbf{1}_{\{G>K\}}\big] + (A-K)^+\mathbf{1}_{\{G<K\}}.$$

and therefore the basket option price is

$$V = e^{-rT}(\xi_1 + \xi_2),$$

with

$$\xi_1 = \mathbb{E}[(A-K)^+ 1_{\{G>K\}}], \quad \xi_2 = \mathbb{E}[(A-K)^+ 1_{\{G<K\}}].$$

We have a closed formula to compute ξ_1, and we use the standard Monte Carlo to estimate ξ_2.

2.4 Hybrid method based on NVC and conditional expectation (NVC C)

We can write :

$$\xi_2 = \mathbb{E}[\mathbb{E}[(A-K)^+ 1_{\{G<K\}} \mid \tilde{Y}]]$$

- A d-dimensional Brownian motion W can be written as $W = LY$, where L comes from the Cholesky decomposition of the covariance matrix, and Y is a standard d-dimensional Gaussian vector.
- $\tilde{Y}$ stands for Y without the last component ($d-1$ dimensional)

This method should reduce the variance as :

$$g\big(\mathbb{E}[X \mid \mathcal{B}]\big) \leq \mathbb{E}\big[g(X) \mid \mathcal{B}\big]$$

By taking g as the square function and then taking the expectation of the inequality, we obtain the desired result.

Knowing $\tilde{Y}$, we can calculate a closed formula of $\mathbb{E}[(A-K)1_{\{G<K<A\}} \mid \tilde{Y}]$. The main trick is to write, G < K as $y_n < k_u(\tilde{Y})$ and K < A as $y_n > kl(\tilde{Y})$. The definitions of $k_u(\tilde{Y})$ and $k_l(\tilde{Y})$ are given in [2] and we get :

$$\xi_2 = \mathbb{E}\big[\sum_{i=1}^{d} w_i S_{i0} e^{(r-\frac{\sigma_i^2}{2})T}\,\mathbb{E}\big[e^{\sigma_i W_i(T)} \mathbf{1}_{\{k_l(\tilde{Y})<K<k_u(\tilde{Y})\}} \mid \tilde{Y}\big] - K\,\mathbb{E}\big[1_{\{k_l(\tilde{Y})<K<k_u(\tilde{Y})\}} \mid \tilde{Y}\big]\big]$$

We can write :

$$\xi_2 = \mathbb{E}\big[z(\tilde{Y})\big]$$

and we estimate this quantity by sampling from the distribution of $\tilde{Y}$ and applying the Monte Carlo process.

3 Comparison of estimators

In this section, we compare the four estimators, varying the initial parameters in order to gauge the effectiveness of each method with respect to different parameter values.

We use two different indicators. The first one is the asymptotic variance, which shows whether the methods effectively reduce the variance of our Monte Carlo procedure. A second indicator is needed to assess the overall effectiveness of the methods, since variance reduction may come at the cost of high computational time. For this reason, we introduce the effort, defined as the product of the variance and a measure of sampling complexity — in our case, the time required to run the Monte Carlo method.

We consider a basket option with the following initial parameters:

$$
\begin{array}{ll}
N = 100{,}000 & \text{(number of Monte Carlo simulations)} \\
d = 4 & \text{(number of assets)} \\
S_0 = (100,\ 95,\ 105,\ 110) & \text{(initial asset prices)} \\
\sigma = (0.20,\ 0.25,\ 0.18,\ 0.22) & \text{(volatilities)} \\
r = 0.05 & \text{(risk-free rate)} \\
T = 2 & \text{(maturity)} \\
K = 100 & \text{(strike)} \\
w = (0.25,\ 0.25,\ 0.25,\ 0.25) & \text{(weights of the basket)} \\
R = \begin{bmatrix} 1.00 & 0.80 & 0.60 & 0.40 \\ 0.80 & 1.00 & 0.65 & 0.75 \\ 0.60 & 0.65 & 1.00 & 0.50 \\ 0.40 & 0.75 & 0.50 & 1.00 \end{bmatrix} & \text{(correlation matrix)} \\
\text{Confidence level:}\quad 1 - \alpha = 0.95 & \text{(probability that the option price lies within the CI).}
\end{array}
$$

3.1 The Strike (K)

In this subsection, we are shifting the Strike range from 100 to 270, with a step size of 10. (Only some representative values of K are displayed.

Table 1: Variance ratio of MV/VC MV/NVC and MV/NVC C for different K values

K	MV/VC	MV/NVC	MV/NVC C
100	2.18e+02	7.35e+03	5.12e+04
150	7.58e+01	1.31e+03	1.10e+04
210	2.90e+01	3.70e+02	1.86e+03
270	1.02e+01	6.00e+01	4.44e+02

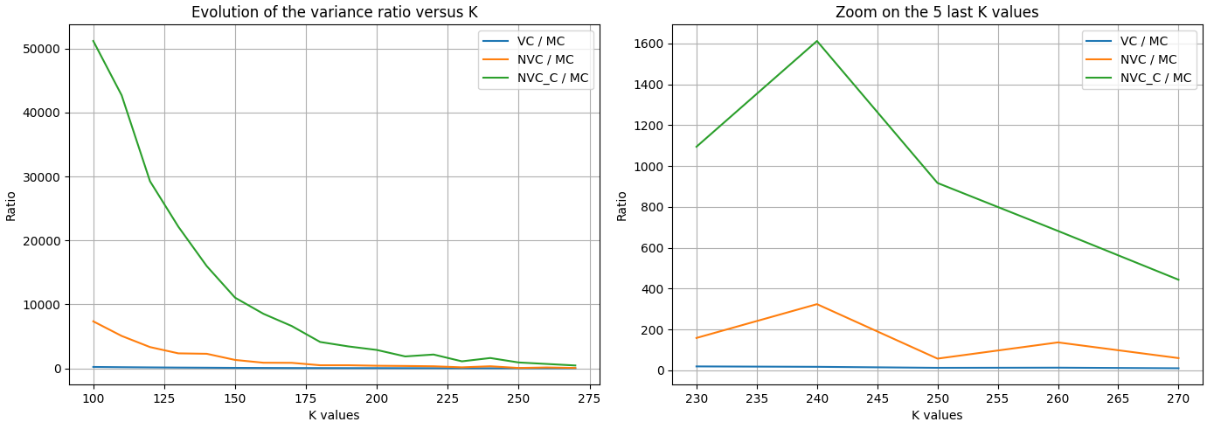

Figure 1: Evolution of the Variance Ratio for K from 100 to 270

Table 2: Effort ratio of MV/VC MV/NVC and MV/NVC C for different K values

K	MV/VC	MV/NVC	MV/NVC C
100	3.43e+01	1.10e+03	7.10e+02
150	1.00e+01	1.79e+02	1.49e+02
210	3.84e+00	4.89e+01	2.41e+01
270	1.16e+00	6.08e+00	4.49e+00

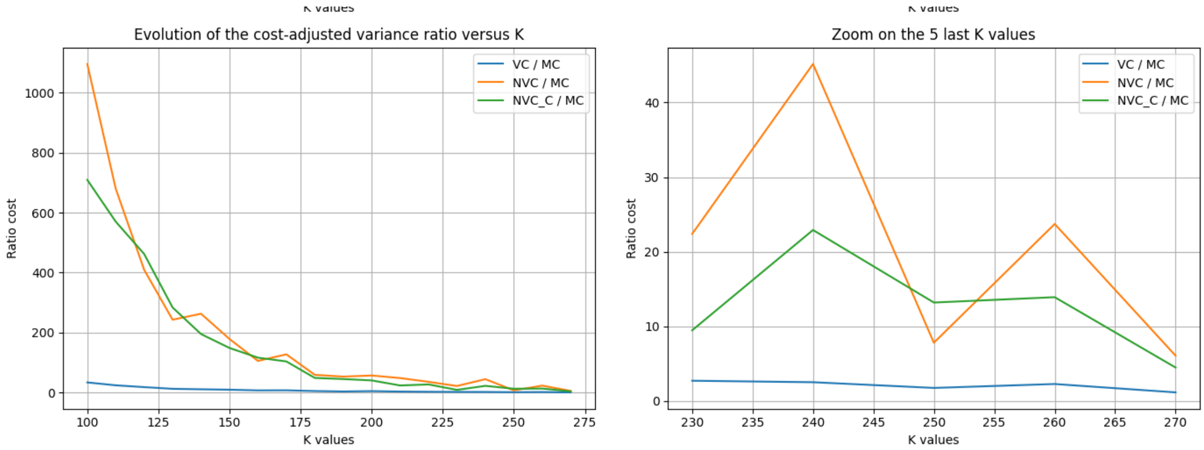

Figure 2: Evolution of the Effort Ratio for K from 100 to 270

From the empirical results, the variance ratio seems to decline as K increases across the three methods. It seems that as the strike K increases, the probability of a positive payoff decreases, which could lead to a smaller variance of the payoff. This might explain why the variance reduction methods appear less effective for higher strikes.

Based on the first indicator, one might think that NVC C is the best estimator. However, it is important to also consider the complexity involved in generating the variables and computing the intermediate functions needed for the closed-form formulas. The effort ratio takes this aspect into account, and we observe that the NVC and NVC C estimators are equivalent under this indicator. This is due to the fact that in the NVC C method, the computation time increases significantly, as the formula for $\tilde{h}(y)$ is quite complex, not to mention that it requires calling norm.cdf($k_u(\tilde{Y})$) and norm.cdf($k_l(\tilde{Y})$) at each step.

We could consider ways to reduce the computation time required for this last method. This would ensure that it becomes the most efficient variance reduction method among those studied.

3.2 The Time horizon (T)

In this subsection, we are shifting the Time horizon range from 0.25 to 4 (years), with a step size of 0.25. (Only some representative values of T are displayed.

Table 3: Variance ratio of MV/VC MV/NVC and MV/NVC C for different T values

T	MV/VC	MV/NVC	MV/NVC C
0.25	8.29e+02	3.10e+04	4.91e+05
1.25	3.20e+02	1.11e+04	9.74e+04
2.50	1.82e+02	6.12e+03	4.23e+04
4.00	1.75e+02	4.03e+03	2.40e+04

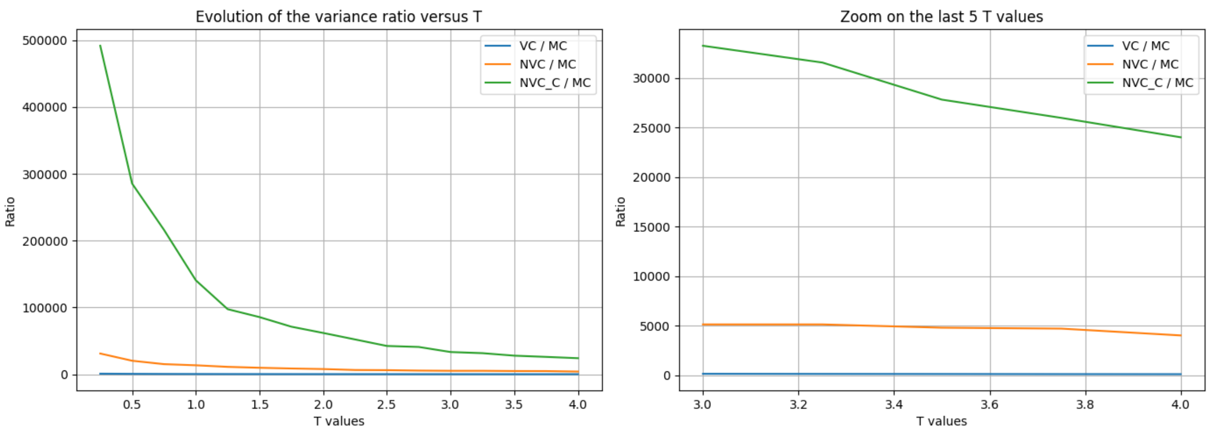

Figure 3: Evolution of the Variance Ratio for T from 0.25 to 4

Table 4: Effort ratio of MV/VC MV/NVC and MV/NVC C for different T values

T	MV/VC	MV/NVC	MV/NVC C
0.25	1.89e+02	7.33e+03	1.30e+04
1.25	4.51e+01	1.50e+03	1.40e+03
2.50	2.36e+01	7.90e+02	6.10e+02
4.00	1.51e+01	4.90e+02	3.43e+02

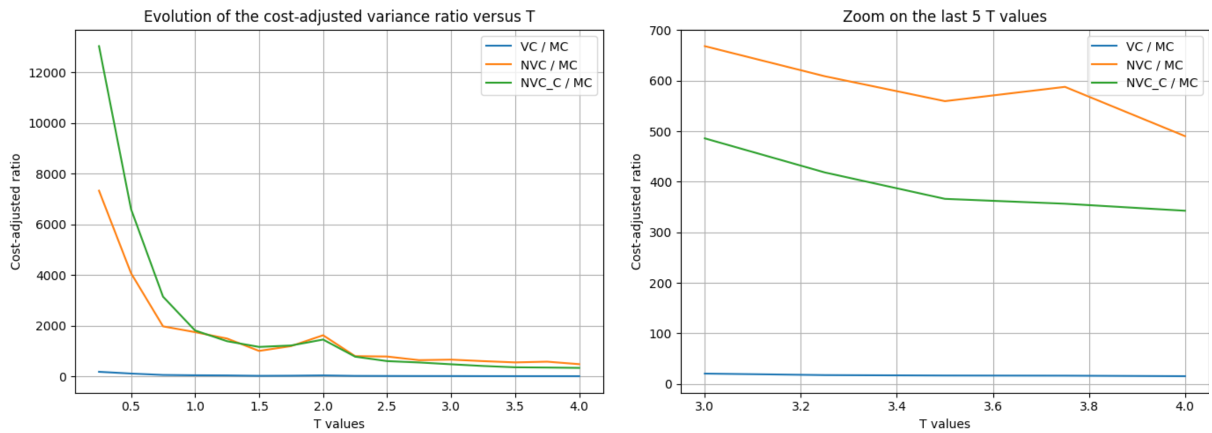

Figure 4: Evolution of the Effort Ratio for T from 0.25 to 4

Empirically, we observe that the efficiency of the variance reduction methods tends to decrease as the time horizon T increases. This can be explained by the fact that the payoff becomes more dispersed for longer maturities, which reduces the correlation with the control variables and, consequently, the relative variance reduction.

Once again, we observe that the effort ratios of MV/NVC and MV/NVC C are similar, even though MV/NVC C achieves greater variance reduction. This is explained by the increased computational time required for MV/NVC C.

It is rather difficult to conduct a study where we vary the correlation matrix and the vector of standard deviations, as this is a high-dimensional problem. However, it is not too hard to observe that when the variance is high, variance reduction methods become less effective. Therefore, increasing the variance of each asset will generally reduce the efficiency of the studied methods.

We can also vary the number of assets, examine how changing the correlation between the assets affect the results. We will examine these effects in the next section, when applying the methods to real data.

We will not study the case where the risk-free rate r is varied, as it is unlikely to deviate significantly from 0.3–0.5.

4 Experimental Study

4.1 Context Description

In this study, we analyze the performance of variance reduction methods applied to a diverse set of financial assets. The assets under consideration include large-cap equities (AAPL, MSFT, TSLA, NVDA), small to mid-cap stocks (SHOP.TO, TWLO, ETSY), a sector-specific ETF (ARKK), and commodity futures (GC=F for gold and NG=F for natural gas). The analysis covers historical price data over the period from January 1, 2018, to December 31, 2023. This selection allows us to observe how variance reduction techniques behave across different asset classes and market conditions.

The data are retrieved from the Yahoo Finance API and saved in an Excel sheet. When prices are missing for some trading days—for example, because SHOP.TO is listed on the Toronto Stock Exchange, which does not always operate on the same days as U.S. exchanges—we fill the missing values using linear interpolation.

Number of missing trading days per asset:

- AAPL: 28
- MSFT: 28
- GAS: 27
- NVDA: 28
- SHOP: 32
- TWLO: 28

Total number of trading days: 2188, the other values have 0 missing days.

Table 5: Sample of daily closing prices first 5 trading days.

Date	AAPL	ARKK	ETSY	GOLD	MSFT	GAS	NVDA	SHOP	TSLA	TWLO
2018-01-02	40.38	36.05	20.83	1313.70	79.33	3.06	4.93	13.21	21.37	25.10
2018-01-03	40.37	36.20	20.41	1316.20	79.70	3.01	5.25	13.52	21.15	25.69
2018-01-04	40.56	36.32	20.23	1319.40	80.40	2.88	5.28	13.75	20.97	25.68
2018-01-05	41.02	36.91	20.25	1320.30	81.40	2.80	5.33	13.73	21.11	25.76
2018-01-08	40.87	36.66	20.32	1318.60	81.48	2.84	5.49	13.95	22.43	26.17

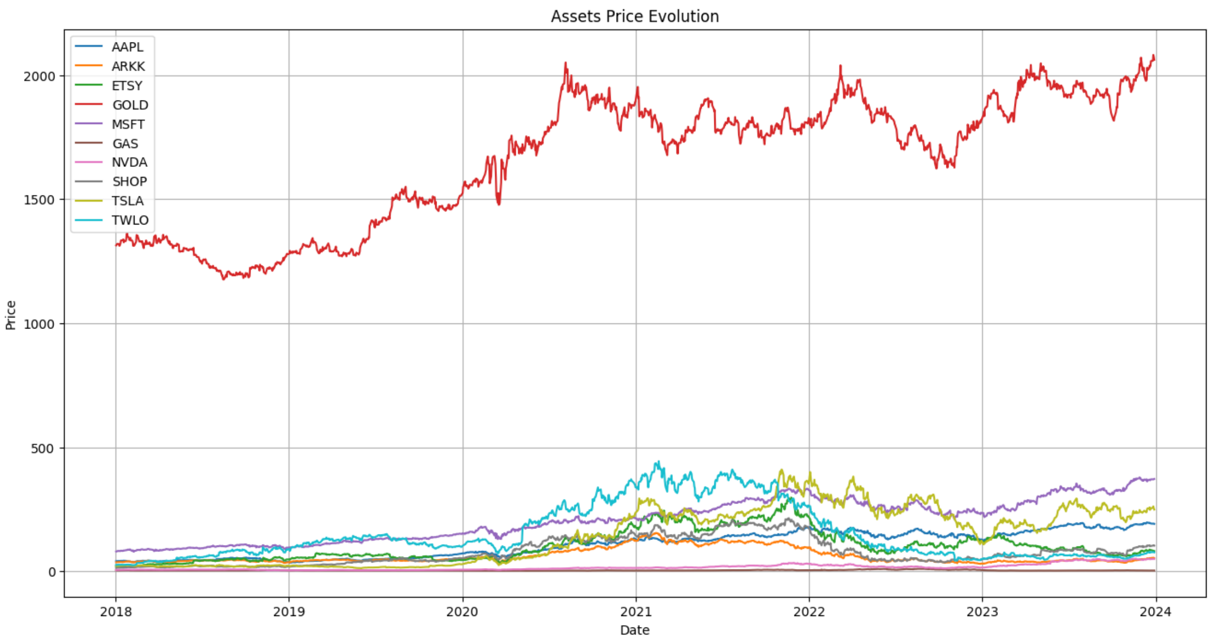

Figure 5: Evolution of the Prices from January 2018 until December 2023.

Prices are not directly comparable across assets, as their scales differ. That's why we are working with the log-returns.

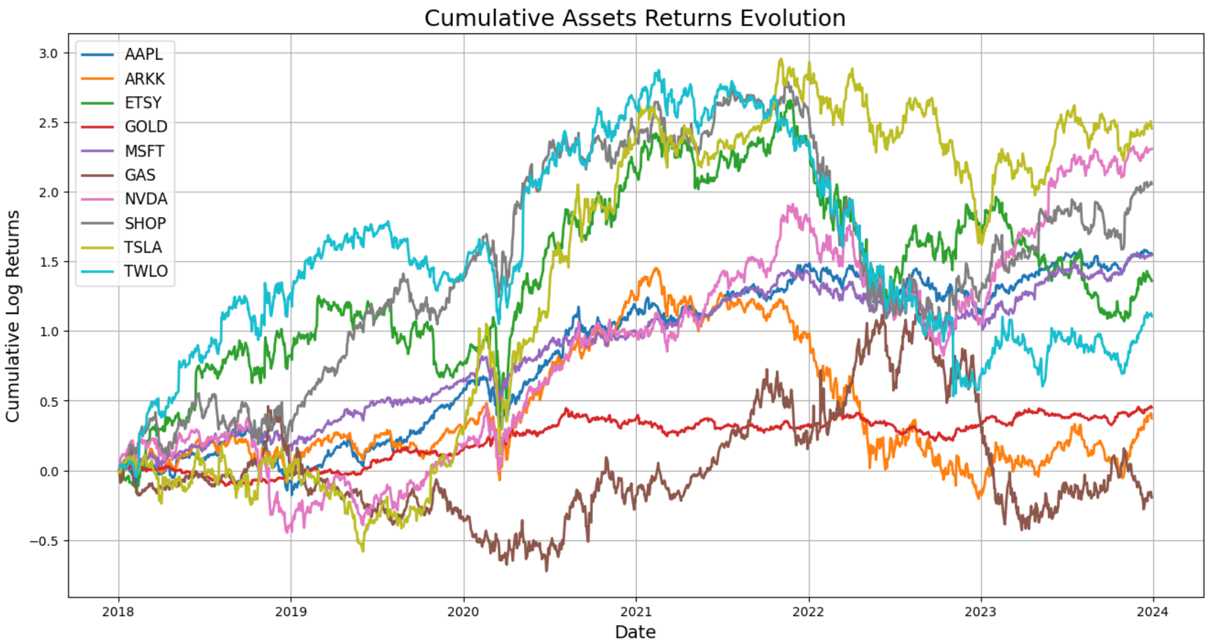

Figure 6: Evolution of the Cumulative assets return evolution from January 2018 until December 2023.

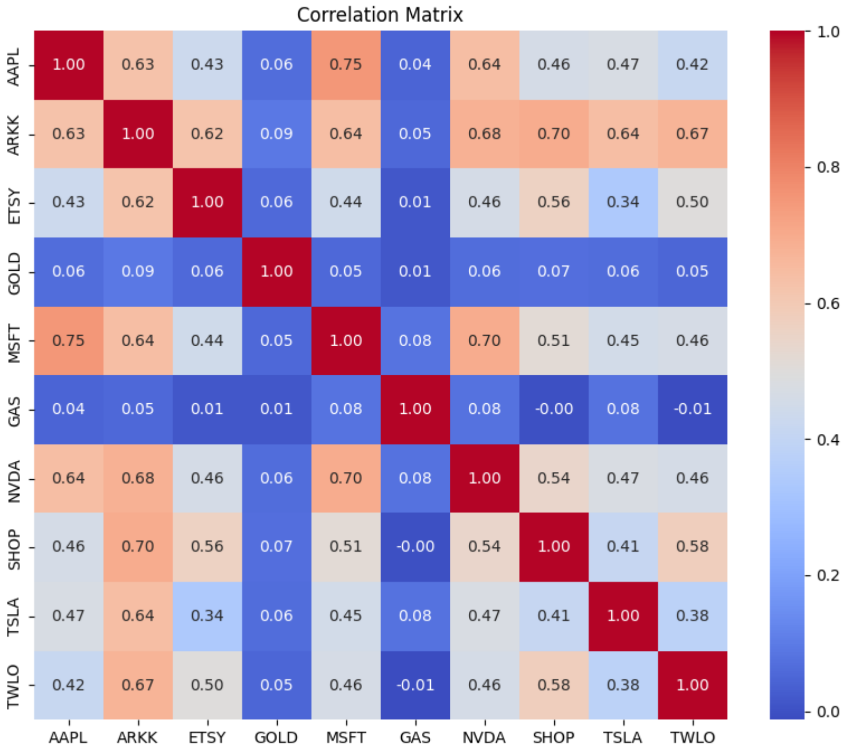

Figure 7: Correlation Matrix.

We consider the following base parameters for the simulation:

$$N = 10000, \quad r = 0.05, \quad T = 1, \quad K = 100.$$

Note that the strike K is likely to change depending on the experiments, and this will be specified when necessary. The correlation matrix used is the one provided above, and if the number of assets changes in a given test, this will also be explicitly mentioned.

4.2 Experimentation

We run our algorithm with the parameters mentioned above. Here are the results:

Estimator	Variance Ratio	Effort Ratio
MC / VC	1.80e+00	1.34e-01
MC / NVC	2.92e-01	2.10e-02
MC / NVC_C	3.45e-01	6.60e-03

Table 6: Variance and effort ratios for different estimators. Strike is set to $K = 100$.

We did not at all expect these preliminary results: the efficiency of the estimators, for both indicators, decreases, whereas we had anticipated the opposite. In fact, this can be easily explained by the fact that the prices of the different assets vary significantly, causing A and G to take drastically different values in terms of magnitude.

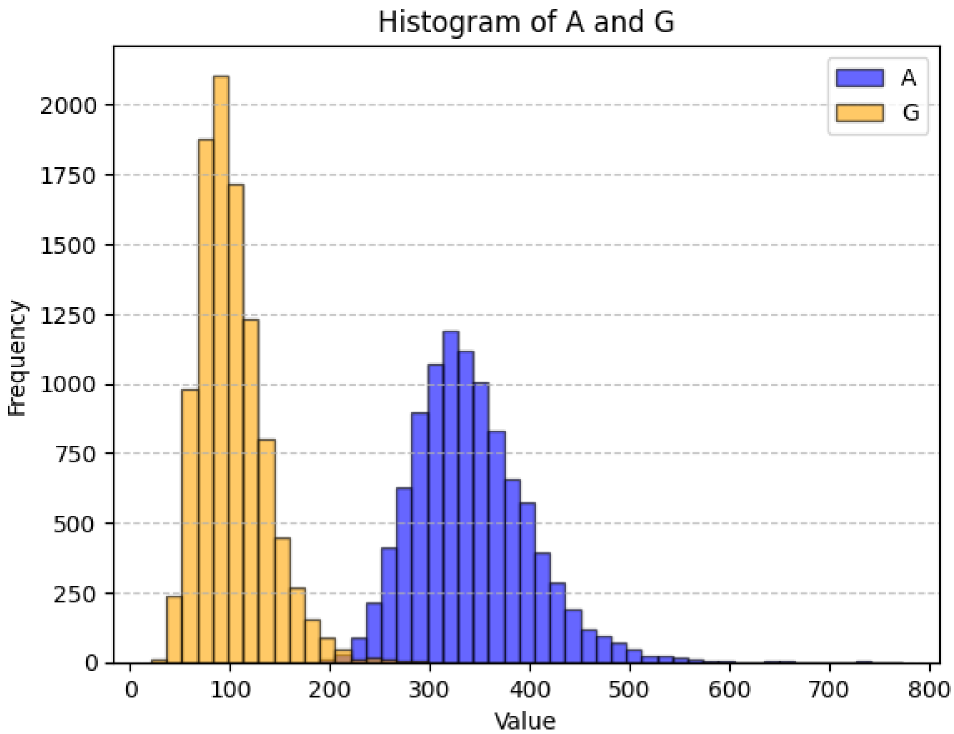

Figure 8: A skew : 0.83, G skew : 1.06, A kurt : 1.73, G kurt : 2.10.

A and G have roughly the same shape but take different values, so the intersection of their supports is almost empty. When we take $K = 100$, A is almost always above K, whereas G is above K roughly half the time. Recall that the control variate based on the geometric basket is $(G - K)^+$ and Curran's control variate is $(A - K)^+ \mathbf{1}_{\{G>K\}}$. As a result, these control variates are no longer strongly correlated with the underlying variable, which explains why these methods are not very effective. This difference in levels arises because, even if A and G are often close, with a basket composed of assets with very different prices, the logarithm flattens the large values and G grows more slowly than A.

We will verify this assertion by computing with a lowered K, it should give us better results.

Estimator	Variance Ratio	Effort Ratio
MC / VC	2.53e+00	1.84e-01
MC / NVC	2.18e+01	1.63e+00
MC / NVC_C	2.93e+01	5.50e-01

Table 7: Variance and effort ratios for different estimators. Strike is set to $K = 40$.

Our results are more consistent with the theoretical expectations, although not fully efficient, since A and G have almost disjoint supports.

We can attempt to remove some assets that drive these differences in order to observe how the estimators behave in a more balanced setting. In this part, we will remove GOLD, GAS, ARKK and NVDA. GOLD has by far the highest price level among the assets, while the three others exhibit the lowest prices.

Estimator	Variance Ratio	Effort Ratio
MC / VC	1.07e+01	9.73e-01
MC / NVC	1.63e+04	1.14e+03
MC / NVC_C	2.16e+04	3.48e+02

Table 8: Variance and effort ratios for different estimators, without 4 assets.

Our results coincide more with the theoretical expectations: by removing these four assets, G becomes much closer to A, leading to improved efficiency of the control variate methods.

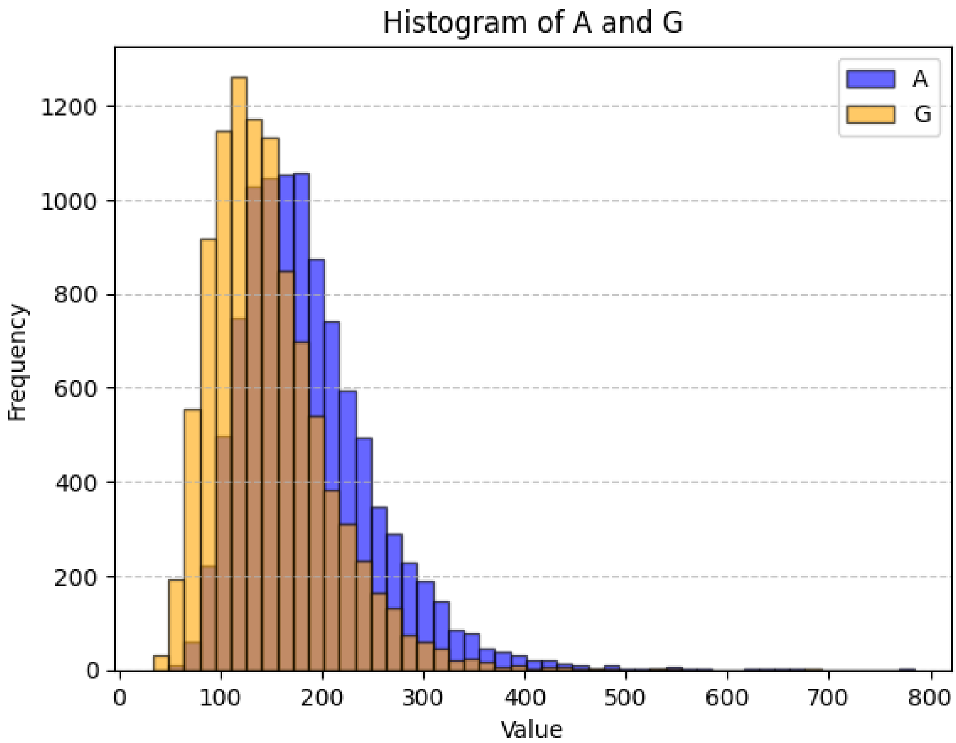

Figure 9: A skew : 1.31, G skew : 1.29, A kurt : 3.50, G kurt : 3.37.

4.3 Discussion of results

Working with a real dataset allows us to better understand the limitations of these methods. While their effectiveness can be remarkable under ideal theoretical conditions—we obtained a variance ratio as high as 5.12×10^4 and an effort ratio up to 1.10×10^3—the performance deteriorates with real data. This is mainly because baskets often include assets with very different prices, causing G and A to lie on different scales. As a result, the control variates become less correlated with the payoff and therefore less effective.

These findings suggest that the success of control variates crucially depends on the structure of the dataset. In particular, when working with a more homogeneous set of assets (in terms of prices), both NVC and NVC C become highly effective. Although NVC C achieves stronger variance reduction, the two approaches are essentially equivalent in terms of computational effort.

5 Conclusion

To improve the current work, we could explore better ways to compute the deterministic values required for NVC C. Reducing its computation time could make this estimator the most efficient among those studied. Additionally, we could investigate which types of control variates would effectively reduce the variance of a Monte Carlo procedure when asset prices are very heterogeneous, or when the strike K is out of the money.

References

[1] M. Curran, *Valuing Asian Options by Conditioning on the Geometric Average*, Management Science, 1994, 40, 1705-1711.

[2] Yongchao Sun and Chenglong Xu, *A hybrid Monte Carlo acceleration method of pricing basket options based on splitting*, Journal of Computational and Applied Mathematics, 2018, 342, 292-304.